The Epistemic Condition

Jan Willem Wieland

This introduction provides an overview of the current state of the debate on *the epistemic condition of moral responsibility*. Its main goal is to offer a framework that contextualizes the chapters that follow. In §1, we discuss the main concepts 'ignorance' and 'responsibility'. In §2, we ask why agents should inform themselves. In §3, we describe what we take to be the core agreement among main participants in the debate. In §4, we explain how this agreement invites a regress argument with a revisionist implication. In §5, we provide an overview of the main responses to the regress argument. In §6, we address the question why blameless ignorance excuses. In §7, we describe further issues that are addressed in the book. In §8, we conclude with some discussion of future directions the debate might take.

1. Introduction

When we act, we are often ignorant of some relevant fact. The question of this book is: are ignorant agents morally responsible for their conduct? Consider the following cases.[1]

> *Director 1.* S is the director of a pharmaceutical company. She knows that they could develop two different treatments for a certain disease, and votes for the one which, unknown to her, has severe side-effects.
>
> *Director 2.* S is the director of a pharmaceutical company. She knows that the drug they currently offer on the market has severe side-effects, but still endorses their product because she is ignorant of better treatments.
>
> *Director 3.* S is the director of a pharmaceutical company. She knows that the drug they currently offer on the market has severe side-effects, and that they could offer a new alternative drug with fewer side-effects, but declines to do so

---

[1] In the following, 'S' stands for an agent; 'X' for anything S can be responsible for; and further abbreviations will be introduced in due course.

because the current one will make more profit, and she is ignorant that this is wrong.

Slaveholder 1. S keeps slaves. She knows she could let them go, yet she does not do this because she hardly pays any attention to them, and is ignorant that they suffer.

Slaveholder 2. S keeps slaves. This time, she knows they suffer. Still, she does not let them go because she is ignorant that she could run her business without them.

Slaveholder 3. S keeps slaves. This time, she knows they suffer, and that she could run her business without them. Still, she does not do this because she is ignorant that she is doing anything wrong. For all she knows, slavery is a given fact of nature, and if she was unlucky enough, she could have been a slave herself.

In these cases, the agents are ignorant of different things. In Director 1 and Slaveholder 1, they are ignorant of the fact that *their conduct has a certain consequence* (harm to patients and suffering). In Director 2 and Slaveholder 2, they are ignorant of the fact that *they could have acted differently* (working on an alternative drug and running one's business without slavery). In these cases, the agents know that their actions have certain consequences but do not know that they could have avoided acting in that way. Finally, in Director 3 and Slaveholder 3, the agents are ignorant of the fact that *their conduct is wrong*, where this ignorance is not based on any relevant non-moral ignorance. In these cases, the agents know that their conduct has a certain consequence, and that they could have acted differently, but they do not know that their actions are morally wrong. The first two kinds of ignorance are called "factual" ignorance, and the third kind "moral" ignorance. Many real-life cases are mixed cases in which agents are ignorant of facts of both kinds. We present these pure cases in order to clarify the conceptual possibilities.

Another relevant distinction concerns not the facts about which agents can be ignorant, but the variety or nature of their ignorance.[2] A first main variant is *false* or *mistaken belief.* For example, one might be ignorant that one's conduct is wrong because one mistakenly believes that it is permissible. A second main variant is *deep*

---

[2] For different types of ignorance, cf. Guerrero (2007), Van Woudenberg (2009), Peels (2014).

ignorance. Here one is ignorant that one's conduct is wrong because one has just never considered the issue in the first place. Next to these variants, there are also cases of *forgetting*. In forgetting cases, one does know one's conduct is wrong, but at the time of action, one is not consciously aware of this fact. The director and slaveholder cases are plausibly classified as mistaken belief or deep ignorance cases, given that they are all either mistaken about or deeply ignorant of the consequences of their actions, the fact that they could have acted differently, or the wrongness of their actions. We will address forgetting cases later on.

As said, the question of this book is: are ignorant agents morally responsible for their actions? If someone unwittingly does something that is morally wrong, are they blameworthy for their conduct, or are they excused by their ignorance?

One straightforward way of addressing these questions is to check whether the ignorant agents are *reckless* or *negligent*. If they are, they can be considered blameworthy (see Matt King's chapter in this book). One is reckless when one is *aware* that one takes a serious and unjustified risk by failing to inform oneself. For example, a doctor is reckless if she is aware that she takes such a risk by failing to review the latest research on a drug she prescribes to a patient. One is negligent, by contrast, when one is *culpably unaware* that one takes a serious and unjustified risk if one does not inform oneself.[3] If the doctor is unaware that she runs a serious and unjustified risk when she prescribes the drug and is culpable for this lack of awareness, then she is negligent.

To apply these concepts to the cases above, we would need more information. Should the director and slaveholder have informed themselves? Did they try to do so? Were they aware that they could do so and that failing to do so posed an unjustified risk of realizing some important harm? If the agents did not know they should have informed themselves, we would also need to know whether this ignorance is blameworthy. When are directors, slaveholders, doctors, and other ignorant agents, blameworthy for their ignorance? Most of the debate has concerned this question, and the details are trickier than it might seem at first sight.

The debate is commonly known as the debate on *the epistemic condition of responsibility*. In this introduction, we will sketch the main moves. As we will see, one basic idea that many philosophers accept is that one is blameworthy for one's ignorant conduct *only if* one is blameworthy for one's ignorance. For example, a doctor is blameworthy for ignorantly harming her patients only if she is blameworthy for her

---

[3] 'Blameworthiness' and 'culpability' mean the same in this debate.

ignorance about the side-effects of the drug. If it is not her fault that she fails to know this, then it is not her fault she prescribed the drug and harmed the patient.

Of course, the epistemic condition is not the only thing that matters to responsibility. Questions about agency (what kind of agents can be considered responsible in the first place), causation (which states of affairs can be considered consequences of one's conduct), and control (whether one is excused by a lack of control or freedom) are also relevant. In the following, we will abstract from these questions by assuming that we are concerned with agents who possess the relevant capacities, who bring about certain consequences, and who acted freely (in any relevant sense that would matter).

These conditions are commonly understood as conditions for moral responsibility. To say that one is morally responsible for an action is to say, roughly, that one is *blameworthy* if the action is wrong and *praiseworthy* if action is right. In this introduction as well as in the book, we will focus on blameworthiness. As many say since Strawson (1962), S is morally blameworthy for X (whether it is an action, omission, consequence, etc.) iff on account of X, S is a proper candidate for the reactive attitudes, such as resentment, indignation, and condemnation.

The claim that an agent is blameworthy is an evaluative claim that should be kept distinct from closely related kinds of evaluation, such as the *wrongness of the action*, *badness of the agent*, *epistemic irrationality*, and *legal punishability*. Most participants in the debate will agree on this, and we will briefly clarify some contrasts. Blameworthiness differs from wrongness given that the former is an evaluative claim about agents whereas the latter is an evaluative claim about actions. Moreover, most accept that one can be blameless for a wrongful action. For example, one might be excused for it by ignorance.[4] Blameworthiness and badness do not entail each other either, at least so long as there are cases where agents are really bad and vicious, but not blameworthy (and thus not proper candidates for resentment). Moral blameworthiness and epistemic irrationality are distinct simply because they concern different domains. In principle, whether agents are rational and justified in their beliefs depends on epistemic standards, and whether agents are morally blameworthy depends on moral standards. Finally, given that agents may be morally responsible even though they are not legally punishable (and vice versa), moral responsibility is also distinct from

---

[4] Another way in which they come apart is when S is blameworthy despite having performed an action that was not wrong. In such cases, S is said to have an 'accuse', rather than an excuse (Zimmerman 2008: 194).

4

legal punishability. Of course, there might be connections between all these things, but they are not obvious. Defending them is a substantive project.

The link between blameworthiness and legal punishment seems relevant, though is complex. Legal punishment can have various justifications, including deterrence (punishing agents might deter others from committing similar crimes), protection (one protects the society from criminals by locking them away), rehabilitation (punishing criminals might help them to become better moral agents), and retribution (punishing criminals gives them what they deserve).[5] The first three functions are forward-looking and consequentialist. The justification for punishment depends on whether it helps realize a certain change (in the criminal's behaviour, in the behaviour of others, or in society). In contrast, retribution is backward-looking and desert-based: agents are punishable depending on whether they deserve blame for what they have done in the past (and not on whether this would likely lead to any desired change).

The central question is of course what wrongdoers deserve, and here philosophers have something to say. The literature contains defenses of several desert-based, or at least backward-looking, accounts. Two of them – *attributability* and *accountability* – will play an important role in this book.[6] Attributability concerns an evaluation of aspects of S's character as they are manifested in her conduct (and is closely related to the concept of 'moral worth'). In this sense, the director deserves blame for prohibiting all research about the side-effects of the drug that her company sells to the extent that this speaks badly of her considered as a moral agent. Accountability, as it is taken in this debate, concerns the appropriateness of responding with reactive attitudes such as resentment to S's failures to meet moral standards that she could reasonably have been expected to meet. In this sense, the director deserves blame for prohibiting all research on a certain drug to the extent that it is fitting to resent her for falling short of some moral standard that would require her to support such research.

Apart from certain exceptions (e.g. Arpaly 2003, Talbert 2013), many philosophers deny that attributability is a substantive or interesting form of blameworthiness. Most of the participants in the debate about the epistemic condition are concerned with responsibility as accountability (cf. Zimmerman 1997, Montmarquet 1999, Rosen 2004, FitzPatrick 2008, Sher 2009, Levy 2009, Clarke 2014, among

---

[5] Here, we rely on Levy (2011, ch. 1).

[6] Two main references here are Watson (1996) and Shoemaker (2011). Clearly, there is disagreement about how the different kinds of blameworthiness are to be defined (and whether they all count as desert-based). For more, cf. Coates & Tognazzini (2013) and Clarke et al. (2015).

others). Hence, the central question will be *when* it is reasonable to expect agents to do better.[7] Sometimes it is, and sometimes it is not. The question is where to draw the line, and here the debate on the excusing conditions becomes relevant (conditions according to which agents are not accountable, in this case, and when the expectations to do better are not reasonable).

Philosophers have long agreed that responsibility has an epistemic condition (at least since Aristotle's *Nicomachean Ethics*), although the predominant focus has been on the control condition. In these debates, philosophers have abstracted away from the epistemic condition and have focused on control (cf. Fischer & Ravizza 1998, Mele 2010). The debate on the control condition has a well-developed tradition (for a recent incarnation of this debate, cf. Kane 2011). Many pages have been devoted to whether or not responsibility is compatible with determinism. In this book, this project is reversed: we will abstract from the control condition (or at least from classical concerns about determinism), and focus on the epistemic condition. The question is not when the absence of control excuses, but when ignorance does.

This book matters for several reasons. First, there is no complete account of moral responsibility absent some epistemic condition. Even if you are more concerned about other issues concerning responsibility, you will not be able to give a final verdict about an agent's responsibility without checking the epistemic condition. After all, as most will agree, the epistemic condition might render agents *fully* blameless all by itself, no matter what other conditions are met.

Second, the question is of interest in its own right. Are the director, slaveholder, and other ignorant agents blameworthy or not? As said, the epistemic condition has only recently attracted the attention of scholars,[8] and the chapters in this book constitute the main recent developments. Of course, there is substantial disagreement about the details of the epistemic condition, and the aim of this book is to further the debate.

Finally, this book is not only valuable for philosophers. We believe its results should inform legal practices of attribution of responsibility and punishment. That is, so long as legal punishment is justified in part on the basis of backward-looking considerations that might justify retribution, we should know whether and how much

---

[7] These expectations are *normative* rather than *predictive*, since they depend on whether S *should* conform to them and not on the likelihood that S will conform to them.

[8] See the journal articles and monographs listed at the end of this introduction. Cf. Peels (2016) for another recent collection on some related themes.

blame wrongdoers deserve, and we cannot determine this without a story on the epistemic condition.

2. Duty of inquiry

Issues of *scope* are important in this debate. In principle, S may be blameworthy for different things, including her omissions to inform herself, her ignorance itself, her ignorant actions, as well as consequences of the latter. Following Holly Smith's classic (1983) terminology, we will call omissions to inform oneself 'benighting acts' (such as failures to keep oneself updated about the latest research on a certain drug), and subsequent ignorant behaviour 'unwitting acts' (such as prescribing the drug). Benighting acts either induce or perpetuate ignorance about the permissibility of unwitting acts. Given the temporal order of things, we will use these abbreviations in what follows:

> A1: benighting act
> A2: unwitting act

In this section, we will say something about the wrongness of these acts and (and turn to issues of blameworthiness afterwards). We can be brief about A2's wrongness. Here we just defer to some ethical theory for a story about this. It may be wrong to keep slaves because it is wrong to harm other people, or because it is wrong to violate their rights, or because of some other reason. In this book, the exact story here does not really matter.

What does matter is the objective/subjective distinction. According to *subjective* accounts of wrongness, the wrongness of A2 rests in part on S's beliefs about A2 such that if S is ignorant about certain features of A2 (its consequences or the availability of alternative options, for example), then A2 cannot be wrong. For example, if a doctor was ignorant about the hazards of a certain drug, it would not be subjectively wrong for her to prescribe it. Indeed, if she believed that it was the best treatment option available, the subjectivist about wrongness might hold that she is even obligated to prescribe it, despite the unknown hazards. According to *objective* accounts of wrongness, in contrast, the wrongness of A2 is not threatened by ignorance. On this view, the doctor's ignorance would not suffice to neutralize the wrongness of prescribing the hazardous drug. We will assume, with most participants in this debate, that some kind of objective account is

true. Objective accounts allow for the cleaner separation of issues of wrongness and blameworthiness.[9]

The interesting question in this context concerns the wrongness of A1. Why are benighting acts wrong? Why is it sometimes wrong not to inform oneself? It is often said that agents who ignorantly bring about some harm "should have known better". Directors of pharmaceutical companies should know better. Slaveholders should know better. Yet, this expression might refer to two different things. It either means that one has duties of *belief* (namely to believe or be aware *that* some A2 is wrong) or that one has duties of *inquiry* (namely to inform oneself *whether* some A2 is wrong). On the former reading, directors of pharmaceutical companies have a duty to know that it is wrong to promote a drug with severe side-effects, and slaveholders have a duty to know that it is wrong to keep slaves. On the latter reading, directors and slaveholders have duties to take certain steps to inform themselves better. Both cases have to do with forestalling or rectifying a certain ignorance. In the inquiry case, though, one should not do this directly by acquiring a specific doxastic attitude, but by undertaking certain actions (such as asking certain questions, gathering further evidence, or reflecting about it) that will likely lead to new doxastic attitudes.

Duties of belief are known to be problematic because of the problem of *doxastic involuntarism* (cf. Alston 1988, Nottelmann 2007). As many philosophers acknowledge, agents cannot directly control what to believe. We cannot simply start believing that some A2 is wrong. Belief formation seems involuntary, and we can only indirectly influence it by conducting certain inquiries and gathering evidence. If belief formation is indeed involuntary, then one's conformity with a duty of belief will be out of one's control. It would then seem implausible to attempt to establish one's blameworthiness for ignorant action on one's involuntary failure to conform to a duty to hold this or that belief. For this reason, many philosophers in the debate on the epistemic condition operate with duties of inquiry rather than belief (cf. Zimmerman 1997, Rosen 2004, Clarke 2014, among others).

Agents might have duties of inquiry for different reasons. In professional contexts, such as medicine and engineering, in which it is clear that holding certain mistaken beliefs poses a threat to the well-being of others, the obligation to inform oneself seem to have clear grounds – it diminishes the threat of harm. Parents and other caregivers might have similar duties for similar reasons. More generally, individuals have

---

[9] For further debate and options, cf. Zimmerman (2008).

moral duties to inform themselves about the moral permissibility of their conduct.[10] Spelling out precisely *when* and *why* moral duties of inquiry obtain is a tricky issue on which not much has been written, and in the following we will only make a few suggestions.

As with any other duty, duties to inquire might be either objective or subjective, and derivative or non-derivative. Holly Smith (2014) plausibly suggests that they are objective and derivative in nature: "the fact that [inquiry] is a duty derives from the subsequent duties it would lead the agent to satisfy." (2014: 20) Pharmaceutical company directors, doctors, engineers, and parents have the duty not to expose certain classes of people to avoidable risk of harm, and the fact that performing certain inquiries would allow them to satisfy this duty is what grounds a duty to carry out these inquiries. In other words, it would be wrong for these agents to expose others to an avoidable risk of harm and to fail to perform prior actions that would diminish this risk of exposure. In short: *the benighting A1 is wrong because the unwitting A2 is wrong.* The wrongness of a doctor's failure to read her journals is derived from the wrongness of prescribing harmful drugs. Cases of moral ignorance are subject to a similar treatment. It is wrong for the slaveholder to fail to question her practices because it is wrong to keep slaves.

This way of deriving duties of inquiry is incomplete, as far as it goes. What is the exact relationship between purported duties to inquire and other duties that motivate them? It seems to be this: S has a duty to inquire D1 because she has some other duty D2 that is such that conforming to D1 will enable her to see that she has D2 (or will enable her to see *how* she can conform to D2). This derivative relationship may also be expressed in terms of the wrongness of the relevant benighting and unwitting acts: A1 is wrong because: A2 is wrong, and performing A1 will prevent S from seeing that A2 is wrong, or prevent S from seeing how she could avoid performing A2.

A complete account will be fleshed out with many more details.[11] First, there must be a way of determining the strength of duties of inquiry. Plausibly, the strength of

---

[10] In the epistemic context, you might have to inform yourself better in order to be epistemically justified in your beliefs (cf. Hall & Johnson 1988, and Feldman 2000 for criticism).

[11] What about cases where A1 seems to be wrong, but A1 does not lead to any particular A2 that is wrong? Perhaps such cases can be accounted for by considering the generality of duties of inquiry. If the wrong that the duty of inquiry will help one avoid takes a *general* description, then the duty of inquiry may be a general duty. A doctor's general duty to keep track of relevant research will derive from the wrong of causing unnecessary harm to her patients (whether or not it causes any such harm in a particular case). If the wrong that the duty of inquiry will help one avoid takes a *specific* description, the duty of inquiry may be quite specific. A doctor's duty to read a specific

a duty to inquire whether A1 is wrong derives from the degree of wrongness of A2. Holding all else equal, the slaveholder's duty to inquire whether slavery is wrong is probably stronger than a shopkeeper's duty to inquire whether her prices are fair (for example). Relatedly, there must be a way of determining whether a duty of inquiry is a prima facie or all-things-considered duty whose strength outweighs other duties. Some philosophers have suggested that these duties should not be overly demanding (cf. Rosen 2004, Gwen Bradford's chapter).

Despite the need for more theorizing on these issues, there is a consensus in the debate that A1's wrongness is *insufficient* for blameworthiness for A2. One cannot simply derive the doctor's blameworthiness for prescribing a drug from the fact that she should have known better. Rather, S's blameworthiness for any unwitting wrong action depends on whether she is *blameworthy* for not knowing better. Next, we put forward an account of this relation.

3. The Orthodoxy

In the following, we will set out what we think forms the core agreement among some main figures in the debate. We will call the full set of these claims the 'Orthodoxy'. The Orthodoxy is useful because it offers a framework that brings into focus many relevant aspects of the epistemic condition, and, as we will see, important issues can be phrased in terms of some part of the Orthodoxy.

All agree that an excuse by ignorance can render S blameless. In cases where S is not ignorant that her action is wrong, an excuse by ignorance is obviously not on the table. In cases where S is ignorant that A2 is wrong (either because of factual or moral ignorance), then two assumptions are taken to hold:[12]

(i)   S is blameworthy for A2 only if S is blameworthy for her ignorance that A2 is wrong; and

(ii)   S is blameworthy for her ignorance that A2 is wrong iff S is blameworthy for a benighting act A1 (at least one such act) that led to A2.

---

article may derive from the wrong of prescribing something harmful to a particular patient at a particular time.

[12] It is important to keep in mind which conditions are considered necessary, and which sufficient. Most of the debate concerns exactly these issues.

For example, if the doctor is ignorant that prescribing the drugs is wrong, then she is blameworthy for doing this only if she is blameworthy for her ignorance that this is wrong, and she is blameworthy for the latter iff she is blameworthy for not informing herself better about the drug. To be sure, philosophers disagree about the conditions of the latter – of blameworthiness for A1 – though (i) and (ii) themselves are widely accepted (cf. Zimmerman 1997, Ginet 2000, Rosen 2004, FitzPatrick 2008, Levy 2009, among others).

(i) states a mere *necessary* condition because there may be other conditions necessary for blameworthiness for A2. Specifically, there might be control conditions, according to which S is blameworthy for A2 only if S is not excused by a lack of control. The doctor might still be blameless for prescribing the drugs, for example, if the patient threatened her to do so. The advantage of separating (i) and (ii) is that it allows one to examine the epistemic requirements for responsibility on their own. That is, if one wants to abstract from non-epistemic requirements on responsibility, one can simply focus on blameworthiness for ignorance rather than blameworthiness for conduct generally (i.e. on (ii) rather than (i)).

Cases where (i) and (ii) apply might be cases of *derivative* or *indirect* blameworthiness. In such cases, blameworthiness for A2 (prescribing the harmful drug) derives from blameworthiness for A1 (the failure to inform oneself). In cases of *original* or *direct* blameworthiness, by contrast, blameworthiness for A2 does not derive from blameworthiness for A1. Agents who are not ignorant that their conduct is wrong, but do it nonetheless may be directly blameworthy. The question is what derivative blameworthiness exactly involves (and what original blameworthiness does not). To explain this, we need two more abbreviations:

B1: blameworthiness for benighting A1 (or the ensuing ignorance)
B2: blameworthiness for unwitting A2 (or a consequence of A2)

We believe that derivative blameworthiness for unwitting conduct involves the following claims:

Claim 1: If B1, and S has no further excuses for wrongful A2, then B2.
Claim 2: B2 only if B1.
Claim 3: B2 (partly) because B1.
Claim 4: Claims 1-3 apply to factual and moral ignorance.

11

Claim 5: Claims 1-4 are to be understood in terms of accountability.

We are the first to systematize derivative blameworthiness in this way. We do not wish to imply that all parties to the discussion on the epistemic condition accept all of its claims. This systemization is nonetheless instructive because it helps clarify what is being denied by those who defend non-derivative accounts of blameworthiness for unwitting acts (see §5). Let us take these claims in turn.

Claim 1 says that if one's ignorance is blameworthy, and one has no further excuses for one's conduct (that is, excuses unrelated to ignorance), then this conduct is also blameworthy. This claim may be called the 'transfer' claim. According to it, B1 may transfer to B2. This claim has so-called 'tracing' implications: if B1 transfers to B2, then B2 traces back to B1.[13]

Claim 2 specifies a necessary condition: B1 is necessary for B2. Thus, the ignorant doctor is blameworthy for unwittingly prescribing a harmful drug *only if* she is blameworthy for her ignorance that this is wrong. This claim entails the excuse by ignorance: if ~B1, then ~B2. If the doctor is blameless for her ignorance about the harmful nature of the drug, she is blameless for her ignorant conduct. Here, one might want to distinguish *full* excuses, which leave agents fully blameless, and *partial* excuses, which mitigate the blameworthiness to a certain degree. According to claim 2, a full excuse for A1 (~B1) implies a full excuse for A2 (~B2). The same goes for partial excuses: a partial excuse for A1 implies a partial excuse for A2.

Claim 3 specifies an explanatory condition: B2 is not only necessary for B1, the former also explains the latter. The doctor is blameworthy for prescribing the drug *because* she is blameworthy for not informing herself about it. This is an additional claim, given that not all necessary conditions are explanatory conditions. It might be condition for B2 that the agent is alive, though this does not explain B2. Still, B1 is not supposed to provide a *full* explanation for B2. For there are further conditions (such as control conditions) that, when met, also form part of the explanation for B2.

Claim 4 is the 'parity' claim. It simply says that these same claims apply to factual ignorance (e.g. that slaves suffer) and moral ignorance (e.g. that it is wrong to keep slaves). One rationale for parity is the existence of duties to inquire into factual and moral issues. If factual ignorance traces to failures to inquire into factual matters, why not think that moral ignorance traces to failures to inquire into moral matters? In his chapter, George Sher notes that the same factors that strike us as relevant for determining

---

[13] For more on tracing, cf. Vargas (2005), Fischer & Tognazzini (2009), among others.

whether it is reasonable to expect S to avoid factual ignorance – the amount and quality of evidence, or the complexity of relevant reasoning – also seem relevant to determining whether it is reasonable for S to avoid moral ignorance.

Claim 5 is that claims 1-4 are meant to be read in terms of accountability (introduced in §1). This implies that, if B1 (or B2) would apply, then S could reasonably have been expected not to perform A1 (or A2). The slaveholder is blameworthy to the extent that she could reasonably have been expected to question her practices and, indeed, to stop them.

One may wonder whether this list should be supplemented with further claims. One suggestion would be to add a *degree* claim, according to which, if claim 1 obtains and B1 transfers to B2, S's overall degree of blameworthiness increases. As Smith (1983) points out, to subscribe to this would be to accept *consequential luck*. Smith asks us to consider two agents, call them 'S1' and 'S2', who perform the same benighting act, but with different consequences. Suppose there are two doctors who fail to keep track of the latest research, but only one of them harms a patient as a result. If the consequences in S1's case are more severe than in S2's case, is S1 more blameworthy than S2? Though it is natural to think that agents can be blameworthy for certain foreseeable and avoidable consequences of their actions, it is controversial that mere luck could explain a difference in degree of blameworthiness between agents who acted similarly (and possibly from similar motives).

Furthermore, Smith argues that one way to avoid this implication is to deny that agents are ever blameworthy for their unwitting acts. Her position is that the scope of blameworthiness might only take one's benighting acts. In his chapter, Peter Graham draws even a stronger lesson, namely he denies that agents are blameworthy for their actions at all. The same luck-vulnerable relation that Smith takes to hold between benighting acts and subsequent unwitting ones, is said to hold between an agent forming an *intention* to act and the subsequent performing of the act. According to Graham, the scope of blameworthiness takes only one's intentions. In our chapter, Philip Robichaud amplify on the issue of consequential luck in the context of the epistemic condition and discuss further ways of responding to it. These include allowing only the scope of blameworthiness to vary, and allowing both the scope and degree of blameworthiness to vary (to some extent).

As we will assume here, the Orthodoxy itself is neutral about degrees of blameworthiness and many other things. One central issue it is neutral about: *the*

conditions for B1. Most of the debate so far has focused on this issue. In the next sections, we highlight the main moves in this debate.[14]

4. Internalism's regress

We will start with one main account: *internalism.*[15] According to internalism, the conditions for B1 and B2 are exactly the same, namely that blameworthiness requires witting wrongdoing. Thus, internalism holds that S is blameworthy for benighting act A1 *only if S believes that A1 is wrong* (or she is blameworthy for her ignorance about this). Alternative views will deny this necessary condition, and propose an alternative sufficient condition that can establish blameworthiness for their unwitting or benighting acts (see §5).

      As Michael Zimmerman (1997, 2008) has shown, internalism is on the road to regress. Here is a reconstruction of the regress:

(1) S is blameworthy for A2 only if
    (2) S believes that A2 is wrong, or
    (3) S is blameworthy for her ignorance that A2 is wrong
(3) only if
    (4) S is blameworthy for the past omission A1 that resulted in her ignorance that A2 is wrong
(4) only if
    (5) S believes that A1 is wrong, or
    (6) S is blameworthy for her ignorance that A1 is wrong
(6) only if etc.

Adherents of the Orthodoxy accept the regress at least until condition (4). (4) is motivated on the basis of control considerations (and the falsity of doxastic voluntarism). We are blameworthy for our ignorance only to the extent that we are blameworthy for our past conduct that resulted in our ignorance. (5)-(6) is internalism's addition, and subsequent clauses can be generated in a similar way. The regress entails that: S is

---

[14] Cf. also Talbert (2016a: ch. 5) for a statement of this debate.

[15] Sometimes called 'volitionism', as contrasted with attributionism (an umbrella term for various views that we describe in §5).

blameworthy for A2 only if (2) or (5) (or …), and so it yields Zimmerman's 'origination thesis':

> Every chain of culpability is such that at its origin lies an item of behavior for which the agent is directly culpable and which the agent believed, at the time at which the behavior occurred, to be overall morally wrong. (2008: 176)

Put in terms of a slogan: blameworthiness requires *akrasia*. S does X akratically iff, at the time of X, S believes that she should not do X, i.e. that X is all-things-considered wrong. Akrasia involves awareness of the wrongness of one's action.[16]

It is debatable whether akrasia is not only necessary for blameworthiness, but also *sufficient*. One might suspect that doing something that you know to be wrong suffices for blameworthiness. But, if control is also necessary for blameworthiness, then akrasia cannot be sufficient. Additionally, if akrasia were sufficient for blame, it would yield the wrong verdict on 'inverse akrasia' cases (cf. Arpaly 2003, 2015). In such cases, S falsely believes she does the wrong thing, while in fact she does the right thing in a way that expresses moral concern. Huckleberry Finn, for example, believes that it is wrong to help the slave Jim escape from slavery. Yet still he recognizes Jim's status as a fellow human being, and he acts with moral concern when, against his own judgment, he does help him. Indeed, in such cases, it is plausible that inverse akratic agents should be considered praiseworthy rather than blameworthy, which further problematizes the view that akrasia is sufficient for blameworthiness.[17]

In any case, important here is that Zimmerman's regress argument rests on akrasia's necessity (rather than its sufficiency). Having established the origination thesis, Zimmerman goes on to argue that akrasia hardly every occurs. In many cases, when we act, we do not occurrently believe that we should not act in that way (that is, on balance). Even if we are aware of certain considerations that disfavor acting as we do, in most cases we believe that our actions are all-things-considered permissible.

It will be helpful to consider some concrete instances of the regress. First, the doctor case: (1) The doctor is blameworthy for prescribing the harmful drugs only if (2) she is aware that this action is wrong, or (3) she is blameworthy for her ignorance that

---

[16] Akrasia is often associated with weakness of will. In the present context, akrasia is in principle compatible with different motives: a bad will, or indeed a good albeit weak will.

[17] Cf. Paulina Sliwa's chapter for an alternative analysis. According to Sliwa, Huck is praiseworthy not because he acts with moral concern, but because he is a good friend.

this is wrong. (3) obtains only if (4) she is blameworthy for not keeping up with the latest research. (4) obtains only if (5) she is aware that this is wrong, or (6) she is blameworthy for her ignorance that this is wrong. (6) obtains only if yet further conditions obtain. It also applies in the slaveholder case: (1) The slaveholder is blameworthy for keeping slaves only if (2) she is aware that this practice is wrong, or (3) she is blameworthy for his ignorance this is wrong. (3) obtains only if (4) she is blameworthy for not questioning her current practice. (4) obtains only if (5) she is aware that this omission is wrong, or (6) she is blameworthy for her ignorance that this is wrong. (6) obtains only if yet further conditions obtain.

Saying that akrasia almost never occurs implies that (2) and (5) are not satisfied in these cases. This is plausible: the doctor is not aware that prescribing the drugs is wrong, nor that her failure to better inform herself is wrong. And the slaveholder is not aware that her practice is wrong, nor is she aware that her failure to question it is wrong. Importantly, if akrasia hardly every occurs, and clauses (2) or (5) are hardly ever satisfied, then it follows that (1) is rarely satisfied and that agents are hardly ever blameworthy for anything whatsoever. The doctor is blameless for prescribing the harmful drugs, and the slaveholder is blameless for keeping slaves. In other words, Zimmerman's argument has a revisionist implication:

> Surely such ignorance does not always excuse one's behavior, since one may be culpable for one's ignorance. This is commonly acknowledged. What is not commonly acknowledged, though, is that culpability for such ignorance is rare. On the contrary, our common practice indicates that we think that such culpability is frequently incurred... I believe that this practice is misguided, however. (Zimmerman 2008: 175)

Similar arguments have been advanced by Rosen (2004) and Levy (2011). Rosen adds an epistemic twist to it. He does not say that (2) and (5) are almost never satisfied, but that it is difficult to determine that they are, and so it follows that we hardly ever *know* whether agents are blameworthy for anything. Most philosophers, though, deny the revisionist implication and consider agents blameworthy in more cases than Zimmerman and other internalists are willing to admit. Next, we will summarize some main responses to Zimmerman's regress argument.

5. Responses to the regress

To see how one might possibly respond to the regress, it will be useful to recall this reconstruction:

(1) S is blameworthy for A2 only if

    (2) S believes that A2 is wrong, or

    (3) S is blameworthy for her ignorance that A2 is wrong

(3) only if

    (4) S is blameworthy for the past omission A1 that resulted in her ignorance that A2 is wrong

(4) only if

    (5) S believes that A1 is wrong, or

    (6) S is blameworthy for her ignorance that A1 is wrong

(6) only if etc.

In terms of this, the four responses we will discuss below can be mapped as follows (though we will not explain them in these abstract terms):

(I)   Accept (1)-(6), but propose a more liberal reading of (2) and (5), which avoids the revisionist conclusion in part;

(II)   Deny (5)-(6) as necessary for (4) in favour of an alternative sufficient condition for (4);

(III)   Deny (2)-(3) as necessary for (1) in favour of an alternative sufficient condition for (1);

(IV)   Deny (4) as necessary for (3) in favour of an alternative sufficient condition for (3) (and in fact for a range of things).

(I) There are the internalist-friendly responses to the regress according to which blameworthiness might not only trace back to akrasia, but also to other mental states. For even if one does not often occurrently hold the beliefs specified in clauses (2) and (5), one might still hold certain weaker counterparts (cf. Haji 1997, FitzPatrick 2008, Peels 2011, Robichaud 2014, among others). For example, even if one is unaware at the time of acting that one's conduct is wrong, one might nonetheless suspect that it is, or hold the belief unconsciously that it is, or have sufficient (even though indecisive) evidence that it is. For example, even if the slaveholder does not occurrently believe that

what she is doing is wrong, she might still vaguely suspect that something is wrong. On the basis of these more liberal readings of what clauses (2) and (5) require, one might suspect that they are more frequently satisfied than Zimmerman supposes.

(II) Externalists take a wholly different approach. Here are two representative quotes:

> The root idea here, it seems to me, is that a certain quality of *openness* to truth- and value-related considerations is expected of persons and that this expectation is *fundamental*, at least in the following regard. The expectation is not derivative of or dependent upon one's (at the moment in question) judging such openness as appropriate (good, required, etc.) – just the opposite: it would include a requirement that one be open to the need to be open, and if one is not open to this, one may be blameworthy precisely for that failure. (Montmarquet 1999: 845)

> Ignorance, whether circumstantial or normative, is culpable if the agent could reasonably have been expected to take measures that would have corrected or avoided it, given his or her capabilities and the opportunities provided by the social context, but failed to do so either due to akrasia or due to the culpable, nonakratic exercise of such vices as overconfidence, arrogance, dismissiveness, laziness, dogmatism, incuriosity, self-indulgence, contempt, and so on. (FitzPatrick 2008: 609)

The basic idea is that blameworthiness may not only trace to akrasia, but also to the exercise of epistemic vices. Epistemic vices such as laziness, arrogance and dogmatism are character traits of the agent that may prevent her from acquiring evidence that other more virtuous agents do acquire, or prevent her from evaluating and weighing evidence in the right way. The exercise of these vices would explain why she is ignorant while more virtuous agents are not. More precisely, externalism proposes that exercises of epistemic vices can be sufficient for blameworthiness even in the absence of any relevant episode of akrasia. On this account, directors of pharmaceutical companies and slaveholders might be considered blameworthy – even if they are not aware that they are doing anything wrong – if their ignorance (and their failure to ask certain questions and inquire into certain things) is the result of their vices and if they would not have been ignorant had they been more virtuous.

According to some externalists, the history of S's vices is relevant. They maintain that S might still be blameless for A1, even if A1 is due to her epistemic vices, namely in the case that S had no opportunities to be less vicious. Consider Susan Wolf's JoJo case (1987). JoJo is the son of a dictator who tortures people in the belief that this is permissible, and this belief is the result of exercising all the vices that externalists refer to. JoJo is arrogant and dogmatic, and he never questions how he treats other people. These vices have a history. JoJo had an extremely warped upbringing, and let us suppose that he really had no opportunities to develop any epistemic virtues.[18] In such a case, externalists might still think that JoJo is excused for his ignorance and hold more generally that: S is blameworthy for A1 *if A1 is due to S's epistemic vices and S had normal opportunities to develop her virtues*. For a debate on this historical condition, see the chapters by William FitzPatrick (who accepts such a historical condition) and Matthew Talbert (who does not).[19]

(III) Another way of avoiding the regress is to break with the Orthodoxy that internalists and externalists both accept. On this approach, S's blameworthiness for A2 does not imply that S is blameworthy for her ignorance that A2 is wrong. In fact, it does not require that any other blameworthiness fact be true of S. George Sher and Randolph Clarke each defend such an account, which we will explain in turn. Sher motivates his view by appealing to cases where agents are intuitively blameworthy for their unwitting conduct. One widely discussed case of his is the following:[20]

> *Hot Dog*. Alessandra, a soccer mom, has gone to pick up her children at their elementary school. As usual, Alessandra is accompanied by the family's border collie, Bathsheba, who rides in the back of the van. Although it is very hot, the pick-up has never taken long, so Alessandra leaves Sheba in the van while she goes to gather her children. This time, however, Alessandra is greeted by a tangled tale of misbehavior, ill-considered punishment, and administrative

---

[18] Or develop any concern for other people. JoJo might be a bad *epistemic* agent (unconcerned about the falsity of his beliefs), or a bad *moral* agent (unconcerned about the wrongness of his conduct), or both. See response (IV) below.

[19] The controversy can also be stated in terms of the relevance of *constitutive luck*, i.e. luck concerning one's vices, which might well be the result of coming up in a social context that tolerated and perhaps even rewarded them.

[20] Sher's case, and the one by Clarke discussed below, are forgetting cases. The agents do seem to know all the relevant facts, but at the time of action they fail to put things together and see what they have to do.

bungling which requires several hours of indignant sorting out. During that time, Sheba languishes, forgotten, in the locked car. When Alessandra and her children finally make it to the parking lot, they find Sheba unconscious from heat prostration. (Sher 2009: 24)

Sher maintains that Alessandra is blameworthy for the omission of leaving Sheba in the hot car and for the consequences of that omission. Moreover, he argues that in this case, as well as similar cases, there is no benighting act to speak of (2009: 35). In Hot Dog, the sequence of events that distracted Alessandra and led to Sheba's harm were unpredictable, and so there is no reason to think that there was some action that Alessandra should have performed in order to shore up her memory that Sheba was at risk. On the basis of this, Sher develops an account that obviates the need to trace blameworthiness for unwitting conduct to any blameworthy benighting act.

Sher's view can be put as follows: S is blameworthy for unwitting A2 if relevant non-epistemic conditions are met, and if S's failure to realize that A2 is wrong (1) falls below a relevant standard and (2) is caused by attitudes constitutive of S. On this view, short of establishing that S is blameworthy for her ignorance, we need only show that there is some standard of which S falls short and that her falling short has a certain causal dimension. Condition (1) ensures that S had good evidence that A2 is wrong, and that S did possess the relevant cognitive capacities to evaluate that evidence. Condition (2) is an agency condition that ensures that S's coming to be ignorant is grounded in her own agency. In Alessandra's case, both conditions are met, and she is thus blameworthy for leaving Sheba to suffer. As to (1), dog-owners are expected to remember their dog when they are sitting vulnerable in a hot car. As to (2), Alessandra's failure to keep Sheba in mind as she deals with the situation in the school plausibly traces to her disposition to keep important things in mind, which in this case was unfortunately weak.

In his chapter, Clarke proposes an account that also eschews the necessity of establishing blameworthiness for ignorance. His main case that motivates his view:

Forgotten Milk. As I'm about to leave my office at the end of the workday, my wife calls to tell me we're out of milk. My regular route home takes me right by a grocery store, and I tell her I'll stop and buy some. Between my office and the store, I start to think about a paper I'm writing on omissions. I continue thinking about my work until I arrive home, where I realize that I've forgotten the milk. Not a big deal, but still, my wife might think, grounds for blame. We'll lack milk

20

for our breakfast in the morning – or someone will have to go back out to get it – and the fault lies with me. (Clarke 2014: 164)

Clarke agrees with Sher that in cases like this, the agent's blameworthiness need not trace to benighting acts. In his analysis of this case, Clarke disputes that he had any obligation to set an alarm or to write a note to remind himself, at least in the absence of a pattern of forgetfulness. It suffices for blameworthiness for unwitting acts if S's unawareness "falls below a cognitive standard that applies to her, given her cognitive and volitional abilities and the situation she is in" (Clarke 2014: 167). In Forgotten Milk, Clarke could reasonably have been expected to remember to buy the milk given his abilities and the situation he was in.

In his chapter, Neil Levy focuses critical attention on the intuitions that motivate Clarke's view, and offers an error theory of them. Levy's main argument is that the adaptive function of our disposition to make blameworthiness attributions admits of the possibility that these attributions are systematically mistaken. When accurate, attributions of blameworthiness are adaptive because they track mental states that make agents unreliable in cooperative contexts. The relevant mental states include those that are constitutive of the lack of due regard. Levy maintains that our disposition to use the undesirable consequences of S's behavior as indirect evidence for these mental states may result in the intuition that Clarke is blameworthy in Forgotten Milk, even though there is no prior benighting act upon which to ground these judgments. If this speculation is right, the intuitive judgments upon which Clarke relies are unreliable.

A different objection has been developed by so-called *quality of will* theorists. Although they do not dispute that the agents in Hot Dog and Forgotten Milk fall short of important standards, they argue that Sher's second condition is too weak. For example, Matthew Talbert (this book) argues that the mere fact that one's failure to realize that her conduct is wrong is grounded in her constitutive attitudes does not suffice for blameworthiness. These failures must also be grounded in *a lack of moral concern*. In Hot Dog, however, it is unclear that Alessandra is insufficiently concerned about her dog, which makes it controversial whether she is blameworthy for what happens. Clarke's view is subject to a similar point (to which he responds in his chapter), namely that in falling short of cognitive standards one might not thereby lack moral concern. Clarke forgets the milk, but that need not indicate that he is insufficiently concerned with his wife and respecting promises made to her.

(IV) The fourth response to the regress is to explain blameworthiness for ignorance in quality of will terms (rather than in terms of prior blameworthy conduct).

Angela Smith (2005) has argued that S can be blameworthy for an attitude, including a failure to recognize or remember something, if her attitude is due to a lack of moral concern on S's part. Smith motivates this account on the basis of an example of someone who forgets her friend's birthday. What is revealed in her forgetting is an insufficient degree of concern for her friend, and on the basis of this Smith argues that it makes sense to claim that she is blameworthy for her ignorance, which in this case is her failure to occurrently believe that it is her friend's birthday.

While response (II) above also holds that agents are blameworthy when they lack a certain kind of concern for avoiding ignorance, this lack of concern might be purely epistemic. They might be dogmatic or lazy thinkers, and need not be unconcerned with morality. Quality of will views, by contrast, accept a requirement that blameworthy agents fail to be concerned with *morality*.

Another influential quality of will account is due to Nomy Arpaly (2003). Roughly, her account says that S is blameworthy for some unwitting A2 iff *A2 is due to a lack of concern for the features that make A2 wrong*. According to Arpaly, the relevant account of moral concern involves is not concern for what an agent thinks is morally relevant to her action (*de dicto* concern), but concern for what actually is morally relevant to her action (*de re* concern). One may have *de dicto* moral concern while lacking *de re* moral concern. To illustrate, consider "the ancient Roman who goes the circus because he heartily enjoys watching people thrown to the lions" (Arpaly & Schroeder 2014: 182). One can imagine that this Roman is simply giving in to his bloodlust and that he thereby shows a lack of concern for the suffering at the heart of the spectacle. On Arpaly's view, he would be blameworthy for going to the circus even if he believed that he cared about morality.

What is exactly involved in being concerned about people's suffering (which, let us assume, makes going to the circus wrong)? It need not require the belief that boycotting the circus is morally significant. Arpaly leaves room for the possibility of inverse akrasia (discussed earlier), i.e. the possibility that one can have moral concern despite one's mistaken belief that one's conduct is wrong. Rather than conceptualizing one's conduct as having moral significance, a Roman with moral concern will have a particular cluster of dispositions (cf. Arpaly 2003: 84-93). In the emotional dimension, he will be disposed to feel sympathy toward those who are suffering; in the cognitive dimension, he will be disposed to notice their suffering; and in the motivational dimension, he will be disposed to boycott the circus.

While the details of different quality of will accounts matter, they are united in their denial that blameworthiness for unwitting acts is derivative (in the sense of claims

1-5, discussed earlier). In particular, they deny that blameworthiness for unwitting acts is to be explained by blameworthiness for benighting acts. Instead, blameworthiness for unwitting acts is to be explained by a lack of moral concern. Moreover, agents might be non-derivatively blameworthy for a whole range of things (unwitting and benighting acts, failures to notice, remember, and other attitudes) so long as they are all explained by a lack of moral concern. For example, in his chapter Gunnar Björnsson defends a quality of will account according to which S is blameworthy for X when and insofar as X is explained by S's quality of will falling short of what can reasonably be expected. When X is *not* explained by a substandard quality of will, S will not be blameworthy for X. The quality of will approach is popular, and many authors in this book defend a quality of will view of some stripe or other (see also Talbert, Harman, Zimmerman, Mason & Wilson, Alvarez & Littlejohn).[21]

In the foregoing, we have summarized several accounts on the epistemic condition (revisionist internalism, plus four main alternative accounts). Given that they contradict each other (certain conditions cannot be sufficient, if others are necessary, and the other way around), the question arises: *which is the best account?* One straightforward criterion of theory choice is intuitions-based. According to it, the best account is the account that covers our intuitions about all of the relevant cases.

Clearly, though, this criterion suffers from familiar problems. First of all: intuitions of the participants in the debate are often theory-laden. In such a case, they simply have the intuitions that match their own theories. Second, as far as we know, not much empirical work has been done regarding the intuitions of non-experts on the epistemic condition.[22] Also, even if we would have such data, it might still not tell us much so long as the intuitions of non-experts are the result of certain cognitive biases or – which is actually rather likely – of conceptual confusions. The folk does not always distinguish between moral blameworthiness, wrongness of the action, badness of the agent, and legal punishment, and it is not easy to see how clean experiments can be constructed to avoid such confusions. Still, even if there are such problems with the intuitions-based criterion, it seems to be a mistake to think that we can do without. In the end, any general account will rely on intuitions about certain cases.

---

[21] It is important to note that some quality of will theorists are interested in attributability rather than accountability. They maintain that if S is blameworthy for X, then X speaks badly of S, even if it would not be reasonable to expect S to have avoided X.

[22] An exception is Faraci & Shoemaker (2014) who examine folk views about praise- and blameworthiness of Huck Finn and JoJo (discussed above).

6. Why does blameless ignorance excuse?

Another potential criterion of theory choice is the following: the best account is the account that offers the best explanation of *why blameless ignorance excuses*. Indeed, why should we think that S is blameless for A2 if S is blameless for her ignorance that A2 is wrong? This is a tricky question, and any account that can answer to it has an obvious advantage over those that cannot. Rosen (2003, 2008) discusses two competing answers to this question:

> *Fairness explanation*:
> Blameless ignorance excuses because it is unfair to consider S blameworthy for A2, which she blamelessly considered permissible.

> *Quality of will explanation*:
> Blameless ignorance excuses because S does not perform A2 with a lack of moral concern.

The fairness explanation has been employed by internalists, and the quality of will explanation by quality of will theorists (and it is an open question whether these explanations can be utilized by other accounts as well). Next, we will briefly reconstruct both explanations.

Rosen states the fairness explanation here: "It is unfair to blame someone for doing something if he blamelessly believes that there is no compelling moral reason not to do it." (2003: 74) Suppose that you requested sugar in your coffee and that I want to give you what you want. If I am blamelessly ignorant that the substance I put in your coffee is actually poison, then I mistakenly believe that there is no compelling reason not to stir the substance into your coffee. Given that I see no such reason, it is unreasonable to expect me to refrain from giving you the coffee, and it would be unfair to consider me blameworthy for doing it and for any resulting harm.

In steps, the fairness explanation can be stated as follows, borrowing further premises from Levy (2009):

(P1)   If ~B1, then A2 can only be avoided on the basis of luck or akrasia.

(P2)   If A2 can only be avoided on the basis of luck or akrasia, then the expectation to avoid A2 would be unreasonable.

(P3)   If the expectation to avoid A2 would be unreasonable, then B2 would be unfair, and so ~B2.

(C)   Hence: If ~B1, then ~B2. [from P1-P3]

We assume that there is some benighting A1 that S should have performed because it would have prevented her from being ignorant about the fact that subsequent act A2 is impermissible. If S is ignorant that she should avoid A2 (and is not blameworthy for her ignorance), then according to (P1) she can only avoid A2 on the basis of *luck* or *akrasia*. In the sugar case, let us assume that I am expected to refrain from putting the substance in your coffee. How can I comply with this expectation if I do not know that I am supposed to comply with it? I might be lucky and accidentally spill the poison from the spoon in my attempt to add it to your cup. Or I might believe that, given your request, I should add sugar to your coffee, but akratically fail to do so. Moreover, if luck or akrasia are the only ways in which I can comply with the expectation, the latter is unreasonable, which yields (P2). Finally, we can take (P3) on board so long as considering S blameworthy for her failure to comply with an unreasonable expectation would be unfair, and so long as blameworthiness cannot be unfair.[23]

Let us next consider the quality of will explanation for why blameless ignorance excuses. Rosen voices a condensed version of this explanation here: "non-culpable ignorance of fact excuses when it does because it defeats the inference from harm done to an objectionable attitude towards others – malice or insufficient regard – on the part of the agent." (2003: 72) Rather than pinning the explanation to assumptions about fairness and rationality, the current explanation focuses on S's moral concern (or regard for the interests of others). Here is a reconstruction of this explanation in steps:

(P1)   If ~B1, then S expresses no lack of concern for A1's wrongmaking features.

(P2)   If S expresses no lack of concern for A1's wrongmaking features, then S expresses no lack of concern for A2's wrongmaking features.

(P3)   If S expresses no lack of concern for A2's wrongmaking features, then ~B2.

(C)   Hence: If ~B1, then ~B2. [from P1-P3]

---

[23] At least in the accountability sense, i.e. the kind of blameworthiness according to which S can reasonably be expected to do better (see §1).

On one quality of will criterion, S is blameworthy for X iff S expresses a lack of concern for features that make X wrong (see §5). (P1) and (P3) follow straightforwardly from this. Furthermore, (P2) follows, so long as the wrongmaking features of A1 and A2 are the same, and this seems to be plausible in many cases.[24] In the sugar case, we can assume that the wrongmaking features of A1 (not informing myself that there is poison in my sugar bowl) and A2 (putting poison in your coffee) involve the harm that would result from them. If I do not act with a lack of concern for these harm-related reasons when I do not check for poison, then I must act without the relevant lack of concern when I unwittingly poison you. Now, *if* I am blameless for failing to check whether the sugar is poison (perhaps there is no reason to suspect that it is tainted), then on the basis of this argument it follows that I am blameless for poisoning you.

Although both explanations share the same conclusion (namely the excuse by ignorance), it is useful to consider cases where they come apart. Rosen's examples here include the ancient slaveholder and the sexist father. These agents are ignorant that their conduct (keeping slaves, oppressing one's daughter) is wrong, but their ignorance may be due to their lack of concern with slavery and sexism. Hence, on quality of will standards, they might well be considered blameworthy.[25] But, the same need not carry over to the fairness explanation, that is, so long as the conditions of B1 are not quality of will conditions. For example, if the conditions were internalist such that agents are blameless for their ignorance when it does not trace back to akrasia, then the slaveholder and the father may not be blameworthy despite lacking the relevant concerns. They simply may not have thought that these were important moral issues that they should question.

In the end, then, we are back at a fundamental disagreement over *the conditions of B1*. What look like competing explanations of *why* blameless ignorance excuses unwitting agents actually depends on competing accounts of *when* they are blameworthy for their ignorance. This suggests that we cannot get traction on the former without having clarity on the latter.

---

[24] As discussed in §2, it is plausible to think that A1's wrongness actually derives from A2's wrongness.

[25] In their chapter, Maria Alvarez & Clayton Littlejohn further take up this argument.

7. Further issues

In this section, we map out further main issues that receive attention in this book. These issues constitute either a refinement of some part of the Orthodoxy (set out in §3) or a challenge to it.

Difficulty. It is plausible to think that difficulty has some effect on blameworthiness. Compare: for one director of a pharmaceutical company, it is very easy to develop alternative drugs (which are more effective and have no harmful side-effects) since all shareholders are open to such alternatives. For a second director, it is rather difficult to develop such drugs; no shareholder wants to hear about them, since they only want to maximize profit. Suppose both directors fail to develop the alternative drugs. Plausibly, the former director is more blameworthy for her failure than the latter, given the difference in degree of difficulty (of acting differently, in this case).

Similar considerations carry over to the epistemic condition. This time, we are not interested in the difficulty of avoiding A2, but in the difficulty of avoiding A1, and in the effect this has on B2. Compare: for one director it is very easy to gather more information about an alternative drug; all the research has been done, and the relevant journals need only be consulted. For a second director, it is rather difficult to inform herself; no research has yet been done, and she would first have to finance long-term projects to obtain the relevant information. Plausibly, the former director is more blameworthy for her failure to develop the alternative drugs than the latter, given the difference in degree of difficulty (of informing herself, in this case).

In their chapters, Gwen Bradford and Alexander Guerrero both develop this idea, albeit in different directions. According to Bradford, difficulty per se does not excuse, but the *source* of the difficulty may do so. Compare two directors for whom it is equally difficult to inform themselves about the best strategy for their pharmaceutical company. Yet for one director, the source of difficulty is her context: it is just very hard to determine the best strategy. For the other director, the source of difficulty is herself: in principle the best strategy is easy to determine (she need only consult an expert, let us suppose), but she is a rather cranky and insensitive person and finds this hard to do, making it equally difficult for her to inform herself. Intuitively, Bradford points out, the latter source of difficulty does not mitigate blameworthiness, at least not in the same way that the former source might.

Guerrero focuses not so much on *sources* of difficulty as on *kinds* of difficulty. It might be very hard to better inform oneself because it requires a lot of *effort* (solving a math problem which requires focus and precision), or because it requires a lot of *skill*

(determining the whereabouts of a dangerous snake), or because it is even hard to *try* to inform oneself in the first place (questioning widespread practices of slavery, and challenging the customs and values of one's culture). As Guerrero argues, these kinds of difficulty affect blameworthiness in different ways. In particular, skill-related difficulty often mitigates blameworthiness, while effort-related difficulty rarely does so (and in trying cases, it depends).[26]

Social context. A related issue is whether S can be excused (for B1, and by implication for B2) just because her ignorance is in some way supported by her culture or social context. Ancient slaveholders, early 20th century sexists, and contemporary purchasers of sweatshop-made products, all of whose social contexts widely tolerate or even encourage false beliefs about the permissibility of their conduct, may be less blameworthy than those raised in social contexts that openly condemn such conduct. The issue of difficulty is also relevant here. If one's social context constantly challenges one's sexist beliefs, for example, one might find it easier to reflect critically on one's views and revise them (and also cease one's sexist conduct).

Arguably, though, social context does not always excuse. As Moody-Adams (1994) has argued, we sometimes face cases of *affected ignorance*. As she writes: "Affected ignorance – choosing not to know what one can and should know – is a complex phenomenon, but sometimes it simply involves refusing to consider whether some practice in which one participates might be wrong." (1994: 296) For example, according to Moody-Adams's analysis, ancient slaveholders were ignorant about the wrongness of their conduct not because they could not have known better, but because it was in their interest to ask no questions. Moreover, insofar as Moody-Adams is concerned, the fact that the false beliefs of ancient slaveholders were so widespread and shared by most of their contemporaries is no bar to their being blameworthy.[27] In their chapters, Elinor Mason & Alan Wilson and George Sher further refine and qualify this analysis.

---

[26] A follow-up issue concerns the question of what *grounds* the excusing effects of difficulty. In terms of the epistemic condition, the question is what would explain the fact that difficulty in conforming to some duty of inquiry, or seeing that one has such a duty, provides an excuse for B1 and/or B2.

[27] In a similar way, Harman (2011) rejects claim 4 of the Orthodoxy, and holds that, in contrast to factual ignorance, moral ignorance (e.g. that slavery is wrong) never excuses. On her view, it does not matter if S fails see certain moral truths even if she tried really hard, and exerted a lot of effort, to inform herself and to get it right. For in many such cases S's ignorance is due to a lack of adequate moral concern (cf. Wieland 2015).

Awareness. According to the Orthodoxy, when it comes to blameworthiness, what matters is awareness or unawareness of the fact that one's conduct is wrong on balance. S is blameworthy for X if she is aware of this, and blameless if she is blamelessly unaware of this. For example, the slaveholder is blameworthy for keeping slaves if she is aware that this is wrong on balance (and thus acts akratically). But the question is whether this is the content that is relevant to blameworthiness.

As Arpaly (2003) and other quality of will theorists argue, when it comes to blameworthiness, all depends on whether S is sufficiently sensitive to the wrongmaking features of her action (such as whether her slaves suffer) rather than on whether S is also aware of the wrongness of her conduct (that keeping slaves is wrong). Similarly, when it comes to praiseworthiness, all depends on whether S is sufficiently sensitive to rightmaking features (such as that Jim is a human being with his own rights) rather than on whether S is also aware of the rightness of her conduct (that saving Jim is the right thing to do). Elizabeth Harman's chapter starts from a similar point, but she argues that the details of this position are trickier than they might seem at first sight. Paulina Sliwa, in contrast, aims to resist this whole approach. In her view, it is insufficient for praiseworthiness (blameworthiness) that S is sensitive to rightmaking (wrongmaking) features. S also has to represent her action as the right (wrong) one. Huck is praiseworthy for saving Jim, then, only if he knows that it is right to do so (and, according to Sliwa, we can explain away intuitions that Huck might be praiseworthy just if he is sensitive to rightmaking features).

Kinds of blameworthiness. According to claim 5 of the Orthodoxy, the operative sense of blameworthiness in B1 and B2 is the accountability sense. In his chapter, Talbert denies this. He defends an attributability take on blameworthiness, according to which S is blameworthy for X when X reflects an objectionable moral orientation on S's part, even when one could not reasonably have expected S to know better. For example, JoJo does not care enough about the interests of the people that he tortures, and his torturing reflects an objectionable moral orientation on his part, even though one could not reasonably have expected him to know better given his upbringing. While it would be appropriate to protest against JoJo's wrongdoing, and consider him blameworthy for his unwitting conduct in the attributability sense, Talbert suggests that he need not be blameworthy, in the accountability sense, for any benighting act.[28]

---

[28] In this sense, Talbert also denies claim 2: B1 is not required for B2.

Others pursue a more pluralist line.[29] In his chapter, Michael Zimmerman argues that careful attention both to kinds of mental states that seem necessary for moral responsibility and to the nature of blameworthiness itself show that there are several varieties of blameworthiness, each with its own conditions and set of fitting responses. If, for example, an agent's utterance of a slur was not motivated by malice, and merely reflected her morally stunted upbringing, then she may not meet the conditions for one kind of blameworthiness. However, given her still-objectionable beliefs, she may meet the conditions for another kind of blameworthiness that requires only that her beliefs and actions reflect her own moral take on things.

Radical denial. Matt King's chapter constitutes a radical denial of the B1/B2 framework at the heart of the Orthodoxy. He argues that we can capture everything we want to say about blameworthiness for ignorant action by attending to the *recklessness* and *negligence* of ignorant agents. If S knowingly flouts an obligation to inform herself and then unwittingly harms someone, her blameworthiness for the latter is established by the reckless adoption of a plan that included the risk of unwittingly causing harm. If, on the other hand, S does not knowingly flout an obligation to inform herself, but simply forgets or neglects to do so, then her blameworthiness for any subsequent unwitting act will rest simply on the fact that the unwitting act fell within the scope of the risk she took in failing to inform herself. According to King, therefore, it would be unnecessary and misleading to trace blameworthiness for unwitting acts to blameworthiness for ignorance.

8. Prospects

In light of the preceding, it seems that further research on the following two issues regarding the epistemic condition would be especially beneficial. First, the debate has resulted in rather different accounts of blameworthiness for ignorance (§§4-5) and associated explanations for why blameless ignorance excuse (§6). The question arises on the basis of which criteria these accounts can be evaluated, and one issue is whether the intuitions which they try to systematize are ultimately reliable. Second, philosophers have recently begun to distinguish between different kinds of blameworthiness (§1), and

---

[29] Mason (2015) distinguishes 'ordinary' blameworthiness, for which claim 5 is true (as well as the rest of the Orthodoxy) and 'objective' blameworthiness, for which 5 is false.

the details of the epistemic condition may well depend on this. The implications of this are to be explored further.[30]

References

Alston, W.P. 1988. The Deontological Conception of Epistemic Justification. *Philosophical Perspectives* 2: 257-99.

Alvarez, M. & C. Littlejohn 2016. When Ignorance is No Excuse. This volume.

Arpaly, N. 2003. *Unprincipled Virtue. An Inquiry Into Moral Agency*. OUP.

⸺ 2015. Huckleberry Finn Revisited: Inverse Akrasia and Moral Ignorance. In R. Clarke et al. (pp. 141-56).

Arpaly, N. & T.A. Schroeder 2014. *In Praise of Desire*. OUP.

Björnsson, G. 2016. Explaining (Away) the Epistemic Condition of Moral Responsibility. This volume.

Bradford, G. 2016. Hard to Know. This volume.

Clarke, R. 2014. *Omissions. Agency, Metaphysics, and Responsibility*. OUP.

⸺ 2016. Ignorance, Revision, and Commonsense. This volume.

Clarke, R., M. McKenna, & A.M. Smith 2015. *The Nature of Moral Responsibility. New Essays*. OUP.

Coates, D.J. & N.A. Tognazzini 2013. *Blame. Its Nature and Norms*. OUP.

Faraci, D. & D. Shoemaker 2014. Huck vs. JoJo: Moral Ignorance and the (A)symmetry of Praise and Blame. *Oxford Studies in Experimental Philosophy* 1: 7-27.

Feldman, R. 2000. The Ethics of Belief. *Philosophy and Phenomenological Research* 60: 667-95.

Fischer, J.M. & M. Ravizza 1998. *Responsibility and Control. A Theory of Moral Responsibility*. CUP.

Fischer, J.M. & N. Tognazzini 2009. The Truth about Tracing. *Noûs* 43: 531-56.

FitzPatrick, W.J. 2008. Moral Responsibility and Normative Ignorance: Answering a New Skeptical Challenge. *Ethics* 118: 589-613.

⸺ 2016. Unwitting Wrongdoing, Reasonable Expectations, and Blameworthiness. This volume.

---

[30] I thank several contributors to this book and colleagues for their feedback. Special thanks go to Phil Robichaud for numerous stimulating discussions on the Orthodoxy, and quite a few phrasings are from his hand.

Ginet, C. 2000. The Epistemic Requirements for Moral Responsibility. *Philosophical Perspectives* 14: 267-77.

Graham, P. 2016. The Epistemic Condition on Moral Blameworthiness, A Theoretical Epiphenomenon. This volume.

Guerrero, A.A. 2007. Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution. *Philosophical Studies* 136: 59-97.

— 2016. Intellectual Difficulty and Moral Responsibility. This volume.

Haji, I. 1997. An Epistemic Dimension of Blameworthiness. *Philosophy & Phenomenological Research* 57: 523-44.

Hall, R.J. & C.R. Johnson 1998. The Epistemic Duty to Seek More Evidence. *American Philosophical Quarterly* 35: 129-39.

Harman, E. 2011. Does Moral Ignorance Exculpate? *Ratio* 24: 443-68.

— 2016. When Is Failure to Realize Something Exculpatory? This volume.

Kane, R. 2011. *The Oxford Handbook of Free Will. Second Edition.* OUP.

King, M. 2016. Tracing the Epistemic Condition. This volume.

Levy, N. 2009. Culpable Ignorance and Moral Responsibility: A Reply to FitzPatrick. *Ethics* 119: 729-41.

— 2011. *Hard Luck. How Luck Undermines Free Will and Moral Responsibility*. OUP.

— 2016. Methodological Conservatism and the Epistemic Condition. This volume.

Mason, E. 2015. Moral Ignorance and Blameworthiness. *Philosophical Studies* 172: 3037-57.

Mason, E. & A. Wilson 2016. Vice, Blameworthiness, and Cultural Ignorance. This volume.

Mele, A.R. 2010. Moral Responsibility for Actions: Epistemic and Freedom Conditions. *Philosophical Explorations* 13: 101-11.

Montmarquet, J.A. 1999. Zimmerman on Culpable Ignorance. *Ethics* 109: 842-5.

Moody-Adams, M.M. 1994. Culture, Responsibility, and Affected Ignorance. *Ethics* 104: 291-309.

Nottelmann, N. 2007. *Blameworthy Belief. A Study in Epistemic Deontologism*. Springer.

Peels, R. 2011. Tracing Culpable Ignorance. *Logos & Episteme* 2: 575-82.

— 2014. What Kind of Ignorance Excuses? Two Neglected Issues. *Philosophical Quarterly* 64: 478-96.

— 2016. *Moral and Social Perspectives on Ignorance*. Routledge.

Robichaud, P. 2014. On Culpable Ignorance and Akrasia. *Ethics* 125: 137-51.

Rosen, G. 2003. Culpability and Ignorance. *Proceedings of the Aristotelian Society* 103: 61-84.

— 2004. Skepticism about Moral Responsibility. *Philosophical Perspectives* 18: 295-313.

— 2008. Kleinbart the Oblivious and Other Tales of Ignorance and Responsibility. *Journal of Philosophy* 105: 591-610.

Sher, G. 2009. *Who Knew? Responsibility Without Awareness*. OUP.

— 2016. Blame and Moral Ignorance. This volume.

Shoemaker, D. 2011. Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility. *Ethics* 121: 602-32.

Sliwa, P. 2016. On Knowing What's Right and Being Responsible for it. This volume.

Smith, A.M. 2005. Responsibility for Attitudes: Activity and Passivity in Mental Life. *Ethics* 115: 236-71.

Smith, H.M. 1983. Culpable Ignorance. *Philosophical Review* 92: 543-71.

— 2014. The Subjective Moral Duty to Inform Oneself before Acting. *Ethics* 125: 11-38.

Strawson, P.F. 1962. Freedom and Resentment. *Proceedings of the British Academy* 48: 187-211.

Talbert, M. 2013. Unwitting Wrongdoers and the Role of Moral Disagreement in Blame. *Oxford Studies in Agency and Responsibility* 1: 225-45.

— 2016a. *Moral Responsibility*. Polity.

— 2016b. Akrasia, Awareness, and Blameworthiness. This volume.

Van Woudenberg, R. 2009. Excusing Conditions for False Beliefs: Ignorance and Force. *American Philosophical Quarterly* 46: 373-86.

Vargas, M. 2005. The Trouble with Tracing. *Midwest Studies in Philosophy* 29: 269-91.

Watson, G. 1996. Two Faces of Responsibility. *Philosophical Topics* 24: 227-48.

Wieland, J.W. 2015. What's Special About Moral Ignorance? *Ratio*.

Wieland, J.W. & P. Robichaud. 2016. Blame Transfer. This volume.

Wolf, S. 1987. Sanity and the Metaphysics of Responsibility. In F.D. Schoeman (ed.), *Responsibility, Character, and the Emotions* (pp. 46-62). CUP.

Zimmerman, M.J. 1997. Moral Responsibility and Ignorance. *Ethics* 107: 410-26.

— 2008. *Living with Uncertainty. The Moral Significance of Ignorance*. CUP.

— 2016. Moral Responsibility and Quality of Will. This volume.