

SKEPTICISM ABOUT MORAL RESPONSIBILITY

Gideon Rosen
Princeton University

1. Introduction

Suppose you believe, as many philosophers do, that if the physical universe is a deterministic system then no one is morally responsible for what he does. Suppose you further believe, as more philosophers probably should, that it is an open empirical question whether the universe is deterministic. Then since you are uncertain whether a necessary condition for responsibility obtains, you should not judge (for example) that Dr. Crippen was morally responsible for the bloody murder of Belle Elmore. That is not to say that you should judge that Crippen was *not* responsible. It is rather to say that, unless you have some further exculpatory information, you should *suspend judgment* as to whether Crippen and his ilk are properly responsible for their transgressions. You should embrace a version of what I shall call **skepticism about moral responsibility**. You should hold that confident positive judgments of responsibility are never justified.¹

Note that a skeptical position of this sort need not amount to full-blown *philosophical* skepticism about responsibility. The claim need not be that thanks to some enduring structural feature of the human predicament, it is impossible in principle for us to know that Dr. Crippen is on the hook. The skeptic might simply claim that given certain contingent and possibly temporary limits on our knowledge, **confident positive judgments of responsibility are, for the moment at least, unwarranted.**

In what follows I sketch an argument for a moderate skepticism of this sort. Indeed in one respect the view I shall defend is even weaker. **My skeptic concedes that the conditions for moral responsibility are satisfiable in the actual world.** In fact, he may concede that they are frequently satisfied. His skeptical claim is quite narrowly focused. It is that whatever we may know at a general level about the possibility of culpable bad action in the actual world, **in any particular case in which we may be tempted to judge that X is responsible for doing A, this particular judgment of responsibility would be unwarranted.**

(In this respect my skeptic is like the sensible paranoid who maintains that while the Man has his spies everywhere, any confident judgment to the effect that *Oricutt* is a spy would be unwarranted for lack of evidence.)

So my skepticism is not as radical as it might be. And yet I should acknowledge that for some readers *any* skepticism on this point is bound to seem perverse. There is a powerful tendency to insist that if we know anything at all about the moral life (as I suppose we do), we know that the cold-blooded murderer or tax cheat who acts from nothing more subtle than ordinary greed simply *has no excuse*. My aim in this paper is to provide grounds for reconsidering this ostensible 'datum'. Unfortunately, the argument is complex, and a full-dress treatment would take a book. So for the present I will be content to provide a sketch of the argument and to mount a preliminary case for its main assumptions. I will be satisfied if by the end the reader finds it less obvious than he once did that we are normally entitled to our run-of-the-mill confident judgments of moral responsibility.

2. What is Moral Responsibility?

Just as legal responsibility amounts to liability to legal sanctions (monetary damages, criminal punishment), so moral responsibility, as I understand it, amounts to liability to 'moral sanctions'. However, this general formula covers a number of distinct phenomena. When your dog escapes by an improbable route and bites a neighbor, despite your having taken every reasonable precaution to keep him in the house, there is a sense in which you are responsible for the injury. He's your dog, after all, and if someone has to bear the costs of his malignity, it should be you. You should apologize and pay the doctor's bill—that's one sort of 'sanction'. It may even be that you should feel badly about the incident, not simply because your neighbor was bitten by a dog, but more specifically, because it was *your* dog that bit him.² This emotional self-laceration is another sort of 'sanction'. And yet if it is clear that you had taken every reasonable precaution to prevent the injury—if it is clear that you were neither negligent nor reckless nor malicious in causing the harm—then there is another sense in which you are not morally responsible for the injury. *It's not your fault; you are not culpable. Your neighbor should not blame you for it—he should not blame anyone—and you should not blame yourself.*

In what follows we shall be concerned with moral responsibility in this second sense. Since I will be focusing exclusively on bad acts and their consequences, we may simply identify moral responsibility with culpability or blameworthiness. Skepticism about moral responsibility is thus the thesis that *confident positive judgments of blameworthiness are never justified.*

To say that X is blameworthy for A is to say that X is liable to blame for having done A. Blame itself is a complex phenomenon; but at its core it consists in a range of emotional responses. Very roughly, you blame X for doing A when you resent him or feel indignant towards him for having done it; you blame

yourself for doing A when you feel guilty for having done it. To say that X is *blameworthy* for doing A is to say that X is *liable* to a negative emotional response of this sort for having done A, or equivalently, that some such response would be *appropriate* or *fitting*.³ To say that a response is appropriate is not to say that it is justified all things considered. Appropriateness is an internal normative relation between an object or an action and a response. A joke is funny if and only if it is appropriate (in the relevant sense) to find it funny; but there may be any number of excellent reasons not to laugh at a funny joke. Similarly, there may be any number of excellent reasons, moral and otherwise, not to blame someone who is in fact blameworthy. (It might be counterproductive; it might be cruel.) And conversely, there may be excellent reasons for blaming someone who is not in fact blameworthy. (If blame plays an indispensable role in moral education or social control, then the costs of forswearing it may be unbearable, even if, as a matter of fact, no one is ever blameworthy in the intended sense.) The skeptic about responsibility takes no stand on whether it makes sense *all things considered* for us to blame one another for what we do. He may even concede that from the standpoint of morality we are sometimes entitled, and perhaps even required, to blame. His position is that be this as it may, confident positive judgments of *blameworthiness* are never justified.

3. Excuses

A theory of moral responsibility is an account of the conditions under which an agent is liable to blame for an act, or more generally, for an occurrence. And while I shall not propose a general theory of this sort, certain rudiments are tolerably clear and for the most part noncontroversial. For example, it is generally agreed that an act is blameworthy only if it is wrong, but that the converse does not hold. When you pull my chair out from under me just for laughs, there is no doubt that the act is wrong. But if you're only five years old, or if you mistakenly believed that I wouldn't mind, then even though the act is wrong,⁴ it may be a mistake for me to blame you for it. When I say that I have an argument for skepticism about moral responsibility, I do not mean that I have an argument for skepticism about morality. In fact I shall assume that *commonsense morality is an excellent guide to right and wrong*. My skeptic concedes that people are often vicious or cruel or inconsiderate, that they often do what they have no right to do, and so on. His characteristic claim will be, "I concede the act was wrong; but I suspend judgment on whether it makes sense to blame the agent for it."

Everyone agrees that there are cases in which a person is not responsible for the wrong he does. Commonsense morality may operate with a presumption of responsibility; but it allows that this presumption may be defeated in many ways. When Bloggs tells an embarrassing story about his wife in public, he acts badly. And until we hear more we are perhaps entitled to suppose that it makes sense for his wife to blame him. But when we learn that Bloggs is on a new

medication that has unforeseeably addled his judgment, we may be inclined to withdraw the claim of blameworthiness without withdrawing the moral claim that it was wrong for him to tell the story.

Let's call any consideration that defeats the standing presumption of blameworthiness an **excuse**. Ordinary morality recognizes an open-ended battery of excuses. The criminal law has incorporated some of them under standard rubrics—insanity, infancy, duress, mistake, and so on. But we should not assume in advance that every genuine excuse finds clear expression in the law or in folk morality. Certain philosophical positions amount to the claim that there exists a single global excuse that the law and common sense both fail to credit. Thus some writers find it plausible that a person should be excused whenever his act is ultimately caused or determined by factors beyond his control—a principle which, given determinism, entails that everyone is automatically excused—even though nothing in ordinary moral thought or practice clearly forces this principle upon us. The argument I want to discuss does not invoke this particular alleged excuse; it has nothing whatsoever to do with determinism or the traditional free will problem. Nonetheless, the structure of the argument is broadly similar. **The plan is to identify an excuse with a toehold in ordinary morality, but whose full force ordinary morality fails to register.** I will suggest that this excuse is nonetheless genuine: that it is clear upon reflection that when the excusing condition obtains, the agent is not culpable for his bad act. The crux of the skeptical argument is then the thesis that while there may well be cases in which this novel excuse is unavailable, in any particular case we cannot exclude the possibility that it is available. This will then entail the skeptical thesis that in any particular case a confident positive judgment of responsibility would be unwarranted.

4. Original vs. Derivative Responsibility

Let's begin with an elementary distinction. **Sometimes it is clear that a person is responsible for a certain action or event, but only because he is responsible for one of its causes.** Suppose Jekyll goes berserk and smashes up the china shop. Suppose it is clear that at the time he was completely (if temporarily) out of his mind. In any real case of this sort we might well suppose that the agent is off the hook; **extreme pathological incapacity normally functions as an excuse.** But now suppose we learn that our man went berserk only because he had ingested a drug designed specifically to boost the destructive impulse while disabling the normal capacity for self-control. There are then two possibilities. If someone slipped him the drug without his knowledge, then he presumably remains off the hook. **But suppose he ingested it deliberately.** Suppose he was a blocked novelist who wanted to know first hand how it really feels to lose control. In that case we would probably conclude that he is culpable for the damage after all. When he regains his senses, we might blame him for it. ("You idiot. Look what you've done!") But this judgment clearly depends on

our determination that he is independently culpable for the prior reckless act of taking the drug. We may be tempted to say that in general, **pathological incapacity constitutes an excuse**—that a person is not responsible for his act if at the time he lacked the capacity to control his conduct. But as the example shows, that is not quite right. It would be better to say that pathological incapacity excuses if, but only if, it is not the (foreseeable) upshot of prior culpable wrongdoing on the agent's part.

In the example, Jekyll is responsible for the damage only because he is independently responsible for taking the drug. His responsibility for the damage is thus an instance of **derivative responsibility**. In general, X's responsibility for A is derivative when X is responsible for A only because he is independently responsible for B—some prior act or omission. Responsibility that is not derivative is **original**.

Now it is plausible that derivative responsibility presupposes original responsibility. This is plausible, but it is not obvious. Someone might hold that there can be ungrounded backwards chains of responsibility—that X might be culpable for A on Friday only because he is culpable for B on Thursday, and culpable for B on Thursday only because he is culpable for C on Wednesday, and so on ad infinitum—though when it comes to ordinary human beings it is hard to see this as a genuine possibility. More plausibly, someone might hold that responsibility is a vague matter, and **that there can be derivative responsibility without original responsibility for the same reason that there can be chickens even if there was no original chicken**. This is in fact a serious possibility, especially given what I shall go on to say. But for now I propose to set it aside. I shall assume as a background principle—and this is the first substantive premise in the skeptical argument—that **derivative responsibility presupposes original responsibility**.⁵ More precisely, I shall assume that if X is responsible for A then either A itself is a locus of original responsibility or there exists such a locus somewhere in A's causal history. The crucial lemma in the skeptical argument will be that in any particular case we cannot determine with any confidence that this condition is satisfied.

5. Action from Ignorance

But this is to anticipate. The skeptical argument begins by calling our attention to an ordinary excuse. Suppose that there is arsenic in the sugar bowl. You ask for sugar. I put a spoonful of arsenic in your tea, and you drink. Am I responsible for your death by poisoning? It depends. If I knew about the arsenic, then perhaps I am. So let's suppose that I did not know. We then have two cases to consider. If the **mistake was altogether innocent** on my part—if it was due to some freak chemical transformation in the sugar or to the devious conduct of some third party—then it may be clear that I'm not responsible. Suppose, however, that the mistake was **not altogether innocent**. Suppose I negligently stored my arsenic and my sugar in identical unlabeled

containers in the kitchen and then simply forgot about what I had done. In that case I am plausibly culpable for the poisoning—though I am obviously less culpable than I would have been if I had done it deliberately; and this is so despite the fact that *at the time of action* I had no inkling that I was administering a dose of poison.

The example is an example of **action from ignorance**. The fact that the sugar bowl contained a deadly poison was a decisive reason for me not to do what I did. When I acted, I acted in ignorance of this fact. So, in the relevant sense, I acted from (a position of) ignorance. The notion of action done from ignorance is in fact an extremely complex one: some of the complexities will emerge in due course.⁶ But for now an intuitive conception will suffice. The point of the examples is to suggest a general principle governing such action, to wit:

If X does A from ignorance, then X is culpable for the act only if he is culpable for the ignorance from which he acts.

In both cases of ignorant action—the case of the innocent mistake and the case of the not-so-innocent mistake—when I put the arsenic in your tea I fail to know that my act is wrong and, hence, that there is a decisive moral reason not to do it; and I fail to know this because I am ignorant of the fact in virtue of which my act is wrong—as it might be, the fact that in performing the act in question I am subjecting you to an unjustifiable risk of gruesome death. When I have been tricked by a third party, it's not my fault that I don't know this; and that's why I am not morally responsible for the act. On the other hand, when my ignorance derives from my own gross negligence, the ignorance itself is palpably *my fault*. And in that sort of case the ignorance does not constitute an excuse.⁷ So to a first approximation we may say that (pertinent) ignorance excuses, but only when it is *non-culpable*. This is the excuse that the skeptical argument will seek to exploit.⁸

If this is right then we may formulate our first intermediate conclusion, viz., that *an action done from ignorance is never a locus of original responsibility*, or equivalently: our responsibility for what we do from ignorance is always a matter of derivative responsibility. For all we have said, I may be culpable for ignorantly (unwittingly) depositing the arsenic in your tea, and hence for your death by poisoning. But an act of this sort cannot be the *first* thing for which I am responsible. I am responsible for this sort of act only if I am independently responsible for something else, in this case, my failure to know what was in the sugar bowl.

6. Culpable Ignorance

Attention thus shifts to the question: When is a person culpable for failing to know what he should have known? In the example, the ignorance is culpable when it is the upshot of prior culpable negligence or recklessness—in this case,

storing the arsenic and the sugar in identical containers. The next premise in the skeptical argument holds that this is an absolutely general fact about culpable ignorance:

X is culpable for failing to know that P only if his ignorance is the upshot of some prior culpable act or omission.

A complete defense of this claim is beyond the scope of the present essay. But let me briefly sketch the picture that lies behind it. We are under a standing obligation to see to it that as we blunder through the world we do not create unjustifiable risks of harm to others. To violate this obligation is to act negligently (or perhaps recklessly or malicious if one is aware of the violation). These obligations of 'due care' are impossible to codify. The law of negligence is a sprawling mess largely for this reason. But even there one seeks in vain for an explicit account. Instead the law appeals to a heuristic device: A precaution is required, we are told, if a *reasonably prudent person in the agent's circumstances* would have taken it. It is worth stressing in this connection that the Man of Reasonable Prudence is not the Man of Maximal Prudence. You are never obliged to take *every possible step*, no matter how costly, to ensure that no one is harmed by what you do. You are required only to take certain reasonable steps. If you do that much and harm results anyway, then in the vast majority of cases the harm must 'lie where it falls'. The law allows certain exceptions to this principle. This is the domain of so-called strict liability. But in morality there is no such thing as strict liability. **If you have taken every reasonable precaution and someone is harmed anyway, the harm is simply not your fault; you're not responsible for it; you're not to blame. No one is. It was just an accident.**

Now among the required precautions against negligent harm are certain **epistemic precautions**. 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**. Again, they are impossible to codify. But again, the person of ordinary prudence provides a serviceable heuristic. In any given case we can ask whether the agent's ignorance derives from a failure to do what any reasonably prudent person in his circumstances would have done in order to see to it that he was adequately informed. As I understand them, these procedural obligations are always obligations to *do* (or to refrain from doing) certain things: to ask certain questions, to take careful notes, to stop and think, to focus one's attention in a certain direction, etc. 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.**

The skeptical argument assumes as a premise that *whenever you are culpable for failing to know some pertinent truth, this is because you are culpable for failing to comply with one of these procedural epistemic obligations*. Ignorance is culpable only if it derives from culpable recklessness or negligence in the management of one's opinion. Take any case of action done from ignorance and let it be stipulated that the agent has been utterly scrupulous in policing his own opinion: he has been as careful and as inquisitive and as reflective as a person in his circumstances should be⁹, and yet he has failed to grasp some crucial fact. I claim that if you bear this stipulation clearly in mind, you will be persuaded in every case that the agent's ignorance is not his fault.¹⁰

While it is not immediately relevant to the larger argument, I will offer a conjecture as to why this should be. The conjecture begins with the observation that in the normal case, *belief revision is a passive matter*. I can take various active steps to influence the content of my opinion: I can consult my guru; I can read a book; I can sit and think about a question. *But once I have taken these active measures, the change in view itself is something that simply happens in me or to me. It is not something I do.*

I conjecture that the principle about culpable ignorance follows from a more general principle about culpability for passive matters. Roughly: 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. *I can be responsible for falling asleep at the wheel. I can be responsible for forgetting your name.* But I am responsible for such things only if they result from some prior culpable act or omission on my part—taking sleeping pills before the long drive, for example, or neglecting to say your name over and over to myself when we were first introduced.

One might wonder why this more general principle should hold. Why shouldn't I be morally responsible for things that *happen in me or to me*, though not as a result of my culpable failure to take precautions? A theorist for whom the urgent question in judgments of responsibility is whether the act or event in question somehow expresses the agent's *character* will of course wish to resist this principle. As Aristotle stressed, the manifestations of character are often passive. In my view, however, this very phenomenon poses a serious challenge to such theories. Suppose Fred finds himself momentarily wracked with *schadenfreude* when he discovers that his rival's book has been remaindered. Since he knows that it's wrong to take pleasure in someone else's misfortune, he instantly resists and (let us say) stifles the emotion. We can still ask whether he merits blame for the initial feeling. Let us grant that it manifests a defect in his character. A better man would not have had to resist the ignoble impulse. But now stipulate that hitherto *Fred has been neither reckless nor negligent in the management of his character*; he has done whatever it is that a person is supposed to do (which may be nothing at all) in order to prevent this minor vice from taking up residence in his personality. *Then I say that it's simply*

obtuse to hold him responsible for his reaction.) He should not feel guilty; his rival should not resent him, etc. (Perhaps he should be ashamed of himself, but that's another matter.) Of course none of this explains *why* we should not be responsible for what simply happens to us in the absence of prior culpable wrongdoing. It is just to suggest that the principle is independently compelling.

To return to the main thread, if what we have said thus far is sound, it follows that that *our responsibility for what we do from ignorance is always doubly derivative*. The agent is responsible for the ignorant act only if he is responsible for the ignorance from which he acts; and he is responsible for the ignorance only if he is responsible for some prior failure to discharge one of his procedural epistemic obligations.

This looks like a recipe for a regress, and so it is, to a point. Suppose a surgeon orders that her type A patient be transfused with type B blood, that she does this only because she is mistaken about the patient's blood type, and that she is mistaken about the blood type only because she neglected to double-check his chart (which had just been updated) immediately prior to surgery, as standard practice requires. Then we know this much. The surgeon is culpable for the bad transfusion (and the ensuing harm) only if she is culpable for her ignorance as to the patient's blood type, and she is culpable for her ignorance only if she is culpable for her negligent failure to double-check the chart. Now focus on *this* negligent failure, which is of course the failure to comply with a procedural epistemic obligation. It is plausible that in this case, as in most similar cases, the negligent failure will itself be an act done from ignorance. When the time is ripe for the precaution, the agent will fail to take it only because she does not then think—and so does not then *know*—that she ought to be taking it. But if the failure is thoughtless in this sense then our principles entail that the agent will be culpable for it only if she is culpable for the ignorance that underlies it. Our principles then further entail that she will be culpable for *this* bit of ignorance only if she is culpable for omitting some required precaution to prevent *it*—e.g., if she is forgetful, asking one of her colleagues to remind her to check the chart.¹¹ And here the same series of questions will arise again. At this point we cannot rule out the possibility that the regress will somehow terminate. (We will see how this is possible in section 7.) For now the point is simply to convey how complex and arcane the inquiry can be when we set out to determine, in light of our principles, whether an agent is responsible for an action done from ignorance.

It may be worth noting that the law of negligence neatly *short-circuits this arcane inquiry*. If the surgeon's failure to double-check the patient's chart constitutes negligence—if any reasonably prudent doctor in her position would have checked—then she is liable in tort for the harm caused by her negligence. The law of torts does not care whether the doctor is *culpable* for her negligent omission. A fortiori, it does not care whether she is culpable for the ignorance that underlies it. There may be a compelling moral argument that in the circumstances it would be a mistake to *blame* the doctor for her failure to

double-check the chart. Perhaps there has been a deliberate conspiracy to keep her from knowing that such precautions are customary. Perhaps her failure was due to a small and otherwise imperceptible stroke that annihilated the neurons encoding the relevant information. It doesn't matter. The surgeon is liable if she departed from an objective standard of reasonable care. Does this show that the law of torts is at odds with our philosophy? Not at all. It is a commonplace that tort liability has nothing to do with moral blameworthiness. **The question in torts is always, Who should bear the cost of this accident?** When the choice is between an innocent victim and a negligent defendant, the law generally shifts the cost to the defendant. There are a number of competing rationales for this aspect of the law. But none of them depends on the presumption that the defendant is morally culpable for his negligence. An adverse tort judgment does not express moral blame in our sense; this aspect of the law is therefore fully compatible with the present view.

6. Normative Ignorance

We have considered a number of principles concerning action done from ignorance. In the examples, for the most part, the ignorance has been ignorance of *fact*. The unwitting host does not know *that there is arsenic in the sugar bowl*. The surgeon does not know *that the blood types do not match*. Nonetheless, I maintain that *our principles hold in full generality, no matter what the content of the agent's ignorance may be*.¹² It might be ignorance of law; it might be ignorance of moral principle. It might be ignorance of the rational force of moral principles or ignorance about what is to be done all things considered. In each case, whenever ever the agent acts from pertinent ignorance, **he is culpable for the act only if he is culpable for the ignorance from which he acts, and he is culpable for the ignorance only if it is the upshot of prior culpable negligence or recklessness in the management of his opinion.**

Now there is a sense in which the examples we have considered already involve normative ignorance. The surgeon does not simply fail to know that the blood types do not match. She also fails to know *that she should not order a transfusion*. But this normative ignorance obviously derives from underlying factual ignorance. My substantive claim at the moment is that the principles we have articulated hold good even when the normative error in question is not the upshot of factual error. Cases to support this point tend to be far out. Thus we might consider the ancient slaveholder who has no false factual opinions about his slaves but nonetheless believes that it is morally permissible for him to buy and sell them, to force them to work and so on.¹³ (Our principles then entail, not that the ancient slaveholder is off the hook, but rather that **he is on the hook only if he is culpable for the moral ignorance from which he acts.**) In fact, however, the phenomenon of nonderivative normative ignorance is ubiquitous in daily life, and in my view reflection on these cases supports the general reading of our principles. The best examples

are line drawing cases: cases in which the agent is aware of every relevant factual consideration and its valence (whether it counts for or against the contemplated act), but is nonetheless mistaken about the weight such considerations should be given. Consider, for example, the ambitious capitalist who is mistaken about where to draw the line between permissibly aggressive business practice and reprehensibly ruthless business practice. He knows that if he outsources the accounting department, he will expose a significant number of loyal employees to profound hardship; but he also knows that the move will improve the bottom line; that his shareholders have certain legitimate expectations, etc. Suppose that as a matter of fact the marginal improvement in the bottom line does not justify the profound harm to his employees, but that our capitalist, after considering the question, fails to see this. It is not hard to imagine a case of this sort in which (a) our capitalist does something seriously wrong without knowing that he is doing something wrong, but in which (b) he knows all the relevant non-normative facts, and (c) he has been adequately reflective, careful, etc. Perhaps the people who taught him how to think about such cases taught him badly. Perhaps it's just a hard case and after thinking seriously about it for a decent interval he has simply arrived at the wrong answer. I contend that if you are careful to bear in mind the stipulation that in reaching his conclusion our capitalist has not been reckless or negligent in the management of his moral opinion, you will find it plausible that his moral ignorance is not his fault. I further contend that if you bear *this* thought fully in mind, you will find it equally compelling that it would be a mistake to blame him for the wrong he does.¹⁴

The most daring claim I will make in this connection is that our principles hold good even when the ignorance in question is what might be called *bare ignorance of the rational force of moral considerations*. Suppose that Bill has gotten himself into trouble and that he is considering a self-serving lie. Unlike the ambitious capitalist in the previous example, Bill knows that the act he is considering would be wrong: he knows that it's just plain wrong to lie to your wife about where you've been. The trouble is that he also knows that if he tells the truth, he will suffer. Let's stipulate that as a matter of fact, the case against the lie is the stronger case, and hence that all things considered Bill should tell the truth and face the music. This is the correct deliberative conclusion—the correct answer to the practical question, All things considered, what should Bill do? But of course there is no guarantee that Bill will reach it. Suppose he doesn't. I claim first that it is conceivable that Bill might be blameless for this error. This is easy to imagine if we suppose that his deliberations have been distorted by an undetected brain anomaly or the like—some blameless condition that makes it seem to him as he is deliberating that this is one of those cases in which the thing to do is to pursue self-interest and to flout the requirements of morality.¹⁵ But we can also imagine (once again) that he has been badly taught. Perhaps he has been led to believe, through no fault of his own, that while moral considerations have some weight, they are not in general decisive: that moral

costs may legitimately be outweighed by non-moral benefits in sound rational deliberation about what to do.

Now suppose that in the grip of some such view, he tells the self-serving lie. What is the appropriate response? Suppose that it's clear upon investigation that in arriving at the verdict, 'All things considered, I should lie,' **Bill has been blameless in every procedural respect.** He has been as thoughtful and reflective and curious as a person should be in the circumstances. Our principles then entail—correctly, in my view—that it is not Bill's fault that he believed that all things considered, he should lie. Now ask whether, given this ignorance, it makes sense to blame him for the lie. The case against blame strikes me as compelling. Rewind the tape to the period just prior to the action. There is Bill, blamelessly (though mistakenly) convinced that the balance of reasons comes down in favor of lying, and blamelessly convinced, we might add, that he has no reason to reconsider this verdict. What do you expect him to do? Tell the truth? It is of course *possible* for a person to act contrary to his reflective judgment about what he has most reason to do. To do so is to exhibit one form of what philosophers sometimes call *akrasia*.¹⁶ We may hope that Bill will act akratically in this case. But is it reasonable to demand or expect that he do so? **Does it make sense to subject someone who blamelessly believes that he should do A, and then does it, to moral sanctions—to recrimination, resentment, righteous anger, contempt?** Focus on the case in which the verdict is the result of a manifestly blameless condition like a brain glitch or bad training, and the force of these rhetorical questions should be apparent.¹⁷

The suggestion, in effect, is that Bill's wife should take the following philosophical view of the matter: "Poor Bill. Through no fault of his own he found himself believing that all things considered, he should lie. Given that he found himself in that state, I can hardly fault him for lying. Holding the judgment fixed, the lie itself was a perfect manifestation of practical rationality. I can fault him for the lie only if I can fault him for believing that in the circumstances, his selfish interests were more important than my moral interests. Since by hypothesis, it is not his fault that he held this view, I have no option but to conclude that he is not properly culpable for his bad action."

7. Akritic action and original responsibility

The rubric 'action done from ignorance' is normally supposed to cover a restricted (if significant) range of cases—cases like the arsenic case in which the agent's practical deliberation involves a factual mistake. It is a crucial premise in the skeptical argument that its proper scope is ultimately much broader. The ambitious capitalist who acts without knowing that he is doing anything wrong acts from ignorance in our sense, as does the self-serving liar who knows that it is wrong to lie and yet honestly believes that all things considered his lie is warranted in the circumstances.

We need not speculate about how much of the bad action we ordinarily encounter amounts to action from ignorance in this broad sense. But it seems clear that much of it must be. **Let us make the simplifying assumption that when an act is morally wrong, there is always a decisive reason for the agent not to do it.** This high-minded position is probably too high-minded. But we may be high-minded enough to think that the exceptions are rare and that ordinary bad behavior is almost never an exception. That is to say that in ordinary cases, the agent who acts badly almost always has a decisive reason to do something else.

Now suppose that some agent has settled on some bad act of this sort as a result of explicit practical deliberation. He has asked himself the practical question, 'What should I do?' and the answer has come down: 'The thing to do is to rob the liquor store'. If he goes on to do the deed, then by hypothesis he acts from ignorance. **It might be ignorance of fact, or ignorance of moral principle. But at a minimum he acts from ignorance about what is to be done all things considered.**

Alternatively, we may suppose that he did not deliberate, or at any rate, that he had no clear view before he acted about what was to be done all things considered. Perhaps he entertained some of his options, noted some of the considerations pro and con, and then simply chose one of them in response to some of the relevant considerations without ever coming to the view that his chosen option was the option that he had most reason to choose. **Well, then again he acts from ignorance. By hypothesis, he does not believe that he has good reason to do something else. A fortiori, he does not know it.**

What would it take for an agent to act badly *without ignorance*? Just this: He would have to know the pertinent facts about his contemplated act. He would have to know that it was wrong. And he would have to know that in the circumstances, all things considered, he should not do it. He would then have to act despite this knowledge. As we have noted, any such act will have the structure of an akratic act. It will be done despite the agent's considered judgment that all things considered he should be doing something else.

This points to a striking intermediate conclusion. Recall our observation that responsibility for action done from ignorance is invariably a matter of derivative responsibility: One is responsible for the act done from ignorance only if one is independently responsible for something else. On the present broad conception of action done from ignorance, this entails that *the only possible locus of original responsibility is an akratic act*. In weakness begins responsibility. Our first sin must be a knowing sin—a sin done in full knowledge of every pertinent fact or principle.

And this in turn entails that every culpable bad action must be, if not itself a knowing sin in this sense, then at least an act whose etiology involves a knowing sin. ***Every culpable bad action must be the causal upshot of a genuinely akratic act or omission.*** Anyone who bears our principles in mind and nonetheless judges that Bill is responsible for lying to his wife, is thereby committed to the view that

somewhere in the story of that lie there exists a full-blown episode of altogether knowledgeable wrongdoing.

8. The skeptical upshot

These considerations do not show that no one is ever responsible for the wrong he does. For all we have said, akratic bad action is possible. And to the extent that it is actual, it is available to terminate the regress we have identified. If there were an independent argument for the thesis that akratic action is never culpable, then we would indeed have a general argument for the impossibility of responsibility. But so far as I know, there is no such argument. The akratic agent cannot plead ignorance; by hypothesis, he does not act from ignorance. Nor in general can he plead incapacity; there is no reason to suppose that in general the akratic agent lacks the power or ability to do the right thing in the circumstances. In the absence of a complete account of the excuses, we cannot conclude with perfect confidence that the typical akratic agent is on the hook. But at this point we have no reason to doubt that he is. We therefore have no reason to believe, from the considerations so far adduced, that moral responsibility is impossible.

Nonetheless I claim—and this completes the skeptical argument—that in light of what we have said, *it would be unreasonable to repose much confidence in any particular positive judgment of responsibility.* What must you think in order to judge that Bill, for example, is responsible for lying to his wife? You must think that at the time of action, either he knew that he had decisive reason not to lie, or if he did not know this, that his ignorance was the upshot of some prior bad action done in full knowledge of every pertinent fact or norm. *You must think, in other words, that his bad action either is, or derives from, an episode of genuine, full-strength akrasia.*

I suggest that given the opacity of the mind—of other minds and even of one's own mind—it is almost always unreasonable to place significant confidence in such a judgment. Sometimes the relevant facts will be straightforwardly inaccessible. The surgeon neglects to check her patient's chart; this is an act done from ignorance. Is she culpable for the ignorance? That depends on whether it derives from some prior culpable act or omission. But how are you (a third party) supposed to approach that question? Perhaps you can identify the prior omission—perhaps you can show that given her track record on such matters, she should have asked a colleague to remind her to check the chart. Still, the question will be whether her failure to do so was itself a *culpable* failure; and that will depend on whether this episode was itself an akratic episode—that she failed to ask for the reminder even though she knew that she ought to ask for it—or traces back to an akratic episode. *And it seems obvious that in any real case it will be impossible to resolve this question with any confidence.*

It is possible to reinforce this line of thought with a more general ground for doubt. The agent is culpable for his bad action only if that bad action is, or

derives from, an episode of genuine akrasia. But genuine akrasia in this sense is extremely difficult to identify. The reason is that it is not readily distinguishable from an impostor: ordinary weakness of will.¹⁸ 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 persuades himself (or finds himself persuaded) that the preferred alternative is at least as reasonable. Moreover, in between these two pure cases there is a continuum of cases; cases in which the agent *suspends* his original judgment without quite rejecting it; or cases in which it is *simply indeterminate* as the agent acts whether he in fact believes that all things considered he should do A. (Real action almost never involves an *explicit* practical judgment; and it is not hard to imagine that it may be indeterminate whether, in the absence of such a judgment, the agent is to be credited with an implicit belief that A is to be done.) If X is to be (determinately) responsible for his bad act, then that responsibility must be grounded in a (determinate) case of genuine akrasia; the impostors—mere weak bad action and its cousins—will not suffice. The skeptical suggestion is that as a matter of fact we are hardly ever—I might say, *never*—in a position to tell the difference with any confidence.

This verdict is based mainly on reflection. When I consider my own case and ask whether some weakish act of mine amounts to genuine akrasia as opposed to ordinary moral weakness, I have no trouble identifying tolerably clear cases of the latter; but I confess that I cannot identify clear examples of the former with any confidence. This is so even when I ask myself the question *as I am acting*—a rare exercise, to be sure. The difficulty is even more vivid when I ask the question of my own actions in retrospect, and still more vivid yet when I ask about the actions of other people. If I ask whether I can identify an episode of genuine akrasia as distinct from simple weakness in, say, my wife or a close colleague, I can only report that I come up with nothing, or at best, with a tentative conjecture.

So here is the difficulty. The culpable bad actions have a distinctive sort of causal history—an *inculcating* history—in which the act either is, or derives from, an episode of genuine akrasia. There may be no obstacle in principle to identifying such episodes. God could manage it; a super-psychologist might manage it. But given the real limitations on our access to the causal histories of human actions and to the states of knowledge and opinion that underlie them, I claim that as a matter of fact we are never entitled to any significant confidence that the bad act under consideration satisfies the necessary condition we have identified. And this is just the thesis of skepticism about moral responsibility.

The suggestion, in effect, is that whenever we consider a particular bad action, we cannot rule out the possibility that the agent should be excused on the ground that he acts from non-culpable ignorance. We cannot exclude the possibility that the agent acts from ignorance because we cannot conclude,

with any confidence, that his act is an instance of genuine akrasia. And we cannot exclude the possibility that his ignorance is non-culpable because we cannot conclude, with any confidence, that it derives, in the right way, from an act of genuine akrasia.

I close with two comments on the general character of the argument. First, it should be clear that unlike the standard grounds for doubt about moral responsibility, the argument involves no assumption about the possibility of genuine free action under determinism (or indeterminism). One way to bring this out is to note that the argument is compatible with the assumption of full-blown agent causation. **We can assume that in some cases the agent who acts badly is not determined so to act by prior causes or states of the world, but that when he acts *he himself* is the cause of his action (whatever this might mean).** It doesn't matter; we can still ask whether we can exclude the possibility that the act was done from non-culpable ignorance. And if we cannot, then the skeptical conclusion will be forced upon us.

Second, we may note that unlike the more familiar grounds for doubt about moral responsibility, the present skeptical conclusion applies only to cases of *bad* action. Classical grounds for doubt about responsibility aim to call into question the propriety of *attributing* the act to the agent, and as is familiar, any such argument, if sound, impugns our judgments of praiseworthiness along with our judgments of blameworthiness. The present argument does not have this feature. While there is good reason to suppose that in many cases the agent who acts badly acts from ignorance—and is thus a potential candidate for the excuse we have identified—there is no reason to suppose that in general the agent who does the right thing acts from ignorance. So there is no reason to suppose that the present skeptical argument should generalize in the familiar way.

In my view, these observations lend indirect support to the skeptical argument. There is a perennial tendency in philosophy to suppose that skeptical arguments can amount at best to puzzles, and this may be true when it comes to arguments that purport to show that confident judgments about the external world, or about other minds, or about the future, are never warranted. That is in part because in these cases the skeptical posture at issue—e.g., thoroughgoing suspense of judgment about the external world—is not a serious option for us. Strawson famously seeks to assimilate skepticism (or rejectionism) about responsibility to such a case.¹⁹ **His central thought, for present purposes, is that genuine suspense of judgment on questions of responsibility would amount to adopting the 'objective stance' towards other human beings—a stance in which one prescind from a vast range of ordinary emotional responses to human action and treats other people as one normally treats animals or lunatics**—as entities that require management or treatment rather than ordinary interaction. Strawson rightly argues that this revision is not only undesirable; it is impossible. And if this were the inevitable upshot of the skeptical argument, this might be enough to warrant assimilating skepticism about moral responsibility

to philosophical skepticism in other areas. From the present standpoint, however, this assimilation is premature. The present argument is compatible with the view that normal adults are for the most part rational; that they normally have the capacity to control their conduct; that their actions normally reflect their choices; that their choices are normally sensitive to their conception of what they have most reason to do, and that this conception is to some extent sensitive to the reasons that in fact exist. Moreover it is compatible with the thought that the positive reactive attitudes—gratitude, pride, admiration, and the like—are for the most part warranted. There is thus ample scope within the skeptical posture here endorsed for a distinction between the attitude we adopt with animals and lunatics, and an engaged stance we might adopt with normal competent adults. This is not to deny that the skeptical view has revisionary implications. Someone who entertains serious doubts about whether his friends and neighbors are blameworthy for their transgressions should presumably stop blaming them; and that would be a change. If he is concerned with such matters he should presumably also take steps to revise or reconceive any social practice whose justification depends on the assumption that people are morally responsible for their bad actions—one salient candidate for which is the practice of criminal punishment. Whether this would amount to a radical revision—an impossible revision—must remain at this point an open question.²⁰

Notes

1. Skepticism about responsibility in my sense does not extend to *negative* judgments of responsibility. So far as the present argument is concerned, we may judge with utter confidence that *Aristotle* is off the hook for the bloody death of Belle Elmore.
2. This would be a vicarious variant of the sentiment Bernard Williams calls 'agent regret' ('Moral Luck', in his *Moral Luck: Philosophical Papers 1973–1980*, Cambridge University Press, 1980, p. 27). In my view it is a mystifying sentiment, but if it is ever appropriate I see no reason not to say that anyone who is liable to agent regret for an occurrence is to that extent, in one sense, morally responsible for it.
3. For an extensive discussion of these reactive 'moral sentiments' and the thesis that moral responsibility is to be understood in terms of them, see R. J. Wallace, *Responsibility and the Moral Sentiments*, Harvard University Press, 1994.
4. There is a tendency to resist this verdict—to insist that in these cases the act was not wrong after all. In my view, this tendency should be resisted. It is not that small children have a right to their malicious pranks. It is rather that given their immaturity, they are not fully culpable for the wrong they thereby do. Likewise, it is not as if someone who mistakenly believes that his victim would consent to an otherwise impermissible act is thereby permitted to do it. It is rather that (in certain cases) the fact that the agent honestly believed that his act was permissible may entail that it would be inappropriate to blame him for it.
5. Henceforth, important premises and intermediate conclusions are in italics.

6. Very roughly, X does A from ignorance if, at the time of action, for each decisive moral reason against doing A, X is either ignorant of the reason or ignorant that it is a decisive reason. I am grateful to Jessica Boyd, Peter Graham and Barry Lam for discussion on this point. All three have important unpublished papers on the topic.
7. Note the analogy with the Jekyll case. A theorist may be tempted to identify certain 'synchronic' conditions as excuses—conditions that obtain at or about the time of action, including various forms of incapacity and ignorance. But in every case this sort of account is strictly inaccurate. Incapacity and (relevant) ignorance do not always excuse; they excuse only when they are not the (foreseeable) upshot of prior culpable wrongdoing on the agent's part.
8. We sometimes mark this distinction by saying that in the latter case, while I did not know at the time that my act was wrong, nonetheless, I *should have known*. However, this idiom can be misleading. Just as we distinguish between the claim that an act is wrong—i.e., the claim that the agent should have *done better*—and the stronger claim the agent is culpable for it, so we should distinguish between the claim that the agent should have known some fact or principle—i.e., the claim that the agent should have *known better*—and the stronger claim that it is the agent's *fault* that he did not know it. These two judgments come apart whenever the agent has an excuse for his failure to know. As will emerge in what follows, it is clear that for present purposes, what matters is not whether the agent *should have known* the fact in question, but rather whether he is *culpable* for his failure to know it.
9. Which is not necessarily to say that he has been as scrupulous as he *could have been*. As above, reasonable prudence is one thing, maximal prudence another.
10. A complete defense of the principle would have to show, in addition, that when the agent's ignorance does derive from the violation of a procedural obligation, the agent is culpable for his ignorance only if he is *culpable* for the violation. The crucial test cases are those in which the agent's failure to comply with his procedural obligations is excused, e.g., by non-culpable incapacity or ignorance. See the discussion of the Surgeon's case below.
11. If she is not normally forgetful then it may be that there was *no* precaution that she was required to take to ensure that she would remember to check the chart. (Normally we are entitled simply to rely on our practical memory for such things.) In that case, our principles entail that when she does forget, the episode of forgetting is non-culpable, as is the ignorance that results.
12. In 'Culpability and Ignorance,' *Proceedings of the Aristotelian Society* 103:1 (2003), I call this the 'Parity Principle'.
13. See 'Culpability and Ignorance', §IV.
14. 'Culpability and Ignorance,' §V.
15. Compare the (unnecessarily hyperbolic) case of Bonnie in 'Culpability and Ignorance,' §IX.
16. 'Akrasia' here is a placeholder for the circumstance in which X does A despite the persisting judgment that all things considered, it makes sense to do something else. I certainly do not claim that this is the only phenomenon that deserves the name.
17. The point of specifying (even if vaguely) a non-rational causal basis for the error is not to encourage the thought that the agent's mistake was somehow inevitable

or beyond his control. That is neither here nor there. The point is simply to make it vivid that a mistake of this sort might arise even in agent who has taken every reasonable step to prevent such errors.

18. For the contrast, see Richard Holton, 'Intention and Weakness of Will', *Journal of Philosophy*, 96 (1999) pp. 241–62.
19. 'Freedom and Resentment', reprinted in G. Watson, ed., *Free Will*, Oxford University Press, 1982.
20. Work on this paper was supported by a Mellon New Directions Fellowship. I am grateful to the Hauser Program in Global Law at New York University for its hospitality, and to audiences at Amherst College, the City University of New York and NYU Law School. I owe a special debt to Jessica Boyd, Peter Graham and David Sussman for extensive discussion of earlier versions of this material.