



**UNIVERSITY OF  
ILLINOIS PRESS**

North American Philosophical Publications

---

KNOWING ABOUT RESPONSIBILITY

Author(s): Simon-Pierre Chevarie-Cossette

Source: *American Philosophical Quarterly*, JULY 2021, Vol. 58, No. 3 (JULY 2021), pp. 201-216

Published by: University of Illinois Press on behalf of the North American Philosophical Publications

Stable URL: <https://www.jstor.org/stable/10.2307/48616056>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*North American Philosophical Publications* and *University of Illinois Press* are collaborating with JSTOR to digitize, preserve and extend access to *American Philosophical Quarterly*

# KNOWING ABOUT RESPONSIBILITY: A TRILEMMA

Simon-Pierre Chevarie-Cossette

## ABSTRACT

Here is a skeptical trilemma. We know that we are responsible. We do not know that determinism is false. Yet, if we do not know that determinism is false, we do not know that we are responsible. If we want to avoid the skeptical conclusion of rejecting the first claim, we ought to reject the third. It does not follow that compatibilism is true. But it follows that libertarianism, the view that we are responsible even though responsibility requires indeterminism, is unknowable, at least for now. To see this, we need to compare different forms of skepticism and their response.

The best argument for compatibilism is that we know better that we are sometimes free than that we ever escape predetermination; wherefore it may be for all we know that we are free but predetermined.

Lewis 2000, 109

## I. INTRODUCTION

Free will skeptics claim that moral responsibility does not exist, in the sense of basic desert or liability for praise or blame (see e.g., Pereboom 2014).<sup>1</sup> As a consequence, they claim that we do not *know* that moral responsibility exists, since knowledge is factive. It is a rarer thing to see them argue for this second claim directly, without invoking the factivity of knowledge. But they could (and they have, under the banner of “agnosticism”<sup>2</sup>). After all, they are *skeptics* and skeptics like to doubt what others take to be knowledge.<sup>3</sup> They could say:

I don’t know that determinism is false. (A)

If I don’t know that determinism is false, I don’t know that I am ever morally responsible for my conduct. (B)

*Therefore:*

I don’t know that I am ever responsible for my conduct.<sup>4</sup> (~C)

Here, “determinism” refers to the view that every event is determined (or necessitated) by antecedent (or future) conditions together with the laws of nature (see Hofer 2016).

This argument echoes the standard skeptical argument in epistemology: I don’t know that I am not in the bad case (being systematically deceived by a Cartesian Demon, living in a deterministic world); if I don’t know this, I don’t know something I took for granted (that I have hands, that I am ever morally responsible for my conduct); therefore, I don’t know that thing I took for granted. And, like the standard skeptical argument in epistemology, it suggests an inescapable trilemma: reject (A), reject (B), or accept (~C), that is reject (C).

In the debate about moral responsibility and determinism, there are three main clans: the libertarians, who claim that moral responsibility exists and is incompatible with determinism; the non-skeptical compatibilists (henceforth compatibilists), who claim that moral responsibility exists and is compatible with determinism; and the skeptics, who claim that moral responsibility does not exist.<sup>5</sup> Rejecting (A) is the natural home of the libertarians. For if they know that moral responsibility is incompatible with determinism and if they know that they are responsible, they should be able to know that determinism is false. Rejecting (B) is the natural home of the compatibilists. For if determinism and responsibility have nothing to do with one another, it is natural to think that knowledge of the one will be unrelated to knowledge of the other. (As we shall see later, things are more complicated.) Rejecting (C) is, as we said above, the skeptic's take.

This argument has an advantage and a disadvantage for the free will skeptic. The advantage is that it bypasses difficult questions related to *indeterminism*. Skeptics who use this argument do not need to argue that indeterminism is compatible with responsibility; in particular, they do not need to show that there is no middle ground between chance and determination. The disadvantage for the skeptic is that the argument does not show that we are *not* responsible. Still, its conclusion is serious. If we do not know that we are responsible, can we really just act *as if* we were? Shouldn't we be more prudent regarding matters as important as responsibility?

Some have responded<sup>6</sup> that we had better just act as if we were responsible because of practical considerations; because, without a belief in responsibility, our lives would be miserable.<sup>7</sup> But this option is not obviously right. The skeptical life is not unappealing;<sup>8</sup> and the thought of potentially living under a

permanent and pervasive illusion is gloomy.<sup>9</sup> So, it is important to look elsewhere, that is to respond to the skeptic by rejecting (A) or (B).

It is worth emphasizing why. We not only care about being responsible, but about knowing that we are. And we care not only about whether responsibility is compatible with determinism, but about whether we know this.<sup>10</sup> In fact, some philosophers have maintained that knowledge matters because it allows us to be guided by the facts and that this guidance is a fundamental desire of human beings (see e.g., Hyman 2015, chap. 8). Others have argued that in general we ought to act in accordance with what we know (see e.g., Hawthorne 2004, 21–32). For instance, it would be a mistake to blame someone on the premise that she is responsible if we do not *know* that she is. At least, if we think that we are ignorant about responsibility, we have good reasons (and perhaps a moral obligation) to modify our conduct and attitudes (see Vilhauer 2009). So even though most of the literature on free will and moral responsibility focusses on moral or metaphysical questions such as, “which theories of moral responsibility allow its existence and are these theories true?” we should also ask, “which theories of moral responsibility allow knowledge of its existence and can we know these theories?”

The trilemma we just considered offers a humble but significant start to this project. In this paper, I consider in turn rejecting (A) and (B) and argue in favor of the second. For reasons of space but not of antipathy, I leave free will skepticism aside here—and thus the option of rejecting (C). Every non-skeptic, including the libertarian, is likely to accept that we do not know that determinism is false and nevertheless insist that we know that we are responsible. But this comes at a cost for libertarians because it shows that their doctrine is unknowable given the actual state of science. If it were knowable, as I shall argue,

we could come to know determinism without the help of physics. And that is absurd. So, by mapping out the available terrain for non-skepticism about free will, we discover that it is to the compatibilist's advantage.

In § 2 and § 3, I consider the two options, rejecting (A) or (B), and argue for the second. Rejecting (B) has the consequence that libertarianism is unknowable. In § 4, I argue that this consequence is unavoidable given that we accept (C). In § 5, I discuss some dialectical points. In § 6, I come back to the parallel between the two skepticisms (regarding free will and the external world) and consider an objection from van Inwagen (1983), which has recently resurfaced (see Steward 2020).

## 2. FIRST WAY: KNOWLEDGE OF INDETERMINISM

The first way out of the trilemma is to reject (A): that is, to claim that we know determinism to be false. This, I hope to show, is impossible for us at the moment. As I have noted, this is a strategy which is mainly appealing for the libertarians, who claim that we are responsible even though incompatibilism is true. It is mainly appealing for them because rejecting (C) seriously constrains the relevance of their doctrine; and rejecting (B) has bad consequences for them, which we will explore in § 3.

On what grounds can we reject (A)? Physics or other empirical sciences do not yet give us a clear answer about whether determinism is true. Some interpretations of quantum mechanics are indeterministic, some are not (see Hoefer 2016, section 4). Crucially, even if in the next few years, a clearer consensus emerges in favor of indeterminism, it will not follow that determinism is false *in the right way* for the libertarian. This is a fairly common point. If indeterminism is required for responsibility, this is because at least some of our mental events are undetermined. Further, if indeterminism is required for

responsibility, this is because *the right kinds of mental events* are undetermined (perhaps self-forming willings<sup>11</sup> or agent-caused decisions<sup>12</sup>). If, for instance, the occurrence of emotions is undetermined, this does not seem to give us any reason to admit the existence of responsibility. Now, we should remain healthily skeptical about the prospect of physics, biology, and psychology revealing that undetermined events occur when and where specific actions are performed. What this means is that *if* physics reveals that determinism is false, the sceptic can rephrase the trilemma by replacing “determinism” with something like “mental determinism.” So, let us ask again: on what grounds should we reject (A)? “Science” is not an answer for now. Perhaps one day philosophers will be able to claim that they know that the right kind of indeterminism is true; and that this allows<sup>13</sup> them to solve our trilemma by granting (B) and (C) and rejecting (A).

Until then, the non-sceptics (in particular the libertarians) might be tempted to try something more audacious<sup>14</sup>:

I know that I am sometimes responsible. (C)

If I don't know that determinism is false, I don't know that I am ever morally responsible for my conduct. (B)

*Therefore:*

I know that determinism is false. (~A)

“The libertarians may know their Moore,”<sup>15</sup> we should reply, “but they can't know that determinism is false without the physics, which, we have just pointed out, does not disprove (A) yet. Without scientific research, they might have *some prima facie reason* to believe in indeterminism (or some practical reason); and perhaps it is even *metaphysically possible* for them to acquire what some have called ‘lucky’ or ‘unreasonable’ knowledge.<sup>16</sup> But they can't ‘go Moore’ on this one. They lack sufficient scientific evidence that determinism (of the right kind) is false.”<sup>17</sup>

We should note that our response does not turn on the claim that we cannot discover an *a posteriori* truth like determinism from *a priori* considerations. For (C) is a good candidate for an *a posteriori* truth: we discover that we are responsible through our moral experience, whether emotional or not. (In § 6, I come back to the parallel between the Moorean argument against skepticism and this Moorean argument.)

Is our response fair? Some might object that we can come to know that the universe (or its laws) is indeterministic without relying on any premises from physics. But we would discover this fact about the universe simply by identifying a counterexample to a generalization. When Europeans woke up after the 1755 Lisbon Earthquake, they perhaps discovered that the universe was *imperfect* in virtue of discovering something about one of its familiar corners. The illusion would be to think that one can't make a discovery about a whole because one has only made a discovery about a part, where the discovery about the whole is precisely that its parts do not *all* have a certain property.

Yet, the case of determinism seems different. To maintain that the laws of nature are indeterministic is not simply to reject a generalization. Instead, it tells against a *conception* of laws of nature, as simple, unified, and perhaps as necessarily binding or governing. (This touches on controversies about the precise way we should define "determinism.") This seems to be a discovery that we cannot make simply by reflecting on the fact that we are morally responsible. That I am responsible for having forgotten the milk on my way home just cannot be a Galilean crux experiment (an experiment which disproves a system) against deterministic interpretations of quantum mechanics. If it were, it seems that scientists should start compiling cases where people have formed knowledge of responsibility to mount a case against determinism in the same way that

Galileo mounted a case against Aristotelean physics by registering the spots on the Sun. Since this is absurd, we need to consider a different way out of our trilemma. (We will come back to rejecting (A) in § 5.)

The libertarian might reply that determinism is not a scientific hypothesis but a philosophical interpretation (and a generalization of) some scientific theories or discoveries.<sup>18</sup> The idea then is that the libertarian indeterministic hypothesis does not falsify any scientific discovery but circumscribes how these discoveries can be interpreted (and generalized).<sup>19</sup>

This is an important point, but one which I believe overplays the separation of science and philosophy. First, in general, it is very strange that reflection on our own moral agency could give us any strong clues as to whether determinism is true or not, whether it is characterized as a scientific theoretical claim or as a philosophical interpretation of a scientific thesis. Philosophy of science, just like science, should be independent of considerations about moral agency.<sup>20</sup>

Secondly, I am inclined to believe that determinism is *not* just a philosophical interpretation of scientific discoveries. This is because whether determinism is true or false seems to have some implications, although not straightforward ones, on what "first-order" scientific facts we can observe. For instance, whether determinism is true or false should in principle allow for *some prediction of human behavior that indeterminism (of the right kind) does not allow for*. Of course, indeterminism is compatible with *some* prediction of human behavior and determinism does not imply predictability of behavior. Yet, determinism combined with certain further plausible assumptions allows that we could predict the behavior of some agents in equipoise, that is who are just as tempted/likely to  $\Phi$  as to refrain from  $\Phi$ ing. But this is not allowed in principle by the kind of indeterminism cherished by libertarians.

### 3. SECOND WAY: IGNORANCE OF INCOMPATIBILISM

“Very well,” responds the non-skeptic, “give me a second shot”:

I don’t know that determinism is false. (A)

I know that I am sometimes responsible. (C)

*Therefore,*

If I know that I am sometimes responsible, it does not follow that I know determinism is false. (~B)

That is, (A) and (C) are compatible. I think that this path is the best one available to non-sceptics (those who want to preserve knowledge of responsibility). In this section, I will argue that, although it is open to libertarians, it implies that libertarianism is unknowable.

As I claimed in the introduction, this path is the home of compatibilists. Suppose that compatibilism is true. Then, since determinism and responsibility have nothing to do with each other, my knowledge that I am responsible does not stand in the way of my ignorance of determinism and vice versa. Is this path only available to compatibilists? Differently put, does the trilemma present an argument for compatibilism? Now, we touch on an important point.

Suppose compatibilism is false. Then it seems that I need to rule out the possibility of determinism to come to know that I am responsible. But since I can’t rule out the possibility of determinism, at least as things stand, I cannot come to know that I am responsible. But I know I am responsible. So, compatibilism is true. In other words:

The Blunt Argument

I know that I am sometimes responsible, and I cannot\* know that determinism is false.

If I know that I am sometimes responsible and I cannot\* know that determinism is false, then compatibilism is true.

*Therefore,* compatibilism is true.<sup>21</sup>

The asterisks remind us of the fact that we cannot know that determinism is false, *so long as science does not reveal that it is*. The Blunt Argument seems to approximate what Lewis had in mind when he said, “The best argument for compatibilism is that we know better that we are sometimes free than that we ever escape predetermination [that determinism is false]; wherefore it may be for all we know that we are free but predetermined.” (2000, p. 109)

The Blunt Argument fails because its second premise is false. From the fact that I know that I am responsible and that I cannot\* know that determinism is false, it does not follow that compatibilism is true. It only follows that *I cannot\* know that compatibilism is false*. Lewis’s argument supports the unknowability of incompatibilism, given that we know we are responsible, not the truth of compatibilism.

This can be illustrated by two independent cases where determinism is false, incompatibilism is true, and I am ignorant of these facts. In the first case, I have never heard of the problem of free will; I come to know that I am responsible for insulting my sibling, simply in virtue of noticing that I did this out of choice—and perhaps in virtue of feeling guilty. In the second (more controversial<sup>22</sup>) case, I have encountered the problem of free will; but the possibility of determinism is remote (perhaps deterministic worlds are metaphysically distant from this one<sup>23</sup>). It is so remote that it does not prevent me from knowing that I am responsible; it is no obstacle to satisfying the safety condition on knowledge.<sup>24</sup> In each case, I know that I am responsible, I don’t know that determinism is false and yet compatibilism is false. We have the confirmation that the Blunt Argument fails.

Yet a subtle counterpart of the Blunt Argument is successful. Here it is. I know that I am sometimes responsible, and I cannot know that determinism is false independently of

science. But I know that if I am sometimes responsible, either compatibilism is true or determinism is false. Suppose, for *reductio*, that I knew that compatibilism were false. Then I could come to know that determinism were false from these (non-scientific) premises. But I cannot. Therefore, I do not know that compatibilism is false. Let's restate this argument:

The Subtle Argument

I know that I am sometimes responsible, and I cannot\* know that determinism is false.

I know that <If I am sometimes responsible and compatibilism is false, then determinism is false.>

Therefore, I don't know that compatibilism is false.

This argument proceeds by two applications of the principle of *single premise closure under known entailment* (henceforth *closure*), that is, if I know that  $p$  and that  $p \rightarrow q$ , and from these premises I competently infer that  $q$ , then I know that  $q$ .<sup>25</sup> I shall come back to this assumption in § 4.

If we accept the Subtle Argument, we must accept that libertarianism is *unknowable\**, that is unknowable provided that we make two restrictions. The first restriction is important, but unsurprising given what we have said: we cannot know that libertarianism is true *so long as physics has not shown that determinism is false*. The second restriction is perhaps surprising, but unimportant: the conclusion might be restricted to those who have made the relevant inferences. I say that this is unimportant because it is no comfort to libertarians that they could only know their theory by refraining to read their rivals' papers.

(That libertarianism *might be knowable\** to those who fail to make the relevant inferences should be no consolation for the libertarian. To make this point clearer, consider the following:

I cannot\* know that determinism is false.

I *must* know that <If libertarianism is true, determinism is false.>

Therefore, I cannot\* know that libertarianism is true.

Suppose for *reductio* that I can know that libertarianism is true. Then, by this and the second premise, I can know that determinism is false. But I cannot know this, *qua* the first premise.)

It is good to pause for a moment to see what difference it makes to endorse the Subtle Argument as opposed to endorsing the Blunt Argument, apart from not accepting a false premise.<sup>26</sup> To endorse the Blunt Argument is to endorse compatibilism, whereas to endorse the Subtle Argument is to claim that libertarianism is unknowable\*, which is not a sufficient reason to accept compatibilism. The Subtle Argument should be appealing to compatibilists, since it offers them a free escape to the trilemma. (It is free for compatibilists because, since they believe libertarianism is false, they already take it to be unknowable.) But the argument is also available to an incompatibilist (libertarian or not). An incompatibilist might believe, *qua* the Subtle Argument, that incompatibilism is unknowable\*; and she might also accept that compatibilism *would be knowable\* if true*. But she takes it that incompatibilism is true, though unknowable\* (and so that compatibilism is also unknowable\* *since it's false*). Of course, this incompatibilist accepts a proposition that she takes to be unknowable\*, which is perhaps an undesirable position to be in.

Note that from the facts that incompatibilism is unknowable\* and that it rules out compatibilism, we cannot infer that compatibilism is true, even if we grant that compatibilism is knowable if true.

To see this, take a proposition like "Life exists or has existed outside the Solar System." If true, this proposition (or rather its truth) might be knowable: if we found unmistakable signs of life on an exoplanet,

we would come to know it. But if this proposition is false, then its negation is plausibly unknowable\*. It seems we cannot come to know that life has never existed outside the Solar System. The only way to know this would be presumably (at least, given what we know about life) to scan the whole universe with respect to all times and see that life is nowhere to be found but here. Now, performing such scans is physically impossible, because the speed of light puts a limit on which spatiotemporal stretches we can observe even in principle. Propositions which logically rule out “Life exists or has existed outside the Solar System” are equally unknowable\*. (For this reason, it is unadvisable to affirm that the question whether there is life outside the Solar System will soon be settled. It might be settled soon *if* there is life outside it.)<sup>27</sup>

What the Subtle Argument seems to reveal is that the proposition that compatibilism is true is like the proposition that life exists or has existed outside the Solar System. It is presumably knowable if true. Yet, if it is false, its negation (“incompatibilism is true”) remains unknowable\* (at least given what we know, e.g., that we are responsible). Propositions which logically entail that it is false (e.g., “libertarianism is true”) are also unknowable\*. In both the case of life outside the Solar System and of compatibilism, the epistemic asymmetry is related to fundamental physical facts.

Let’s recap. We considered in this section whether we could respond to the sceptic’s trilemma by accepting that we can know that we are responsible even if we do not know that determinism is false (~B). Having shown that the Blunt Argument is unsound, we can claim that this path—rejecting (B)—is not strictly reserved to compatibilists. However, having shown that the Subtle Argument is sound, we must claim that this path implies that incompatibilism is unknowable\*, given that we know we are responsible. In a word, it implies that libertarianism is unknowable\*.

But it does not imply that compatibilism is unknowable\*. *So, the trilemma leaves so far the compatibilist in a more comfortable position than the libertarian.*

#### 4. THIRD WAY: ANTI-CLOSURE AND CONTEXTUALISM

Some libertarians might share my solution to the trilemma and accept:

If I know that I am sometimes responsible, it does not follow that I know determinism is false. (~B)

And yet, they might want to chart a “third way,” away from the Subtle Argument. They might maintain that they can know that incompatibilism is true and that they are responsible but not that determinism is false. How? By rejecting closure or by endorsing epistemic contextualism.

Starting with closure, we may note that there are strong philosophical sentiments and arguments in its favor (see e.g., Hawthorne 2005). After all, we can extend our knowledge by drawing inferences and closure offers an excellent explanation of how this is possible. But more importantly for our purpose, there is a distinct problem for libertarians here. If they abandon closure, they need to replace it with something else. But the possible replacements—the two main anti-closure accounts of knowledge—are not libertarian-friendly.

Fred Dretske (1970) offered a first anti-closure account of knowledge, *error ruling-out*. It features a principle according to which I know a proposition only when I rule out the relevant possibilities of error. Consider the following classic scenario. I am in a zoo and observe a zebra in a pen. I know that this is a zebra. And I know that if this is a zebra, this is not a cleverly disguised mule. Yet, I can’t know that this is not a cleverly disguised mule. If, by contrast, I had been in a dodgy zoo in Cairo,<sup>28</sup> I would have perhaps not even been able to know that the animal in front of me was a zebra, even if it were. For, then, the

possibility that it be a cleverly disguised mule would be a relevant possibility of error.

Dretske's account, appealing as it might be, cannot help our third-way libertarian. Our third-way libertarian claims that she knows she is responsible, that incompatibilism is true and yet that she does not know that determinism is false. On Dretske's account of knowledge, this is possible only if determinism is not a relevant possibility of error. Determinism is perhaps not normally a relevant possibility of error: Dretske's account might allow that if I voluntarily break a promise, I come to know that I am responsible without ruling out the possibility that my action was determined. But when I form the belief that *libertarianism* is true, determinism *is* a relevant possibility of error. This is simply because determinism is part of the *definition* of libertarianism. By contrast, not being a cleverly disguised mule is not part of the *definition* of being a zebra (and not being determined is not a part of the *definition* of being responsible). Therefore, on Dretske's account, no one can come to know that libertarianism is true without having ruled out determinism; and since this is impossible independently of science, on Dretske's account libertarianism is unknowable independently of science. So Dretske's account leads the libertarian to the conclusion of the Subtle Argument. This is not a "third way".

Robert Nozick (1981) presented another anti-closure account of knowledge, *truth-tracking*. It features a principle, sensitivity, according to which I know a proposition only if I would not believe it if it were false. Yet again, this is of no help for the third-way libertarian. If libertarians lacked moral responsibility, they might have still believed that they were responsible. And if libertarianism had been false, they might have still believed it.<sup>29</sup> (Remember that sensitivity does not allow us to discard remote worlds. It is precisely why it implies that closure fails.) Thus, libertarians do not know that they are responsible, and

they do not know that libertarianism is true. It's best to keep closure.<sup>30</sup>

On top of what I have just said, there is an independent reason for the libertarians to keep closure. In fact, the Consequence Argument, probably the main incompatibilist argument ever championed, uses a principle which resembles closure:

β: If *S* has no control over *A* and *A* necessarily implies *B*, then *S* has no control over *B*.

β is intuitive, but those who find the Consequence Argument appealing typically invoke the fact that other closure principles are correct as evidence for β (see Fischer (1994)).<sup>31</sup> This makes sense if having control is understood as a kind of possibility; if, for instance, having control over a state of affairs just is for this state of affairs to be agentially possible.<sup>32</sup>

I now quickly turn to contextualism. *Epistemic contextualism* claims that the strength of epistemic position that is required of *S* for the utterance of "*S* knows that *P*" to be true will be higher in some contexts of utterance than others. So, libertarians could argue that they know they are responsible and that they know incompatibilism is true, but that they cannot know both things in the same context of utterance. Roughly, they know the first in theoretical context and the second in practical contexts. Because of this, they are never in a position to infer that determinism is false. Closure *within a same context* might remain true, but the context shifts as one goes through the Subtle Argument.

On a popular version of epistemic contextualism, a context of utterance is determined by conversational moves. Bringing up skeptical hypotheses raises the standards. The idea here, I take it, is that bringing up determinism would raise the standards of knowledge of responsibility. But what this means is that libertarianism is unknowable, for the same reason that it is on Dretske's account. For determinism is part of the definition of libertarianism;

the latter cannot be brought up without the former. Moreover, this conversational contextualism has the strange consequence that I could avoid being the appropriate target of blame if I raise the possibility of determinism at the right moment. This is implausible.

On another kind of epistemic contextualism, a context of utterance is determined by objective or conventional features of the situation, for instance whether the discussion occurs in a philosophy classroom or in a court of law. The idea is that in the latter but not in the former, one can know that they are responsible. But this version of contextualism does not offer a ‘third-way’ solution: it does not discard the possibility that, in a court room or in other practical contexts, I could come to know that determinism is false (without the help of science). It merely offers a more complex version of the first way out of the trilemma, which we rejected in § 2.

In the last section (§ 3), we argued that the best way to escape our skeptical trilemma is to endorse:

If I know that I am sometimes responsible, it does not follow that I know determinism is false. (~B)

From this, we cannot conclude that compatibilism is true, as the Blunt Argument would have it. Instead, we can conclude that libertarianism is unknowable, independently of science. And we confirmed this conclusion in this section (§ 4), by demonstrating that, independently of their respective merit, contextualism and anti-closure accounts of knowledge do not offer to the libertarian a third way out of the trilemma.

### 5. THREE DIALECTICAL POINTS

We come back to the Subtle Argument and its consequences for the libertarians. Its weakest point is the claim that I cannot know that determinism is false independently of science. This takes us back to the strategy that we rejected in § 2: accepting that we can

come to know that determinism is false on the grounds that we know that we are responsible.

To those who still accept this strategy, I can offer three dialectical points. The first point is that we seem to be in a context where compatibilists and incompatibilists are in a deadlock.<sup>33</sup> On both sides, there are convincing arguments. It is true that incompatibilist arguments are persuasive, but not comparatively enough to warrant that we disregard the counter-intuitive consequence that we could come to know that determinism is false without the help of physics. Differently put, the libertarians would need to be in a much stronger position to accept the audacious claim that they know that determinism is false independently of science.

The second dialectical point is that it would be disingenuous on the libertarian’s part to champion the conclusion that we know that determinism is false as a surprising though necessary anti-skeptical response. For it is unnecessary: we can reject (B) to escape the trilemma; and we can do so without endorsing compatibilism. All we need is to accept that we cannot know\* libertarianism. It does not mean that libertarians should stop fine-tuning their arguments—they might come handy if one day we discover that determinism is false. Or, they might make it harder for others to embrace compatibilism.

The third dialectical point is that we can explain how it is that convincing libertarian arguments (typically combination of incompatibilist arguments and of one anti-skeptical premise) cannot give us knowledge of indeterminism. We cannot come to know that determinism is false on the basis that (1) we are sometimes responsible and (2) incompatibilism is true. For, absent our knowledge that determinism is false, (1) and (2) are evidence against one another. They are evidence against one another because if incompatibilism is true, it is harder to be responsible (some of our actions need to be undetermined); and if we are responsible, it is

less likely that incompatibilism is true (since it is not any less likely that determinism is true). As such, incompatibilist arguments are, for libertarians, a double-edged sword: the better their arguments, the more fragile their knowledge that they are responsible.

#### 6. AN OBJECTION FROM EXTERNAL WORLD SKEPTICISM

Let us now go back where we started this conversation: with the parallel between external world and free will skepticism. Remember: in both cases, I don't know that I am not in the bad case (being systematically deceived by a Cartesian Demon, living in a deterministic world); if I don't know this, I don't know something I took for granted (that I have hands, that I am ever morally responsible for my conduct); therefore, I don't know what I took for granted.

This exact parallel was used by Peter van Inwagen to argue for the view that we can in fact come to know that determinism is false without the help of science. His argument is that, just like we should accept, on empirical grounds (e.g., by raising our hands *à la Moore*), that we can come to know that we are not being universally deceived by a Cartesian demon, we should also accept, on moral grounds, that we can come to know that we are not determined:

You say that our having good reasons for believing in moral responsibility is not the sort of thing that could provide us with good reasons for believing that determinism is false. But you *also* say that our having good reasons for believing that our knowledge-claims are generally correct *is* the sort of thing that can provide us with good reasons for believing that there is no Universal Deceiver. But isn't it true that you are responding to philosophical problem that are essentially the same in arbitrarily different ways? (van Inwagen 1983, p. 213)

This objection is ingenious, for it is hard to resist the analogy (for a very similar objection, see Steward (2020)). Coffman (2016, pp.

89–90) offers three possible responses; let me mention the two more simple ones. The first response is that we know directly that we are not deceived by a Cartesian Demon, but not that determinism is false. This means that we do not come to know that we are not deceived on the basis of a problematic Moorean inference (from empirical grounds). Therefore, we can cast doubt on the validity of the Moorean inference from responsibility to indeterminism without having to worry about external world skepticism. The second response is that we have abductive evidence for the claim that we are not being deceived (we can explain our world better with the non-skeptical hypothesis), but we have no such abductive evidence for indeterminism. There is no data that is left unexplained without the hypothesis of indeterminism. So, again, we do not come to know that we are not deceived from a problematic Moorean inference and can therefore cast doubt on the validity of the Moorean inference from responsibility to indeterminism.

These suggestions are epistemologically controversial. (If they are right, though, my conclusion stands.) I prefer a different solution. I want to insist that there is a different key asymmetry between the two philosophical problems under consideration.

Take the epistemic skepticism case. I know that I have hands; but I know that if I have hands, then I am not being universally deceived; therefore, I know that I am not being universally deceived by a Cartesian demon lurking at the end of the universe. There is no doubt that this conclusion comes as a surprise. How could I know this? We should not discount this surprise because we are used to taking the hardline against skepticism. If we take our surprise seriously—the surprise of coming to know something about a Cartesian deceiver at the end of the universe from our sofa—we should say this: the fact that we infer something so abstrusely is (defeasible) *evidence* that we got something wrong in our deduction.

Compare: suppose that during dinner I get in an argument about whether Algerian cuisine offers the best African food. Imagine that I somehow end up inferring that Algeria is the largest country in Africa, *from purely culinary premises*. I should take this as *evidence* that I made a bad inference along the way or that I said something false. Perhaps I slipped from “dessert” to “desert” and inferred that since Algeria has the most impressive des(s)erts in Africa, it is the largest African country. We might want to admit a rule of thumb like the following: if we infer a proposition of kind *X* from a proposition of kind *Y*, then the more (significantly) different *X* and *Y*, the stronger our suspicion should be about the inference.

Suppose I face the skeptical hypothesis. I consider my options. First, I can abandon the claim that I know that I have hands—but this is a no go.<sup>34</sup> Second, I can abandon the claim that *if I know that I have hands, I know that I am not deceived*. But the only way to do this is to abandon closure or endorse contextualism, two options that careful analysis reveal to be implausible. I am left with one option: to accept that I know that I am not being deceived. And, this, it seems, I can only know (this is what Coffman would deny) in virtue of a Moorean inference. But it does not mean that I need to accept the Moorean inference “as if it was nothing,” that is as if there was nothing epistemically strange in coming to know things about remote deceivers from my knowledge that I have hands. It *is* strange to come to know something about a universal deceiver in this way. But this remains my best option. It’s a better option than skepticism, the rejection of closure, or the endorsement of contextualism. Again, the strength of the Moorean inference gave me pause before accepting closure. But now I see that I must admit the inference.

By contrast, suppose I face the skeptical hypothesis about free will. I consider my options. First, I can abandon the claim that I know that I am responsible—but this is a no

go (according to our working assumption). Second, I can abandon the claim that *if I know that I am responsible, then I know that determinism is false*. Here lies the difference between the two skeptical challenges: in the case of free will but not of knowledge, I can plausibly abandon this claim without denying closure or endorsing contextualism.<sup>35</sup> As we have argued, I can simply accept that incompatibilism is unknowable\*. Now, *this* is a good option. It is a better option than to accept that I can come to know that determinism is false without the help of scientific inquiry.

I think that what makes van Inwagen’s argument so appealing is the proneness that we have to reject epistemic skepticism. We see it and our philosophical training convinces us that we must be able to refute it. Therefore, we conclude that Moorean inferences are innocent. We remain blind to the fact that ordinarily (but not always) when we make one such surprising inference, we did something wrong.

## 7. CONCLUSION

I have presented an unusual skeptical challenge, according to which since we do not know that determinism is false, we do not know that we are responsible. Like many skeptical arguments, this one reveals the epistemological commitments of the non-skeptics (and non-agnostics). It shows that we can know that we are responsible despite not knowing that determinism is false. It would be too hasty, as I have argued *pace* the Blunt Argument, to conclude from this that compatibilism is true. Rather, we must accept that, given the present state of science, libertarianism is unknowable. It is still an open question whether, now that we are aware of this, we can maintain a reasonable belief in libertarianism (perhaps for pragmatic reasons, as James (1912) suggested). It is also an open question whether we can reasonably believe in compatibilism or in free will skepticism. Some philosophers (e.g., Epicurus, *The*

*Extant Remains*; Kant, *Groundworks* 4:448; Lockie 2018) have argued that these views are self-defeating to believe in. I do not think that these arguments are sound, but they should be taken seriously (see Chevarie-Cossette 2019a; 2019b).<sup>36</sup>

Might the incompatibilist retaliate by showing that a mirror image of the Subtle Argument<sup>37</sup> afflicts compatibilism? If this were the case, this would show that non-skeptical compatibilists and libertarians are in the same boat. I think however that most compatibilists would agree that there is no such argument, on the grounds that the existence of moral responsibility does not imply the *truth* of determinism. Admittedly, the view that it

does has had its defenders—possibly Hume (1739) and Hobart (1934). And compatibilists tend to overplay how destructive indeterminism would be for freedom when discussing libertarianism. Until physics tells us that determinism is *true*, this form of compatibilism is also unknowable—just like libertarianism. For, otherwise, we could come to know that *determinism* is true on non-scientific grounds. In any case, we should be prudent before requiring the cooperation of the microphysical world for us to be responsible.<sup>38</sup>

*Université de Neuchâtel and  
King's College London*

## NOTES

I wish to thank Ralf Bader, Charles Côté-Bouchard, Taylor Cyr, Marc-Kevin Daoust, Samuel Dishaw, Tanya Goodchild, Alison Hills, John Hyman, Clayton Littlejohn, and two anonymous reviewers of the *American Philosophical Quarterly* for their invaluable comments on previous drafts of this paper.

1. “Free will” and “moral responsibility” are different concepts, although it is common nowadays to define “free will” as the power necessary for morally responsible conduct (see Pereboom 2014; cf. Fischer 1994). I prefer to talk in terms of moral responsibility to include in the discussion philosophers who find free will obscure.
2. But see Byrd (2010, pp. 412–413) and Kearns (2015) who argue explicitly in favor of this view.
3. This is not a mere wordplay. Free will sceptics share a certain methodology with other sceptics, that is they typically do not revise their premises when these lead to revisionist conclusions.
4. See Sehon (2013, p. 371) for the same argument about free will.
5. Again, these views could be defined in terms of free will rather than moral responsibility. For the first, see for example Chisholm (1964) and van Inwagen (1983); for the second, Strawson (1974) and Kenny (1976); for the third, Waller (2011) and Pereboom (2014).
6. See for instance, Smilansky (2000); Nichols (2007).
7. See Strawson (1974).
8. See for instance, Pereboom (2014); Waller (2011); Nadelhoffer and Tocchetto (2013).
9. See Wolf (1981).
10. Some readers might deny the importance of knowledge and favor *justification*. I think it is plausible to replace ‘know’ with ‘believe with justification’ in most of my arguments.
11. Self-forming willings are at the center of Kane’s (1996) libertarian conception of free will and responsibility.

12. See Chisholm (1964); Pereboom (2014a).
13. Kearns (2015, p. 242) declares that he sees no reason for thinking that if I know I have free will (responsibility), discovering that incompatibilism is true could remove this piece of knowledge. I disagree: the truth of incompatibilism could be a defeater.
14. See Steward (2020), who takes this “Moorean anti-skeptical strategy.”
15. I am referring here to G.E. Moore’s (1939) infamous response to skepticism.
16. See Lasonen-Aarnio (2010).
17. Maier (2014) and Kearns (2015) have argued for this view. Coffman (2016) has replied, it seems, by exploiting the fact that sometimes we come to know things because we have not paid attention to defeaters. I don’t think that this possibility is a cause for comfort for the libertarian.
18. This is suggested by Steward (2020).
19. I am grateful to an anonymous reviewer of the *American Philosophical Quarterly* for raising this issue.
20. Even if determinism were not a thesis *in the philosophy of science*, it would still have important implications for it; but it would be troubling if moral considerations had important implications for philosophy of science.
21. Kearns (2015, p. 242) suggests a similar argument. He claims that if incompatibilism is true, no one knows there is free will. He claims later that if he is wrong to claim that we cannot know that free will exists, then the compatibilist can use the Blunt Argument, in terms of free will (2015, p. 246).
22. The first example, but not the second, is compatible with an internalist framework in epistemology.
23. This seems especially plausible if determinism is incompatible with agency. See Steward (2012).
24. *Safety* claims that *S* knows *p* only if there is no close possible world where *S* believes *p* falsely.
25. Proof. From the two premises, I first infer that <if compatibilism is false, determinism is false>, which I come to know because of closure. But then, since I do not know that determinism is false—and could not come to know it in the process—I can infer that I do not know that compatibilism is false.
26. I am grateful to an anonymous reviewer of the *American Philosophical Quarterly* for raising this issue.
27. I am very grateful to Kolja Keller for this parallel.
28. For real cases, see <https://www.bbc.co.uk/news/world-africa-44968509>.
29. These counterfactuals are perhaps harder to assess than I make it; but this difficulty is an obstacle to a third-way libertarianism.
30. See Sehon (2013, pp. 375–379), who makes a similar argument about several theories of moral responsibility.
31. In a similar spirit, Dretske (2005) argues against closure under known entailment by arguing that other closure principles (memory closure, proof closure, perception closure, etc.) are false.
32. See List (2014) and Kenny (1976) for a discussion.
33. It is this deadlock that leads several philosophers to agnosticism (see Double 1996; Byrd 2010; Kearns 2015).
34. It is a no go to doubt that I have hands for multiple reasons: because it is obvious that I have hands, because trying to doubt this proposition would not be a real doubt since it is a hinge proposition, because it self-defeating to doubt such a proposition.

35. Ralf Bader asks ingeniously whether, in the case of knowledge, idealism might play the role that compatibilism plays here. Just like we cannot know that compatibilism is true, we could not come to know that idealism is false. This is supposed to show that the second premise of the skeptical argument is false. I am unsure that this works as well for knowledge as it does for free will, and not just because idealism is less plausible than compatibilism. The idealist response might work if the argument is about the *existence of the external world*, i.e., if the rejected premise is “if I know that I have hands, I know that there is an external world.” But the premise can be “if I know that I have hands, I know that I am not *being universally deceived*.” Does the idealist response work here? If the answer is yes, it seriously speaks against idealism. If the answer is no, it does not solve our problem.

36. I am grateful to anonymous reviewers of the *American Philosophical Quarterly* for raising these issues.

37. Kearns (2015, p. 241) explores a similar parallel, when he argues that if compatibilism is true, we cannot know that we have free will. I think the argument fails because we could discover that compatibilism is true by discovering that determinism is, assuming that we know we are responsible. Also, if there is a gap between philosophical or ethical considerations and scientific considerations, this is *strong evidence* for compatibilism.

38. Of course, if we can know that we are responsible, there are ‘remote’ conclusions that we will have to accept, for instance that we can know that there is no manipulator undermining our freedom or that our choices are not epiphenomenal.

## REFERENCES

- Byrd, Jeremy. 2010. “Agnosticism about Moral Responsibility,” *Canadian Journal of Philosophy*, vol. 40, no. 3, pp. 411–432.
- Chevarie-Cossette, Simon-Pierre. 2019a. “Self-Defeating Beliefs and Misleading Reasons,” *International Journal of Philosophical Studies*, vol. 27, no. 1, pp. 57–72.
- . 2019b. “Is Free Will Scepticism Self-Defeating?” *European Journal of Analytic Philosophy*, vol. 15, no. 2, pp. 55–78.
- Chisholm, Roderick. 1964. *Human Freedom and the Self* (Kansas City, KS: *The Department of Philosophy, University of Kansas*).
- Coffman, E. J. 2016. “Incompatibilist Commitment and Moral Self-Knowledge: The Epistemology of Libertarianism,” *Philosophical Issues*, vol. 26, no.1, pp. 78–98.
- Double, Richard. 1996. *Metaphilosophy and Free Will* (Oxford, UK: Oxford University Press).
- Dretske, Fred I. 1970. “Epistemic Operators,” *Journal of Philosophy*, vol. 67, no. 24, pp. 1007–1023.
- . 2005. “The Case Against Closure,” In *Contemporary Debates in Epistemology*, ed. M. Steup and Ernest Sosa (Malden, MA: Blackwell), pp. 13–25.
- “Egypt zoo accused of painting donkey to look like a zebra,” 2018. *BBC*, July 26. <https://www.bbc.co.uk/news/world-africa-44968509>.
- Fischer, John Martin. 1994. *The Metaphysics of Free Will: An Essay on Control* (Oxford, UK: Blackwell).
- Hawthorne, John. 2004. *Knowledge and Lotteries* (Oxford, UK: Clarendon Press).
- Hawthorne, John. 2005. The case for closure. In *Contemporary debates in epistemology*, ed. M. Steup and Ernest Sosa (Malden, Ma: Blackwell), pp. 26–43.
- Hobart, R. E. 1934. “Free Will as Involving Determination and Inconceivable Without It,” *Mind*, vol. 43, no. 169, pp. 1–27.
- Hume, David. 1739. *A Treatise of Human Nature* (London, UK: John Noon).
- Hyman, John. 2015. *Action, Knowledge, and Will* (Oxford, UK: Oxford University Press).
- Inwagen, Peter van. 1983. *An Essay on Free Will* (Oxford, UK: Oxford University Press).

- James, William. 1912. *The Will to Believe: And Other Essays in Popular Philosophy* (Auckland, New Zealand: Floating Press).
- Kane, Robert. 1996. *The Significance of Free Will* (Oxford, UK: Oxford University Press).
- Kearns, Stephen. 2015. "Free Will Agnosticism," *Noûs*, vol. 49, no. 2, pp. 235–252.
- Kenny, Anthony J. P. 1976. *Will, Freedom, and Power* (Oxford, UK: Blackwell).
- Lasonen-Aarnio, Maria. 2010. "Unreasonable Knowledge," *Philosophical Perspectives*, vol. 24, no. 1, pp. 1–21.
- Lewis, David. 2000. *Papers in Ethics and Social Philosophy* (Cambridge, MA: Cambridge University Press).
- Lockie, Robert. 2018. *Free Will and Epistemology: A Defence of the Transcendental Argument for Freedom* (New York, NY: Bloomsbury Academic).
- Maier, John. 2014. "The Argument from Moral Responsibility," *Australasian Journal of Philosophy*, vol. 92, no. 2, pp. 249–267.
- Moore, G.E. 1939. "Proof of an External World," *Proceedings of the British Academy*, vol. 25, no. 5, pp. 273–300.
- Nadelhoffer, Thomas, and Daniela Goya Tocchetto. 2013. "The Potential Dark Side of Believing in Free Will (and Related Concepts)," in *Exploring the Illusion of Free Will and Moral Responsibility*, ed. Gregg D. Caruso (Lanham, MD: Lexington Books).
- Nichols, Shaun. 2007. "After Incompatibilism: A Naturalistic Defense of the Reactive Attitudes," *Philosophical Perspectives*, vol. 21, no. 1, pp. 405–428.
- Nozick, Robert.  
 ———. 1981. *Philosophical Explanations* (Oxford, UK: Clarendon Press).
- Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life* (Oxford, UK: Oxford University Press).
- Sehon, Scott. 2013. "Epistemic Issues in the Free Will Debate: Can We Know When We Are Free?" *Philosophical Studies*, vol. 166, no. 2, pp. 363–380.
- Smilansky, Saul. 2000. *Free Will and Illusion* (Oxford, UK: Oxford University Press).
- Steward, Helen. 2020. "Free Will and External Reality: Two Scepticisms Compared," *Proceedings of the Aristotelian society*.
- . 2012. *A Metaphysics for Freedom*. (Oxford, UK: Oxford University Press.)
- Strawson, P.F. 1974. *Freedom and Resentment and Other Essays* (Abingdon, UK: Routledge).
- Vilhauer, Benjamin. 2009. "Free Will and Reasonable Doubt," *American Philosophical Quarterly*, vol. 46, no. 2, pp. 131–140.
- Waller, Bruce N. 2011. *Against Moral Responsibility* (Cambridge MA: The MIT Press).
- Wolf, Susan. 1981. "The Importance of Free Will," *Mind*, vol. 90, no. 359, pp. 386–405.

2