



# Is Fallible Knowledge Attributable?

E. J. Coffman<sup>1</sup>

Received: 16 September 2020 / Accepted: 5 August 2021 / Published online: 18 August 2021  
© Springer Nature B.V. 2021

## Abstract

Here are two *prima facie* plausible theses about propositional knowledge: (i) a belief could still constitute knowledge even if the belief is justified in a way that's compatible with its being either false or accidentally true; (ii) each instance of knowledge is related to its subject in a way similar to that in which each intentional action is related to its agent. Baron Reed (2007, 2009) develops and defends a novel argument for the incompatibility of (i) and (ii). In this paper, I clarify and critically assess Reed's incompatibility argument. Against the backdrop of an example in which an action is non-intentional due to the role that an accidentally true belief plays in the action's etiology, I argue that Reed's incompatibility argument defeats itself: two of its premises are themselves jointly incompatible.

Here are two *prima facie* plausible theses about propositional knowledge:

**Fallibilism:** Possibly, a belief constitutes knowledge even though the belief is justified in a way that's compatible with its being either false or accidentally true.

**Attributabilism:** Each instance of knowledge is related to its subject in a way similar to that in which each intentional action is related to its agent.

Baron Reed (2007, 2009) develops and defends a novel argument for the incompatibility of fallibilism and attributabilism.<sup>1</sup> In this paper, I clarify and critically assess Reed's incompatibility argument. Against the backdrop of an example in which an action is non-intentional due to the role that an accidentally true belief plays in the action's etiology, I argue that Reed's incompatibility argument defeats itself: two of its premises are themselves jointly incompatible.

<sup>1</sup> On the basis of (i) his incompatibility argument and (ii) his case for the claim that the conjunction of fallibilism and attributabilism is "of central importance to the epistemological theories of our day" (2007: 245), Reed concludes that "knowledge, as we now conceive of it, is impossible" (2007: 261; cf. 2009: 103).

---

✉ E. J. Coffman  
ecoffma1@utk.edu

<sup>1</sup> Department of Philosophy, University of Tennessee, 801 McClung Tower, Knoxville, TN 37996, USA

## 1 |

First, I need to justify my clarificatory statement of the thesis that Reed labels “attributabilism.” I’ve explicitly stated attributabilism in terms of intentional action. While Reed does not so state attributabilism, such a statement constitutes the most charitable interpretation of Reed’s (somewhat less clear) formulations of the thesis.<sup>2</sup>

Here is Reed’s initial statement of Attributabilism:

For any state of knowledge *k* that a person *S* possesses, the possession of *k* is attributable to *S*. (2007: 238)

As Reed immediately acknowledges, if the material after the comma is equivalent to (something like) “knowledge of *k*’s propositional content can be accurately ascribed to *S*,” then attributabilism is a truism. But if attributabilism is a truism, then Reed’s incompatibility argument is tantamount to an attempt to establish that fallibilism is self-contradictory—an extremely implausible thesis which Reed clearly does not endorse (cf. 2002; 2007: 236–7; 2009: 94, 102).

In order to forestall such a misinterpretation of attributabilism specifically and his incompatibility argument more generally, Reed provides the following commentary on his initial statement of attributabilism:

But what is meant by “attributable” is something stronger [than ‘can be accurately ascribed to’]. A comparison with the attributability of actions is instructive here.

There are different ways of explaining the special relationship an agent bears to her action. According to simple versions of compatibilism, ...the action must be caused by the agent’s will. For more sophisticated versions of compatibilism, the action must come from... the “Real Self,” where some philosophers take this to be an appropriately structured will while others argue it is the agent’s faculty of reason. Among libertarians, ...we find similar variety: some of them require only that the action be indeterministically produced by the agent (perhaps on the basis of reasons the agent has), while others hold the agent to be a special sort of cause. But what is common to all of these philosophers, compatibilists and libertarians alike, is the attempt to show how the action came from the *person* in question. It is not enough for the person to be merely the locus of a causal process terminating in something like an action. A reflexive leg-kicking, for example, is not something that one *does*—it fails to belong, in the relevant sense, to the person who kicks.

Many epistemologists have made a similar point about knowledge and its relation to the person who has it. (2007: 238-9)

According to the attributabilist, then, propositional knowledge is similar to a certain practical relation—that is, a certain relation that people bear to (at least some of) their actions. In the above passage, Reed expresses the pertinent practical relation

<sup>2</sup> For insightful recent discussion and endorsement of the view that propositional knowledge and intentional action are similar in terms of personal attributability, see Sosa 2015 and Williamson 2017.

with “came from,” an emphasized “does,” and “belong to.” Elsewhere, in one of his clearest and most concise statements of attributabilism, Reed connects “the relevant sense” of “belong to” with practical moral responsibility (i.e., moral responsibility for action)<sup>3</sup>: “[Knowledge] must belong to the person as her own doing, much as an action for which a person is responsible must belong to her” (2009: 103).

Note that, contrary to what the third sentence of the longer quotation in the last paragraph may suggest (“the special relationship an agent bears to her action”), the practical relation to which the attributabilist assimilates propositional knowledge is not that of mere (action) performance. If the pertinent practical relation is mere performance, then both (a) the emphasized “does” (near the end of the longer quotation in the last paragraph) and (b) the phrase “for which a person is responsible” (in the shorter quotation at the end of the last paragraph) are not only unnecessary but also potentially misleading.<sup>4</sup> Charity thus demands that we interpret the attributabilist as appealing to a practical relation that’s logically stronger than that of mere performance.

The clarificatory material provided in the last couple paragraphs yields the following revised statement of attributabilism:

Each instance of knowledge is related to its subject in a way similar to that in which each action that belongs to its agent in a stronger-than-mere performance sense required for moral responsibility is related to its agent.

Now there’s a problem with this revised statement of attributabilism: Unless we read “moral responsibility” as “nonderivative moral responsibility” (where one is “non-derivatively morally responsible” for an action iff one’s moral responsibility for that action does not derive from one’s moral responsibility for some other action), the revised statement of attributabilism simply fails to pick out a practical relation and thus fails to assimilate knowledge to such a relation. While *nonderivative* moral responsibility does require a practical relation that’s logically stronger than mere performance—plausible candidates include intentional performance and metaphysically free performance—*derivative* moral responsibility does not require a practical relation that’s logically stronger than mere performance. To see this, consider a drunk driver who inadvertently strikes a pedestrian. We can imagine this case such that the driver is derivatively morally responsible for striking the pedestrian even though the driver’s striking the pedestrian was merely something that the driver did: the driver didn’t strike the pedestrian intentionally, or for a reason, or freely, or knowingly, etc.

Charitably interpreted, then, the attributabilist assimilates propositional knowledge to a certain practical relation that is (a) required for *nonderivative* moral responsibility and (b) logically stronger than mere performance. One such practical relation, of course, is

<sup>3</sup> Unless otherwise noted, in what follows “moral responsibility” abbreviates “practical moral responsibility.”

<sup>4</sup> On (a), just as an emphasized “knows” often expresses an epistemic relation that’s logically stronger than mere knowledge, an emphasized “does” often expresses a practical relation stronger than mere performance.

nonderivative moral responsibility. But interpreting attributabilism in terms of nonderivative moral responsibility conflicts with Reed's claim that the practical relation to which the attributabilist likens propositional knowledge does not entail moral responsibility. Writes Reed: "Just as an action... may be attributable to a person who does not deserve credit for it..., so an instance of knowledge... may be attributable to a person who does not deserve credit for it" (2007: 242; cf. 2009: 103).

I see four additional candidates for the practical relation to which the attributabilist likens propositional knowledge: metaphysically free performance; self-controlled performance; intentional performance; and apt performance—that is, performance whose success manifests some or other of the agent's competences (cf. Sosa, 2015).<sup>5</sup> Interpreting attributabilism in terms of metaphysically free performance conflicts with Reed's claim that the relevant practical relation does not entail (what John Martin Fischer [2008: 57] calls) regulative control—that is, control *over* (rather than mere control *of*) an event or state. (It's extremely plausible to think that an action one performed was metaphysically free only if one was free with respect to—and therefore had control over—the action's occurrence.) Writes Reed (2007: 242): "...knowledge is attributable to us, not because we have ultimate control over it, but because it comes out of a cognitive 'Real Self'." Now the sentence just quoted may suggest that attributabilism should be understood in terms of self-controlled performance. But such an interpretation of attributabilism conflicts with the obvious fact that an action for which one is nonderivatively morally responsible needn't also be an instance of self-control—think, for example, of certain akratic or weak-willed actions (cf. Mele, 2012).

Unlike the previously rejected interpretations, statements of attributabilism that focus on either intentional action or apt action fit well with relevant material in Reed's papers. For starters, an intentional action can be accurately described as "coming from an agent" and "belonging to an agent as their own doing" (Reed's expressions), as can an apt action. Moreover, either of "does intentionally" and "does aptly" would be a charitable interpretation of the emphasized "does" in the following sentence (quoted earlier): "A reflexive leg-kicking... is not something that one *does*—it fails to belong, in the relevant sense, to the person who kicks" (Reed, 2007: 238–9). Finally, each of the following two theses is extremely plausible: (i) one is nonderivatively morally responsible for an action one performed only if one performed the action intentionally; (ii) any intentional action is an apt action.<sup>6</sup> Accordingly, both intentional performance and apt performance are practical relations that are stronger than mere performance and plausibly regarded as required for nonderivative (practical) moral responsibility.

We have before us two tenable interpretations of attributabilism: one that invokes intentional action, and one that invokes apt action. If one of the indicated practical relations is logically stronger than the other, then we should focus first on the interpretation of attributabilism that invokes the stronger relation, since Reed's

<sup>5</sup> Note that mere performance doesn't suffice for apt performance. My rolling "boxcars" (two sixes) with a pair of fair dice is a non-apt action, as is the focal action in a "secondary causal deviance" case (cf. Mele & Moser, 1994: 47ff.).

<sup>6</sup> Here's a promising argument for thesis (ii): any intentional action is done for a reason (cf. Mele & Moser, 1994); any action done for a reason is an apt action (cf. Neta, 2019; Sosa, 2015); hence, any intentional action is an apt action.

incompatibility argument would be more likely to succeed on the stronger interpretation of attributabilism (as we strengthen the interpretation of attributabilism, we weaken Reed's incompatibility thesis, thereby making the latter easier to establish). Is one of the indicated practical relations stronger than the other?

Arguably, yes: while all intentional actions are apt actions, some apt actions are non-intentional actions. Consider the following case, which is inspired by Alfred Mele and Paul Moser (1994: 41) as well as Duncan Pritchard (2012: 260)<sup>7</sup>:

An informant whom I justifiably believe to be extremely reliable tells me that the winning numbers in tomorrow's pick-your-own-number lottery will be 8675309. On this occasion, my informant is in fact just guessing about tomorrow's winning numbers. But my informant's guess turns out to be accurate, for—unbeknownst to both me and my informant—there exists a sufficiently powerful and knowledgeable third party whose job is to ensure that tomorrow's winning numbers turn out to be whatever I think they will be. Through my informant's testimony, I acquire a justified but accidentally true belief that the winning numbers will be 8675309. I acquire an intention to win the lottery by picking 8675309. I intentionally pick those numbers. I win the lottery.

This example can be understood such that I non-intentionally won the lottery—that is, I won the lottery but did not do so intentionally (cf. Gibbons, 2010; Mele, 1987; Mele & Moser, 1994).<sup>8</sup> In support of the ascription of aptness to my lottery win—and to echo what Mele (1992: 208) says about a relevantly similar case—note that in winning the lottery, I followed the plan component of my intention to do so and that there was no deviant connection between my intention and my picking the pertinent numbers or between my picking those numbers and my winning the lottery. Accordingly, I seem to have carried out or executed my intention to win the lottery, which in turn seems to suffice for my lottery win's manifesting some or other pertinent competence, though not for my lottery win's being something that I did intentionally (cf. Mele, 1992: 207–8).<sup>9</sup>

<sup>7</sup> This case blends aspects of Mele and Moser's second "Fred" case (which features a misread handwritten code) with aspects of Pritchard's "Temp" case.

<sup>8</sup> Note that one's performing an action, A, *non-intentionally* doesn't suffice for one's performing A *unintentionally* (cf. Mele & Moser, 1994: 45).

<sup>9</sup> Sosa (2015: 158) sketches and discusses a somewhat similar case:

A prisoner is told by his jailer that throughout the coming night his jail cell will be unlocked, but the jailer's testimony is only a cruel joke. The prisoner does form the belief, though, and by acting on it that night, he escapes, since completely by accident the door *was* unlocked. Is his escape intentional? Is it apt? It is hard to deny that the prisoner escapes intentionally (by design). Moreover, his escape does seem plausibly enough to manifest a degree of competence already seated in him as he lay in bed prepared to make his move.

I find the ascription of intentionality to the prisoner's escape counterintuitive, and I suspect that many other action theorists would as well (cf. Gibbons, 2010; Mele, 1987, 1992; Mele & Moser, 1994). I hasten to add that, as he (2015: 10ff.) continually notes and as the parenthetical material in the above passage indicates, Sosa's use of "intentionally" is somewhat stipulative, "restricted to that of 'by design'" (10n4). Depending on what exactly Sosa means by "by design" and related expressions ("on purpose," "according to plan," etc.), it may turn out that he and I don't genuinely disagree about the intentionality status of his example's focal action. In any event, I'm confident that Sosa's (currently relatively under-described) case could be amplified so as to illustrate the possibility of apt yet non-intentional action. For another case that illustrates the indicated possibility (though perhaps not as clearly as the one that I focus on in the main text), see Mele and Moser's (1994: 51) "Arnold" case.

We'll return to this lottery case below (section 3). For now, we just need to take from it the lesson that some apt actions are non-intentional. Assuming that's right, it follows that intentional action is logically stronger than apt action and that we should focus first on the interpretation of attributabilism that invokes intentional action, since Reed's incompatibility argument is more likely to succeed on that (relatively strong) interpretation of attributabilism. If Reed's argument were to succeed on the "intentional action" interpretation of attributabilism, we could then go on to consider whether his argument also succeeds on the "apt action" interpretation of attributabilism.<sup>10</sup>

## 2 II

Having justified my clarificatory statement of attributabilism, I can now start moving toward a full formal statement of Reed's argument for the incompatibility of fallibilism and attributabilism. The following passage—in which Reed is discussing an arbitrarily selected instance of fallibly justified knowledge—provides a rough initial statement of his argument:

The result, then, is that the person's cognitive performance is not good enough to allow her to reach the truth. Her performance must be supplemented by something extrinsic—and therefore unattributable—to her. Thus, [fallibly justified] knowledge itself is not attributable to the person, either. Attributabilism is therefore false, as long as we are committed to fallibilism. (2007: 248)

To see a problem with this rough initial statement of Reed's argument, notice that the second claim in the passage entails the third only on the following "transfer of non-attributability" assumption:

If S's knowing that P depends on the satisfaction of a condition whose obtaining is not attributable to S, then S's knowing that P is not attributable to S.<sup>11</sup>

Since one knows that P only if P is true, the indicated assumption entails the extremely implausible thesis that one could not have attributable knowledge of a proposition whose truth is not itself attributable to one (cf. Reed, 2007: 248–9)—for example, a necessarily true proposition which is both self-evident and obvious.

Fortunately, Reed's argument is in fact much more nuanced and, accordingly, much more promising than is the one suggested by the passage just quoted. We can state Reed's argument more fully and formally as follows:

### 2.1 Reed's Incompatibility Argument (2007: 249–258; 2009: 94–101)

- (1) If Attributabilism is true, then there could not be a belief whose status as knowledge depends on the satisfaction of a condition whose obtaining is both "extrin-

<sup>10</sup> For brief critical discussion of the "apt action" interpretation of Reed's argument, see note 13.

<sup>11</sup> This assumption bears some resemblance to the implausibly strong (cf. Zimmerman, 1987: 377–8) thesis about moral responsibility that Thomas Nagel (1979: 25) calls the "condition of control": "...people cannot be morally assessed for... what is due to factors beyond their control."

- sic to the subject's performance and external to the subject's awareness" (2009: 99; cf. 2007: 249–50).
- (2) If Fallibilism is true, then there could be a belief whose status as knowledge depends on its exemplification of justification-independent non-accidental truth—that is, non-accidental truth whose absence is compatible with the pertinent belief's being justified in the way that it actually is.
  - (3) Justification-independent non-accidental truth is both extrinsic to the subject's performance and external to the subject's awareness.
  - (4) If fallibilism is true, then there could be a belief whose status as knowledge depends on the satisfaction of a condition whose obtaining is both extrinsic to the subject's performance and external to the subject's awareness [2,3].
  - (5) If both fallibilism and attributabilism are true, then there both could *and* could not be a belief whose status as knowledge depends on the satisfaction of a condition whose obtaining is both extrinsic to the subject's performance and external to the subject's awareness [1,4].
  - (6) It's false that there both could *and* could not be a belief whose status as knowledge depends on the satisfaction of a condition whose obtaining is both extrinsic to the subject's performance and external to the subject's awareness.
  - (7) Either fallibilism is false or attributabilism is false [5,6].

This argument is valid, and we can't question (2) or (6). But what about (1) and (3)?

Unlike the "transfer of non-attributability" principle discussed earlier, (1) arguably allows for the possibility of one's having attributable knowledge of a proposition whose truth is not itself attributable to one. Reed addresses (1)'s compatibility with the truth requirement on knowledge in the following two passages:

Let me begin... by returning to the attributability of actions—for example, climbing a mountain. This is something that can be fully attributable to a climber (assuming nothing unusual—for example, being carried part of the way by a sherpa guide—happens). Nevertheless, in order to climb a mountain there must be a mountain, a fact that is surely not attributable to the average climber. This does not, however, undermine the attributability of the climber's achievement. The existence of the mountain provides the goal at which she aims. *Given* that the mountain is there, her reaching the top of it is something that can be entirely her own doing. Similarly, in the case of knowledge, the sense in which the truth of a belief is not attributable to the person who has it is irrelevant epistemically. In order to know that it is raining outside, I do not have to make it rain. *Given* that it is raining, however, I may come to know it. The truth of the proposition provides the goal at which I aim. (2007: 249)

The widespread intuitive response to Gettier cases shows that knowledge must be more than merely justified belief that is also true. The something more is that there has to be the proper connection between the subject's justification and the truth of her belief. When that proper connection obtains..., the truth is the attained end of the subject's performance... [T]he truth may be *extrinsic* to

the subject's performance, but, when that performance is properly connected to the truth, the truth then is not *external* to the subject's awareness. (2009: 99)

On the basis of what Reed says in these passages, I'm happy to concede (1)'s compatibility with the truth requirement on knowledge. Moreover, on the basis of what Reed says in the following passage, I'm happy to concede (1)'s compatibility with the extremely plausible claim that knowledge is (at least) reliably produced/sustained belief:

Given that reliability supervenes, not only on the subject's belief-forming processes but on general facts about the surrounding environment as well, we will have to grant that the subject is constituted in part by that environment. Those general facts are external to the person's subjective awareness, but they are *not* extrinsic to the subject's cognitive nature. (2007: 247)

In sum, then, I'm happy to concede that—unlike the previously discussed “transfer of non-attributability” principle—(1) is compatible with the truth and reliability requirements on knowledge.

We'll return to (1) in the next section. As for (3), here's what Reed says in support of the thesis that justification-independent non-accidental truth is extrinsic to the subject's cognitive performance:

[General facts about the subject's surrounding environment] are *not* extrinsic to the subject's cognitive nature. However, the same cannot be said about the environmental factors whose presence may give rise to problematic accidentality. These are different in kind from the much more general features of the surrounding environment that help constitute reliability and are therefore intrinsic to the subject. So, when the factors that might lead to accidentality are absent, this is not something that can be attributed to the subject. (2007: 247-8; cf. 2009: 97)

On the basis of what Reed says in the above passage (and related ones), I'm happy to concede that justification-independent non-accidental truth is extrinsic to cognitive performance.

What does Reed say about (3)'s other element, the thesis that justification-independent non-accidental truth is external to the subject's awareness? Focusing on a subject (Bridget) who forms (under normal conditions) a visual belief that there's a barn in the field, Reed says the following:

The subject's performance is *not* aimed at ensuring that [the relevant anti-accidentality condition] is satisfied... Bridget's performance is directed at determining whether there is a barn in the field in front of her. She is not also trying to determine whether the surrounding fields have genuine barns or mere barn facades. In that sense, Bridget is blind as to whether the anti-accidentality condition has been met... The simple fact that there aren't any facades around her makes no epistemic difference to her. What is problematic about the anti-accidentality condition, then, is that it is both extrinsic to the subject's performance and external to the subject's awareness. Truth and



justification (conceived in an externalistic way) may be one or the other, but they are not both. (2009: 99)

The above passage suggests the following argument for the “external to awareness” portion of premise (3):

In simply forming a belief that P, one isn’t trying to determine (learn, ascertain) whether one is forming a non-accidentally true belief that P. If one isn’t trying to determine whether one is forming a non-accidentally true belief that P, then one isn’t trying to form a non-accidentally true belief that P. So, when one simply forms a belief that P, one isn’t trying to form a non-accidentally true belief that P. So, belief doesn’t aim at non-accidental truth. So, a belief’s exemplification of non-accidental truth is external to the subject’s awareness.

This argument has a false premise—namely, the claim that if one isn’t trying to determine whether one is forming a non-accidentally true belief that P, then one isn’t trying to form a non-accidentally true belief that P. Your simply forming a non-accidentally true belief that P is a different state of affairs from your *determining whether* you’ve formed such a belief. Hence, the indicated premise is false: possibly, one aims to form a non-accidentally true belief that P without also aiming to *determine whether* one is forming such a belief.

Despite this reservation about the “external to awareness” portion of (3), I’m happy to concede premise (3) to Reed. Nevertheless, I believe that Reed’s incompatibility argument fails. Indeed, I believe that the argument defeats itself: premises (1) and (3) are themselves jointly incompatible. Let me explain.

### 3 III

Recall the lottery example described and discussed above (section 1):

An informant whom I justifiedly believe to be extremely reliable tells me that the winning numbers in tomorrow’s pick-your-own-number lottery will be 8675309. On this occasion, my informant is in fact just guessing about tomorrow’s winning numbers. But my informant’s guess turns out to be accurate, for—unknownst to both me and my informant—there exists a sufficiently powerful and knowledgeable third party whose job is to ensure that tomorrow’s winning numbers turn out to be whatever I think they will be. Through my informant’s testimony, I acquire a justified but accidentally true belief that the winning numbers will be 8675309.<sup>12</sup> I acquire an intention to win the lottery by picking 8675309. I intentionally pick those numbers. I win the lottery.

<sup>12</sup> See Coffman (2017: 16–21) for discussion and defense of the thesis that the salient beliefs in cases like this one are “gettiered”—that is, fall short of knowledge in the way illustrated by cases like those Gettier (1963) described. Prominent theorists who would so classify such “hidden helper” examples include Plantinga (1993), Hiller and Neta (2007), Turri (2011), Greco (2012), and Schafer (2014).

As we've already noted, this example can be understood such that I aptly but non-intentionally won the lottery. Notice, though, that we can convert the above case into one in which I intentionally win the lottery by making my belief about the winning numbers non-accidentally true:

An informant whom I justifiably believe to be extremely reliable tells me that the winning numbers in tomorrow's pick-your-own-number lottery will be 8675309. On this occasion, my informant does indeed know that the winning numbers will be 8675309. Through my informant's testimony, I acquire a justified and non-accidentally true belief that the winning numbers in tomorrow's lottery will be 8675309. I acquire an intention to win the lottery by picking 8675309. I intentionally pick those numbers. I win the lottery.

This example can be understood such that I intentionally won the lottery (cf. Mele & Moser, 1994: 60f.; Gibbons, 2010: 81). Reflection on the above pair of cases thus reveals the following modal fact: *There could be an action whose status as intentional depends on a pertinent belief's exemplifying justification-independent non-accidental truth.*

Now, to begin to see the conflict between premises (1) and (3) of Reed's argument, consider the following argument against (1):

### 3.1 An Argument Against (1)

- (8) There could be an action whose status as intentional depends on the satisfaction of a condition whose obtaining is both extrinsic to the subject's performance and external to the subject's awareness.
- (9) If both (8) and attributabilism are true, then there could be a belief whose status as knowledge depends on the satisfaction of a condition whose obtaining is both extrinsic to the subject's performance and external to the subject's awareness.
- (10) Even if attributabilism is true, there could be a belief whose status as knowledge depends on the satisfaction of a condition whose obtaining is both extrinsic to the subject's performance and external to the subject's awareness (hence, ~ 1) [8,9].

This argument is valid, and (9) is extremely plausible.<sup>13</sup> But what about (8)?

It turns out that if (3) is true, then so is (8), and the above argument against (1) succeeds. That's how (1) and (3) conflict. Here's an argument for the claim that (3) entails (8):

If (3) is true, then justification-independent non-accidental truth is a condition whose satisfaction is both extrinsic to the subject's performance and external

<sup>13</sup> Reflection on cases like Sosa's (2011: 26) "magnetic target" example reveals that there could be an action (e.g., hitting a target with a metal-tipped arrow) whose status as apt depends on the satisfaction of a condition whose obtaining is both extrinsic to the subject's performance and external to the subject's awareness (e.g., a nearby magnet's being inoperative). Accordingly, a duly modified version of the above argument against (1) refutes a version of (1) whose antecedent involves "apt action" Attributabilism (rather than "intentional action" attributabilism).

to the subject's awareness. But there could be an action whose status as intentional depends on a pertinent belief's exemplifying justification-independent non-accidental truth. So, if (3) is true, then there could be an action whose status as intentional depends on the satisfaction of a condition whose obtaining is both extrinsic to the subject's performance and external to the subject's awareness (= 8).

This valid argument's first premise is trivially true, and its second premise is just the modal fact revealed earlier by reflection on the two lottery cases. We can conclude, then, that (3) entails (8). But, as we saw in the last paragraph, (8) entails the denial of (1). Hence, (3) entails the denial of (1). I conclude, therefore, that Reed's incompatibility argument fails due to self-defeat.

## References

- Coffman, E. J. (2017). 'Gettiered belief.' In R. Borges, C. de Almeida, and P. Klein (eds.), *Explaining Knowledge: New Essays on the Gettier Problem*. Oxford University Press.
- Fischer, J. M. (2008). 'Compatibilism.' In J. M. Fischer, R. K. Derk Pereboom, and M. Vargas (eds.), *Four Views on Free Will*. Blackwell Publishing.
- Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23, 121–123.
- Gibbons, J. (2010). Seeing What You're Doing. *Oxford Studies in Epistemology*, 3, 63–85.
- Greco, J. (2012). A (Different) Virtue Epistemology. *Philosophy and Phenomenological Research*, 85, 1–26.
- Hiller, A., & Neta, R. (2007). Safety and epistemic luck. *Synthese*, 158, 303–313.
- Mele, A. (1987). Intentional action and wayward causal chains: The problem of tertiary waywardness. *Philosophical Studies*, 51, 55–60.
- Mele, A. (1992). *Springs of action: Understanding intentional behavior*. Oxford University Press.
- Mele, A. (2012). *Backsliding: Understanding weakness of will*. Oxford University Press.
- Mele, A., & Moser, P. (1994). Intentional action. *Nous*, 28, 39–68.
- Nagel, T. (1979). 'Moral luck.' In Thomas Nagel (ed.), *Mortal Questions*. Cambridge University Press.
- Neta, R. (2019). The basing relation. *Philosophical Review*, 128, 179–217.
- Plantinga, A. (1993). *Warrant and proper function*. Oxford University Press.
- Pritchard, D. (2012). Anti-luck virtue epistemology. *The Journal of Philosophy*, 109, 247–279.
- Reed, B. (2002). How to think about fallibilism. *Philosophical Studies*, 107, 143–157.
- Reed, B. (2007). The long road to skepticism. *The Journal of Philosophy*, 104, 236–262.
- Reed, B. (2009). A new argument for skepticism. *Philosophical Studies*, 142, 91–104.
- Schafer, K. (2014). Knowledge and two forms of Non-accidental truth. *Philosophy and Phenomenological Research*, 89, 373–393.
- Sosa, E. (2011). *Knowing full well*. Princeton University Press.
- Sosa, E. (2015). *Judgment and agency*. Oxford University Press.
- Turri, J. (2011). Manifest failure: The Gettier problem solved. *Philosophers' Imprint*, 11, 1–11.
- Williamson, T. (2017). 'Acting on knowledge.' In J. Adam Carter, E. Gordon, and B. Jarvis (eds.), *Knowledge first: Approaches in epistemology and mind*. Oxford University Press.
- Zimmerman, M. (1987). Luck and moral responsibility. *Ethics*, 97, 374–386.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.