Unable to Do the Impossible

Anthony Nguyen
University of Southern California
anthony.nguyen.1@usc.edu

1. Introduction

Jack Spencer argues that we can perform actions that are metaphysically impossible to perform on the basis of cases like (Simple G):

> *Simple G*: Suppose that determinism is true. Let h be the complete specification of the initial conditions of the universe. Let l be the complete specification of the deterministic laws of the universe. Let $h \wedge l$ be their conjunction. Suppose that G has not, does not, and will not believe that $h \wedge l$. G never finds herself reading a book or listening to a radio programme about the initial conditions or the laws of nature; G was home from school and sick with the flu on the day that her physics teacher covered the initial conditions and the laws of nature in class, and the physics teacher never bothered to go over the material again. We may suppose that it is fairly common knowledge in G's community that $h \wedge l$, that matriculating high school seniors are expected to know that $h \wedge l$, that many of G's classmates know that $h \wedge l$, and that G is one of the brightest students in her class. (Spencer 2017, p. 468)

It is metaphysically impossible that G know $h \wedge l$. In any metaphysically possible world, either G believes $h \wedge l$ or G does not. Suppose G believes $h \wedge l$. Necessarily, $h \wedge l$ is true only if G does not believe $h \wedge l$. After all, $h \wedge l$ describes the initial conditions and laws of nature holding in a deterministic world. So, in any possible world in which G believes $h \wedge l$, $h \wedge l$ is false. And truth is necessary for knowledge. So, if G believes $h \wedge l$, she does not know $h \wedge l$. On the other hand, suppose that G does not believe $h \wedge l$. But belief is necessary for knowledge. So G does not know $h \wedge l$. So, whether G believes $h \wedge l$ or not, she does not know $h \wedge l$. Hence, there is no metaphysically possible world in which G knows $h \wedge l$. It is metaphysically impossible for her to know $h \wedge l$. But, according to Spencer, G can know $h \wedge l$.[1]

[1] I use 'is able to' and 'can' interchangeably. I am only concerned with what Vihvelin calls 'wide abilities' (2013, pp. 11-12). Roughly, an agent's wide abilities are the abilities that she has is in virtue of her intrinsic properties and some relevant facts about her surroundings. A pianist with access to a piano has the wide ability to play the piano; on the other hand, a pianist who is imprisoned in a room without a piano does not.

I will assume that compatibilism is true.[2] (Simple G) should only trouble us if compatibilism is true. There is some intuitive appeal to Spencer's judgment on (Simple G). G, after all, is at least as competent as her peers who know $h \wedge l$. But if G can know $h \wedge l$, then the Poss-ability Principle is false:[3]

> *Poss-ability Principle*: For any agent S and action A, if S is able to perform A, then it is metaphysically possible that S perform A. (Spencer, p. 465)

The Poss-ability Principle makes a rather intuitive claim. It would be surprising if we had to give it up. The Poss-ability Principle is, in fact, taken for granted by the two most influential sorts of analyses of ability statements: counterfactual analyses and possibility analyses.

Counterfactual analyses state that, for any agent S and action A, S can perform A iff a relevant counterfactual of the form 'If Ψ had been the case, S would have performed A' is non-vacuously true. G.E. Moore, for instance, tells us that we 'mean by "could" merely "would, *if* so and so had chosen"' (2005, p. 110, emphasis added). For Moore, S can perform A iff S would have performed A had S chosen to perform A.

If G can know $h \wedge l$, then all counterfactual analyses are false. Any counterfactual of the form 'If Ψ had been the case, then S would have performed A' is non-vacuously true only if there is a possible world in which Ψ is the case and S performs A. But since there is no possible world in which G knows $h \wedge l$, there is no non-vacuously true counterfactual with 'G knows $h \wedge l$' as its consequent. But, according to counterfactual analyses, G can know $h \wedge l$ only if some such counterfactual is true. So, according to Spencer, all counterfactual analyses are false. And, historically at least, counterfactual analyses have enjoyed wide popularity.[4]

If G can know $h \wedge l$, possibility analyses share the same fate as counterfactual analyses. Possibility analyses state that, for any agent S and action A, S can perform A iff S performs A in some possible world in a relevant class of possible worlds. David Lewis, for instance, tells us that 'to say that something can happen means that its happening is *compossible* with certain facts' (1976, p. 150, emphasis added). Let W be the

---

[2] I relax this assumption in §4. (Simple G) only poses a threat to the Poss-ability Principle if compatibilism is true. Determinism is true in the world of (Simple G). So, if compatibilism is false, G cannot perform any action that she does not, in fact, perform. So, if compatibilism is false, G cannot know $h \wedge l$.

[3] I grant Spencer that knowing, or learning, is an action.

[4] Cross reports that '[f]or years, philosophers have tinkered with conditional analyses of the meaning of "can"' (1986, p. 58). For a famous counterexample to counterfactual analyses, see Lehrer (1968, p. 32). For a recent descendent of the counterfactual analysis, see Vihvelin (2004). For a critique of this 'new dispositionalism,' see Clarke (2009).

class of possible worlds in which these facts obtain.[5] Then, for Lewis, S can perform A iff there is some possible world in W where S performs A. And Lewis is not alone. Many other philosophers find possibility analyses appealing.[6] Spencer reports that 'if any conception of ability deserves to be called the prevailing view, it is the…possibility analysis' (p. 482).

But, if G can know $h \wedge l$, then all possibility analyses are false. There is no possible world in which G knows $h \wedge l$. So, if any possibility analysis were true, G could not know $h \wedge l$.

There is a second intuitively plausible principle that we have to give up if we share Spencer's judgment on (Simple G):

> *Counterfactual Principle*: For any agent S and action A, if, on any occasion, S would fail to perform A no matter what S tried, then S cannot perform A.[7] (Wasserman 2017, p. 116)

After all, if G can know $h \wedge l$, then 'there is no counterfactual of the form 'S would (or might) φ if S Ψ-ed' that is a necessary condition on S's being able to φ' (Spencer, p. 482). The counterfactual 'S would fail to perform A no matter what S tried' is true if there is no possible world in which S performs A.[8] So, G would fail to know $h \wedge l$ no matter what G tried. So, the Counterfactual Principle entails that G cannot know $h \wedge l$. Therefore, if G can know $h \wedge l$, the Counterfactual Principle is false.

There is a third intuitively plausible principle we have to give up if we share Spencer's judgment on cases like (Simple G):[9]

> *Intensionality Principle*: For any agent S and actions A and B, *if* both
> (i) Necessarily, S performs A iff S performs B; and
> (ii) S can perform A;
> *then*
> (iii) S can perform B.

[5] '*Which* facts? That is determined, but sometimes not determined well enough, by context' (Lewis 1976, p. 150).

[6] See, for example, Kratzer (1977) and Cross (1986). Mele reports that '[p]hilosophers happy to talk in terms of possible worlds will say that an agent in a world *W* is [able] to *A* at *t* if and only if she…*A*-s at *t* in some relevant possible world' (2003, p. 450). For a close relative—the 'cluster analysis'—of possibility analyses, see Brown (1988). The cluster analysis is also false if G can know $h \wedge l$.

[7] For a statement of a closely related principle, see Vihvelin (1995, p. 320).

[8] I ignore the vacuous case. This is acceptable here because, in some possible world, G tries to know $h \wedge l$.

[9] Thanks to John Hawthorne for pointing this out to me.

Assume that G can know $h \wedge l$. Now assume, for *reductio*, that the Intensionality Principle is true. Necessarily, G knows $h \wedge l$ iff G proves that 2 is odd. After all, it is impossible for G to know $h \wedge l$ or to prove that 2 is odd. But since G can know $h \wedge l$, then the Intensionality Principle entails that G can prove that 2 is odd. G, however, cannot prove that 2 is odd. No one can. So, by *reductio*, the Intensionality Principle is false. Therefore, if G can know $h \wedge l$, the Intensionality Principle is false.

There is a fourth intuitively plausible principle we have to give up if we accept that G can know $h \wedge l$:[10]

> *No Self-Undermining Knowledge*: For any agent S and proposition P, if P entails that S does not know P, then S cannot know P.

If G can know $h \wedge l$, then G can know something that entails that she does not know it. After all, $h \wedge l$ entails that G does not know $h \wedge l$.[11] So, if G can know $h \wedge l$, the No Self-Undermining Knowledge Principle is false.

If we accept that G can know $h \wedge l$, then we must reject four appealing principles: the Poss-ability Principle, the Counterfactual Principle, the Intensionality Principle, and the No Self-Undermining Knowledge Principle. To ease discussion, I will henceforth call the disjunction of these four principles 'the Tetrad'.[12] I think it desirable that we be able to consistently maintain the Tetrad. Therefore, in my view, we should reject Spencer's judgment that G can know $h \wedge l$.

Spencer disagrees.[13] Spencer would happily forgo the Tetrad. But my goal is not to convince those who *already* agree with Spencer. My goal is to provide the rest of us a way of avoiding his conclusion.[14]

I proceed as follows. In §2, I critically evaluate reasons to believe G can know $h \wedge l$ in (Simple G). In §3, I offer three explanations for why we might have the misleading intuition that G can know $h \wedge l$. In §4, I discuss a couple of Spencer's cases that are importantly different from (Simple G).

---

[10] Thanks to John Hawthorne for pointing this out to me.

[11] Noah Gordon originally brought this to my attention. Spencer recognizes that $h \wedge l$ entails that G does not know $h \wedge l$. See Spencer (p. 490).

[12] I am interested in the disjunction since, if G can know $h \wedge l$, we must reject the disjunction—and not merely the conjunction—of these four principles.

[13] Spencer explicitly denies the Poss-ability Principle and the Counterfactual Principle. See Spencer (pp. 481-482).

[14] Similarly, it seems foolish to try to convince committed skeptics that skepticism is false. But it is nonetheless a fruitful philosophical endeavor to provide non-skeptics a reasoned way of avoiding skepticism.

2. Cropping the Snapshot Principle Out of the Picture

In my diagnosis, the central theoretical motivation for claiming that G can know $h \wedge l$ in (Simple G) is the Snapshot Principle:

> *Snapshot Principle*: For any action A and agents S and T, *if* S and T are intrinsic duplicates governed by the same laws of nature and are in qualitatively identical situations, *then* S can perform A iff T can perform A.[15] (Vihvelin 2011)

Some of G's worldmates are qualitatively similar to G and can know—and, in fact, know—$h \wedge l$. G is as smart as many of these worldmates. These worldmates, however, are luckier than G insofar as $h \wedge l$ entailed that they know $h \wedge l$. Spencer suggests this sort of reasoning when justifying his claim that G can know $h \wedge l$ in (Simple G): 'In fact, in the vignette above [Simple G], I think that G has the unexercised ability to know that $h \wedge l$. G is able to know that $h \wedge l$, I think, *no less than her fellow classmates*' (p. 469, emphasis added).

The Snapshot Principle, however, is weaker than what is required to share Spencer's judgment on (Simple G). After all, in (Simple G), there is no mention of any intrinsic duplicates of G who are in situations qualitatively identical to G's. There is only mention of intrinsically *similar* worldmates who are in situations qualitatively *similar* to the one G is in. Strictly speaking then, only something like the Strong Snapshot Principle implies that, in (Simple G), G is able to know $h \wedge l$:

> *Strong Snapshot Principle*: For any action A and agents S and T, *if* S and T are sufficiently similar intrinsically, governed by the same laws of nature, and are in situations that are sufficiently similar qualitatively, *then* S can perform A iff T can perform A.

Let us stipulate that, in (Simple G), G is sufficiently similar intrinsically to her peers who know $h \wedge l$. Moreover, let us stipulate that G is in a situation that is sufficiently similar qualitatively to the situations of her peers who know $h \wedge l$. From the Strong Snapshot Principle, it then follows that G is able to know $h \wedge l$.[16]

The Strong Snapshot Principle clearly entails the (regular) Snapshot Principle. Perfect similarity counts as similar enough. Therefore, if the (regular) Snapshot Principle is false, then so is the Strong Snapshot Principle. But if we are to maintain the Tetrad, we must reject the Strong

[15] Vihvelin (2011) denies the Snapshot Principle. She states the Snapshot Principle while arguing that time travelers are unable to kill their younger selves.

[16] Thanks to Troy Cross for helpful discussion here.

Snapshot Principle. I propose, then, to focus our attention on the (regular) Snapshot Principle. Is it true?

The Snapshot Principle is, admittedly, an intuitive principle. If my intrinsic duplicate and I are governed by the same laws of nature, are in qualitatively identical situations, and my intrinsic duplicate can perform some action A, what could stop me from being able to perform A?

But we should be suspicious of the Snapshot Principle. If we accept it, we seem compelled to accept that G can know $h \wedge l$. And if G can know $h \wedge l$, then we have to give up at least four intuitively plausible principles: the Poss-ability Principle, the Counterfactual Principle, the Intensionality Principle, and the No Self-Undermining Knowledge Principle. It's four against one. It looks as if we may be best off cropping the Snapshot Principle out of the picture.

Moreover, there are counterexamples to the Snapshot Principle. To start off with the exotic: *If* we accept the metaphysical possibility of backwards time travel, accept that no adult time traveler can kill her younger self, and also accept that an intrinsic duplicate of Adult Suzy can kill Baby Suzy, *then* an intrinsic duplicate of Adult Suzy can perform some action—killing Baby Suzy—that the time traveler Adult Suzy cannot perform. Adult Suzy cannot kill Baby Suzy. Baby Suzy must live through childhood in order for Adult Suzy to exist.[17] But there is no similar obstacle to an intrinsic duplicate of Suzy's ability to kill Baby Suzy. An intrinsic duplicate of Adult Suzy can kill Baby Suzy. But Adult Suzy unambiguously cannot kill Baby Suzy.

While I am inclined towards this solution to the 'autoinfanticide paradoxes' of time travel, it is controversial.[18] Thankfully, the case against the Snapshot Principle need not rest on any controversial view on the abilities of time travelers. There are more mundane counterexamples to the Snapshot Principle. Here is one:[19]

> *Autobiography*: Ann cannot write an autobiography without writing an account of Ann's life. Alice, Ann's intrinsic duplicate, can write an autobiography without writing an account of Ann's life. Alice need only write an account of her own life and not Ann's.

---

[17] Worlds in which Baby Suzy would be resurrected if Adult Suzy tried to kill Baby Suzy are simply irrelevant to the assessment of Adult Suzy's abilities. These worlds are too remote from the actual world. See Vihvelin (1995, pp. 321-322).

[18] Vihvelin (1995) defends this position on the autoinfanticide paradoxes. For helpful discussion, see Wasserman (2017, pp. 114-130). For a well-known opposing position, see Lewis (1976).

[19] Vihvelin (1995, p. 330 fn 4) develops a similar counterexample.

In (Autobiography), two intrinsic duplicates are such that one of them can perform some action but the other cannot. This is so even if the intrinsic duplicates are governed by the same laws of nature and are in qualitatively identical situations. Therefore, the Snapshot Principle is false.

One might try weakening the Snapshot Principle so as to get around this counterexample. But this cannot be done while also ending up with a principle that yields the verdict that G knows $h \wedge l$ in (Simple G). In (Autobiography), there is an agent S who cannot perform some action A *because* S's performing A would generate a contradiction. S's duplicate, however, can perform A without generating any contradiction. In particular, in (Autobiography), Ann cannot perform the following action that her intrinsic duplicate Alice can perform: *writing an autobiography without writing an account of Ann's life*. Suppose Ann performs this action. Writing an autobiography conceptually requires writing an account of one's own life. Hence, both Ann does and does not write an account of her own life. That a contradiction would be true if Ann were to perform the action in question suffices to show that Ann cannot perform this action.

A perfectly fine explanation of Ann's inability to write an autobiography without writing an account of Ann's life is this: It is metaphysically impossible for Ann to perform this action. However, it would be question-begging against Spencer to appeal to this explanation of Ann's inability. Spencer's view is that one can do the metaphysically impossible. I, however, am appealing to the distinct commonsensical thesis that no one can do the *logically impossible*. This is weaker than the claim that no one can do the metaphysically impossible. It is logically possible for Donald Trump to be a lizard, but this is metaphysically impossible. (Autobiography) is explained if no one can do the logically impossible.[20]

This observation suggests the following weakened version of the Snapshot Principle:[21]

---

[20] Spencer might bite the bullet and insist that some agent *can* do the logically impossible. Not only is this intuitively implausible, this maneuver would preclude a straightforward explanation of why, in (Autobiography), Ann cannot write an autobiography without writing an account of Ann's life.

[21] One might be tempted to take a different line and weaken the Snapshot Principle by restricting it to 'qualitative' abilities, where a qualitative ability is, roughly, an ability whose description need not make reference to any particular individual. Call the resulting principle 'the Snapshot* Principle'. (Autobiography) is not a counterexample to the Snapshot* Principle; this case concerns non-qualitative abilities.

But if we are free to add exceptions to the Snapshot Principle, why not also exclude G-cases from its purview? Let the Snapshot** Principle be just like the Snapshot* Principle except that G-cases are also excluded. Some reason to prefer the Snapshot** Principle is given by (Simple G) itself. If G had the ability to

Weak Snapshot Principle: For any action A and agents S and T, *if* both

(i) S and T are agents that are sufficiently similar intrinsically, governed by the same laws of nature, and are in situations that are sufficiently similar qualitatively, and

(ii) Neither S nor T would generate a contradiction by performing A,

then

(iii) S can perform A iff T can perform A.

But (Simple G) is relevantly similar to (Autobiography). If G were to know $h \wedge l$, then a contradiction would be generated. Suppose G knows $h \wedge l$. Since knowledge is factive, $h \wedge l$ is true. But $h \wedge l$ entails that G does not know $h \wedge l$. And truth is closed under logical entailment. Therefore, if G knows $h \wedge l$, both she knows $h \wedge l$ and she does not know $h \wedge l$! Therefore, just as it is logically impossible for Ann to write an autobiography without writing an account of her life in (Autobiography), it is logically impossible for G to know $h \wedge l$ in (Simple G).[22] If either action were performed, a contradiction would be generated. Therefore, the Weak Snapshot Principle is far too weak to show that G is able to know $h \wedge l$.

It seems that there is no true variant of the original Snapshot Principle that entails that G can know $h \wedge l$ in (Simple G). And I do not

---

know $h \wedge l$, then both (1) the Tetrad would be false and (2) we would lose a natural explanation of why Ann is unable to perform the non-qualitative action in question in (Autobiography)—that no one can do something such that she would generate a contradiction by doing it.

A Spencerian might consider all this question-begging. I do not. Recall the dialectic here. I am entertaining the possibility of justifying Spencer's intuition on (Simple G) in a principled manner. If the best way to do this is to appeal to the Snapshot* Principle, then the burden of proof is on the Spencerian to show why we cannot instead accept the Snapshot** Principle—or some even weaker variant of the Snapshot Principle. The Spencerian has to show that the Snapshot* Principle is preferable to the Snapshot** Principle. Why not also exclude G-cases if we are already excluding cases involving non-qualitative abilities? And in her answer, the Spencerian cannot appeal to G-cases. After all, she is using the Snapshot* Principle to justify her Spencerian intuitions on G-cases. Her task seems hard. Thanks to Troy Cross and Jack Spencer for helpful discussions on this point.

In any case, there may be counterexamples to the Snapshot* Principle. Suppose G discovers curium. Then none of her intrinsic duplicates can discover curium. But discovering curium is plausibly qualitative. Thanks to Cameron Domenico Kirk-Giannini for this potential counterexample.

[22] Thanks to an anonymous referee for pressing me to clarify this point.

see any other appealing theoretical grounds for claiming that G can know $h \wedge l$.[23]

Having realized that nothing like the Snapshot Principle will do the trick, one might try to justify the claim that G can know $h \wedge l$ in (Simple G) by a direct appeal to intuition. One can—it is claimed—directly intuit that G is able to know $h \wedge l$!

Such a direct appeal to intuition here, however, is underwhelming. I like intuitions as much as the next analytic philosopher, but whatever intuition there is that G can know $h \wedge l$ is outbalanced by the intuitive and theoretical weight of the Tetrad. Between the Tetrad and the speculative claim that G can know $h \wedge l$, there is no contest. Keep the Tetrad. Even Spencer seems to concede that it would be underwhelming to directly appeal to the intuition that G can know $h \wedge l$: 'Of course, *Simple G*, is a strange, somewhat artificial example, and such examples should not be asked to pull much philosophical weight' (p. 469).

3. Simple Error Theories for Simple G

But why do we have the intuition that G can know $h \wedge l$ at all? In my diagnosis, there are at least three sources of this misleading intuition.

The first source is the Snapshot Principle itself. We admittedly use something like the Snapshot Principle in ordinary life. If we know that someone very much like me intrinsically can run 10 miles per hour, then we would normally be happy to accept this as strong evidence that I can run 10 miles per hour.

But as I have already argued, the Snapshot Principle is false. Moreover, (Simple G) seems relevantly similar to the counterexample I raised against the Snapshot Principle. That is, G would generate a contradiction if she knew $h \wedge l$. Therefore, while the Snapshot Principle may mislead us into believing that G can know $h \wedge l$, it does not provide any good reason to believe that G can know $h \wedge l$.

A second possible source of the intuition that G can know $h \wedge l$ is that G can do something *extremely close* to knowing $h \wedge l$:

> *Close*: G can know $h \wedge l^*$, where l^* is minimally different from l in such a way that, necessarily, $h \wedge l^*$ is true only if G knows $h \wedge l^*$.

---

[23] Spencer (pp. 475-477) discusses a plausible principle, the Revealing Principle:

> *Revealing Principle*: For any agent S and action A, if both S can perform A and there are enough representative attempts across modal space by S to perform A, then at least one of these representative attempts is a success.

Spencer employs the Revealing Principle in order to explain how, *if* there are counterexamples like (Simple G) to the Poss-ability Principle, such counterexamples to the Poss-ability Principle could arise. I, however, deny that such counterexamples arise in the first place.

(Close) states that G can know some proposition very similar to $h \wedge l$. This proposition is $h \wedge l^*$. Here, l^* is a complete specification of the deterministic laws of nature holding in some world w very close to the world of (Simple G) such that, in w, G knows $h \wedge l^*$.

One might object that since $h \wedge l^*$ is false in the world of (Simple G), it follows that (Close) is false when evaluated at the world of (Simple G). After all, only true propositions are known.

Recall, however, that I am assuming compatibilism.[24] And if compatibilism is true, an agent S in a deterministic world v can perform some action A such that, if S were to perform A, then either the initial conditions or the laws of nature would be different than those holding in v (Lewis 1981, p. 114). Assume, for a toy example, that the actual world is deterministic. If compatibilism is true, I can stand now even though I am, in fact, now sitting. The initial conditions or the laws of nature would just be different than they actually are if I were to stand now.

Suppose that if G were to learn the initial conditions and laws of nature,[25] the laws of nature would be different such that l^*, but not l, is true. Given that (Close) is true, G can do something very similar to knowing $h \wedge l$. G can know $h \wedge l^*$. And given that two actions A and B are very similar, we naturally infer that any agent S can perform B if she can perform A. After all, if it is known that I can run 9.9 miles per hour, it is natural to infer that I can run 10 miles per hour. It does seem that, if actions A and B are very similar, an agent's being able to perform A provides *prima facie* evidence that she can also perform B.

But it may be that I can run, *at most*, 9.9 miles per hour. In that case, I can run 9.9 miles per hour but not 10 miles per hour. Just because an agent S can perform action A and action B is very similar to A, it does not follow that S can perform B. Therefore, just because G can know $h \wedge l^*$, it does not follow that she can know $h \wedge l$. It may nonetheless be tempting, in an unreflective mood, to make this inference.

A third potential source of the intuition that G can know $h \wedge l$ is that the following argument seems sound when evaluated at G's world:[26]

(1) $h \wedge l$ states the initial conditions and laws of nature.[27]
(2) G can know the initial conditions and laws of nature.
(3) Therefore, G can know $h \wedge l$.

---

[24] If compatibilism is false, then the correct judgment to have on (Simple G) is that G cannot know $h \wedge l$. After all, $h \wedge l$ describes a deterministic world.

[25] Give the description 'the initial conditions and laws of nature' narrow scope!

[26] For this method of explaining away the intuition that G can know $h \wedge l$, I am indebted to Noah Gordon.

[27] I will sometimes speak loosely and use 'the initial conditions and laws of nature' to mean *the proposition that completely states the initial conditions and laws of nature, and nothing more*.

(1) is true by stipulation. By stipulation, $h \wedge l$ completely states the initial conditions and laws of nature holding in G's world (and nothing more). (2) seems plausible because G seems to have the cognitive capacity and resources to learn the initial conditions and laws of nature if she tried hard enough. And (3) seems to be a logical consequence of (1) and (2). But (3) states precisely what I have been at pains to deny—that G can know $h \wedge l$.

But the argument is either unsound or invalid. Which is the case depends on what scope is assigned to the definite description 'the initial conditions and laws of nature' in premise (2). If it is given wide scope, then the argument is valid, but unsound. I have already argued that the initial conditions and laws of nature—$h \wedge l$—are *not* such that G can know them.

But if the definite description is instead given narrow scope, then the argument has true premises but is invalid. To see that this is so, consider the analogous argument from (4) and (5) to (6), where 'the tallest man alive' is given narrow scope in (5):

(4) Sultan Kösen is the tallest man alive.[28]
(5) Jones can become the tallest man alive.
(6) Jones can become Sultan Kösen.

Suppose that Jones is a man, that Jones ≠ Kösen, that Kösen is the tallest man alive, and that Jones can grow up to be significantly taller than Kösen's current height. The above argument is then obviously fallacious even though premises (4) and (5) are true.

Analogously, if 'the initial conditions and laws of nature' is given narrow scope in (2), the argument from (1) and (2) to (3) is invalid. Therefore, no matter what scope is assigned to 'the initial conditions and laws of nature' in (2), the argument from (1) and (2) to (3) is unsound. Appearances to the contrary, we may not infer (3) from (1) and (2).

I have identified three sources of the intuition that G can know $h \wedge l$. We might—in ways that are not *obviously* foolish—come to mistakenly believe that G can do the impossible.

That G cannot know $h \wedge l$ is, I believe, well-hidden in (Simple G). In order to see that G cannot know $h \wedge l$, we must appreciate the point that there is *no way* for G to know $h \wedge l$. It seems intuitive that G can know $h \wedge l$ only if this point is not fully appreciated. But, it might be protested, G could have gone to class and paid attention! Right, but this is not truly a way in which G could have come to know $h \wedge l$. After all, the truth of $h \wedge l$ necessitates that G does not do this.

I think (Simple G) is relevantly similar to (Sculptor G):

[28] 'Sultan Kösen' is a proper name of an actually living person, who was, in 2011, measured to be 8 feet and 2.8 inches tall.

> *Sculptor G*: G is an accomplished artist. G desires, more than anything else, to sculpt a beautiful statue. As it so happens, the only metaphysically possible ways to sculpt a beautiful sculpture are by sculpting a statue of shape S_1, or S_2,…, or S_n. An evil neuroscientist, however, has recently implanted a chip in G's brain that makes her psychologically incapable of sculpting all and only statues of shapes S_1, or S_2,…, or S_n.[29]

In (Sculptor G), can G sculpt a beautiful statue? Of course not! There is no way for G to sculpt such a statue. After all, G cannot sculpt anything of shapes S_1, or S_2,…, or S_n. Analogously, in (Simple G), there is no way for G to know $h \wedge l$. No roads G could have taken would have led her to knowledge that $h \wedge l$. The main difference between (Simple G) and (Sculptor G) is that the lack of any route to the performance of the relevant action is better hidden in the former than the latter. The inability is there. It is just expertly camouflaged.[30]

4. The Other Cases: Between a Rock and an Actually Hard Place

Spencer presents other cases like (Simple G). Call these cases 'deterministic G-cases'. In deterministic G-cases, G can purportedly perform some action A even though both (i) anyone can perform A only when $h \wedge l$ is true and (ii) $h \wedge l$ entails that G does not perform A. My response to (Simple G) generalizes to other deterministic G-cases. The main theoretical reason to accept that G can know $h \wedge l$ in any factive G-case is the Snapshot Principle. But no variant of the Snapshot Principle is both true and entails G can know $h \wedge l$. Moreover, whatever intuition there is that G can do the impossible—in any deterministic G-case—is outbalanced by the theoretical weight of the Tetrad.[31] Finally, the three sources of the intuition that G can know $h \wedge l$ in (Simple G) also serve as sources of the corresponding intuition in the other deterministic G-cases.

Spencer develops two more cases against the Poss-ability Principle. I consider them last for two reasons. First, unlike deterministic G-cases, neither requires the truth of determinism in any world. So, incompatibilists should care about these cases. The second reason why I con-

[29] For this case, I am indebted to Troy Cross.

[30] Thanks to Troy Cross and Paul Hovda for helpful discussion on this point.

[31] There is one caveat. In one of Spencer's cases, (Teacherly G), G knows $h \wedge l$ but fails to teach—in a factive sense—anyone $h \wedge l$. So it is metaphysically impossible for G to teach anyone $h \wedge l$. According to Spencer, G can nonetheless teach someone $h \wedge l$. See Spencer (p. 478). (Teacherly G) does not threaten the No Self-Undermining Knowledge Principle. But it does threaten the following related principle:

> *No Self-Undermining Teaching Principle*: For any agent S and proposition P, if P entails that S does not teach P, then S cannot teach anyone P.

sider these two cases last is that they do not concern an agent's ability to know something. So they are not even putative counterexamples to the No Self-Undermining Knowledge Principle, one of the disjuncts of the Tetrad. Let 'the Triad' name the disjunction of the Poss-ability Principle, the Counterfactual Principle, and the Intensionality Principle. Spencer's last two cases do not threaten the Tetrad per se, but they do threaten the Triad. This is motivation enough to resist Spencer's judgments if we can. I evaluate Spencer's final cases one at a time. First is (Rocky H):

> *Rocky H*: H, a normal human, is walking along a path in Siberia. Just to his right is a large rock. There is an object, Rocky, co-located with the rock and just like the rock except it is modally fragile. Rocky has the same color and size as the rock. But Rocky has all of its properties essentially. Nobody ever sees the large rock. Hence, nobody ever sees Rocky.[32] (Spencer, p. 480)

Spencer judges that H can see Rocky, even though it is metaphysically impossible that anyone see Rocky. Rocky has all of its properties essentially. So, it is metaphysically impossible that it be seen. For (Rocky H) to plausibly be metaphysically possible, what Bennett calls 'bazillion thing-ism'[33] has to be true (Bennett 2004, p. 356). So, one who rejects such a plenitudinous ontology will likely rule out (Rocky H) as metaphysically impossible.[34]

For the sake of argument, however, assume that (Rocky H) is metaphysically possible. We can still resist Spencer's claim that H can see Rocky. H cannot see Rocky. But H can perform a very similar action. H can see the rock that Rocky is co-located with. The rock and Rocky are the same color and size. But H's being able to see the rock does not threaten the Triad. After all, the rock is a perfectly ordinary object. There are some possible worlds in which H does see the rock.

Moreover, it lies in Rocky's very essence that no one sees it. It would not be strange if it turned out that no one had the ability to see it. Analogously, if it lies in H's essence that H be human, it would not be strange if no one had the ability to make it so H were not human. I see no cost to denying Spencer's judgment that H can see Rocky. Let us move on, then, to Spencer's final case, (Actual G):

---

[32] Spencer does not present the case exactly as I do. He, for instance, does not call the agent 'H' or the modally fragile object 'Rocky'. But I have not changed any important feature of the case.

[33] For some defenses of bazillion thing-ism, see Fine (1999), Hawthorne (2006), and Leslie (2011).

[34] Spencer is careful to rest little weight on cases like (Rocky H): 'It is controversial whether these sorts of modally fragile objects exist, and I will not assume that they do' (p. 480).

Actual G: G is an actual person, a competent college student, who, as a matter of fact, will never come to believe that the actual world is actual. (Spencer, p. 479)

Spencer's intuition is that (Actual G) describes the actual world.[35] That is, Spencer judges that some actual college student has the ability to believe that the actual world is actual even though it is metaphysically impossible that she believe that the actual world is actual.

I see three ways to understand 'G believes the actual world is actual'. This sentence can be used to report that G believes propositions (7), (8), or (9), where '@' names the actual world:

(7) The world that is actual is actual.
(8) The actually actualized world is actual.
(9) @ is the actual world.

But no matter how we understand 'G believes the actual world is actual', we can resist Spencer's judgment that G has an ability that, necessarily, G fails to exercise.

G cannot fail to believe (7).[36] *Maybe* G would not accept—assent to—the sentence 'The world that is actual is actual'.[37] But that would likely only be because G has not been introduced to the philosophical terms of art 'actual' and 'possible world'.[38] If this philosophical jargon were adequately explained to G, then she would accept 'The world that is actual is actual'. But it is not as though G would have learned anything metaphysically deep. At best, she would have learned another way to express a previous belief of hers. If G can believe (7), then she believes (7). In her philosophical innocence, she might express this belief using the sentence 'The way the universe in fact is *is* the way the universe is'.

(8) is supposedly an improvement on (7), for Spencer's purposes, because (8) is contingent (Spencer, p. 479 fn 19). 'The actually actualized world' rigidly designates the actual world, @. But it is contingent that @ is actual. Another world might have been actual.

As with (7), I have my doubts that any agent may have the ability to believe (8) without, in fact, believing it. For an analogous example, consider proposition (10):

---

[35] Following Lewis (1986, pp. 92-93), I assume that 'actual', and its cognates, are indexicals. At any world *w*, 'actual' picks out *w*.

[36] Spencer agrees. He says that '[w]hat is needed to make the counterexample work is a maximally contingent proposition, not the tautology that the world that is actualized is actualized' (p. 479 fn 19).

[37] Roughly, for any agent A to accept any sentence S is for A to believe that S is a true sentence.

[38] I assume that G is a competent English speaker.

(10) The actual President of the United States, if a unique one exists, is the President of the United States.[39]

I believe that anyone who can believe (10) believes it. (10), like (8), is contingent. It is contingent that whoever is the actual President of the United States be the President of the United States. G, a competent college student, may believe (10) even if she would not accept 'The actual President of the United States, if a unique one exists, is the President of the United States'. Perhaps G just has not been introduced to the philosophical term of art 'actual' yet. If this philosophical jargon were adequately explained to G, then G would accept 'The actual President of the United States, if a unique one exists, is the President of the United States'. But it is not as though G would have learned anything metaphysically deep. At best, G would have learned another way to express a previously held belief. If G can believe (10), then she believes (10). In her philosophical innocence, she might express this belief using the sentence 'Whoever is in fact the President of the United States, if a unique one exists, is the President of the United States'.

Similarly, I deem it impossible for an agent to have a necessarily unexercised ability to believe (8), the proposition that the actually actualized world is actual. G might not accept 'The actually actualized world is actual'. But she nonetheless has always believed (8). In her philosophical innocence, G might express this belief using the sentence 'The way that the universe in fact is, whatever it happens to in fact be, *is* the way the universe is'.

So, neither (7) nor (8) serve to show that G, in (Actual G), can do the metaphysically impossible. Will (9) do the trick? That is, can G believe the singular proposition that @—the actual world—is actual even if it is impossible for G to believe this singular proposition?

I assume that possible worlds are maximal consistent sets of propositions.[40] A set S of propositions is maximal just in case, for any proposition P, S either entails P or S entails not-P. A set S of propositions is consistent just in case it is metaphysically possible for all propositions in S to be jointly true. A set is maximal consistent just in case it is maximal and it is consistent. The actual world, then, is a maximal consistent set of only true propositions.

If an actual agent fails to believe (9), then I deny that she is able to believe it. Assume that G does not believe (9). Then it comes at little

[39] I include 'if a unique one exists' so as to guarantee that (10) is apriori.

[40] For one early statement of the actualist view that possible worlds just are maximal consistent sets of propositions, see Adams (1974). I believe that close analogs of what I say below in the main text hold if possible worlds are maximal consistent sets of sentences, maximal consistent propositions, possibly instantiated world-sized properties, or maximal consistent states of affairs.

cost to deny that G *can* believe (9). @, a maximal consistent set of true propositions, entails that G does not believe (9). If G were to actually—in @—believe (9), a contradiction would be generated. G would both believe (9) in @ and not believe (9) in @.

Perhaps one will object that there may be some possible world *w* besides @ such that, in *w*, G falsely believes (9), the singular proposition that @ is actual. If G were to believe (9), then some close world like *w*—not @—would be actual. This is no contradiction.

For the sake of argument, let us grant that there is such a world *w*. I agree that no contradiction would be generated were *w* actual. But then it is no longer impossible for G to believe (9). By assumption, there is a possible world *w* in which G believes (9). Therefore, even if G were able to believe (9), this does not show that G can do the impossible.

In summary, whether we understand 'The actual world is actual' as expressing (7), (8), or (9), Spencer's case (Actual G) provides no reason to believe that G can do the impossible.

5. Conclusion

None of Spencer's cases provides any compelling reason for us to reject the Triad. Moreover, none of Spencer's deterministic G-cases provide any compelling reason to reject the Tetrad, which is even weaker than the Triad. Perhaps we are unable to do the impossible.[41]

[41] I wish to thank David Boylan, Troy Cross, Noah Gordon, John Hawthorne, Paul Hovda, Cameron Domenico Kirk-Giannini, Ryan Ravanpak, and Jack Spencer for helpful discussions on the topic of this paper. Moreover, I wish to thank audiences at a 2017 Reed College Philosophy Colloquium and at the 2018 Harvard-MIT Graduate Philosophy Conference, where I presented this paper, for helpful discussions. I also wish to thank an anonymous referee and an anonymous editor from *Mind* for their helpful comments. Most of all, I wish to thank Kadri Vihvelin, who gave me numerous encouraging and insightful comments throughout the paper-writing process.

References

Adams, Robert 1974, 'Theories of Actuality', in *Noûs* 8

Austin, J.L. 1970, 'Ifs and Cans', in J.O. Urmson and G.J. Warnock (ed.), *Philosophical Papers* (Oxford: Oxford University Press)

Bennett, Karen 1974, 'Spatio-Temporal Coincidence and the Grounding Problem', in *Philosophical Studies* 118

Brown, Mark 1988, 'On the Logic of Ability' in *Journal of Philosophical Logic* 17

Clarke, Randolph 2009, 'Dispositions, Abilities to Act, and Free Will: The New Dispositionalism' in *Mind* 118

Cross, Charles 1986, '"Can" and the Logic of Ability' in *Philosophical Studies* 50

Fine, Kit 1999, 'Things and Their Parts', in *Midwest Studies in Philosophy* 23

Hawthorne, John 2006, *Metaphysical Essays* (Oxford: Oxford University Press)

Kratzer, Angelika 1977, 'What "Must" and "Can" Must and Can Mean', in *Linguistics and Philosophy* 1

Lehrer, Keith 1968, 'Cans Without Ifs', in *Analysis* 29

Leslie, Sarah-Jane 2011, 'Essence, Plenitude, and Paradox', in *Philosophical Perspectives* 25

Lewis, David 1976, 'The Paradoxes of Time Travel', in *American Philosophical Quarterly*, 13

——1981, 'Are We Free to Break the Laws?' in *Theoria* 3

——1986, *On the Plurality of Worlds* (Oxford: Blackwell)

Mele, Al 2003, 'Agents' Abilities', in *Noûs* 37

Moore, G.E. 2005, in William H. Shaw (ed.), *Ethics* (Oxford: Oxford University Press)

Spencer, Jack 2017, 'Able to Do the Impossible', in *Mind* 126

Vihvelin, Kadri 1995, 'What Time Travelers Cannot Do', in *Philosophical Studies* 81

——2004, 'Free Will Demystified: A Dispositional Account', in *Philosophical Topics* 32

——2011, 'Ability, "Can", and Counterfactuals', in *Kadri Vihvelin's Philosophy Blog*,<http://vihvelin.typepad.com/vihvelincom/2011/05/ability-can-and-counterfactuals.html>

——2013, *Causes, Laws, & Free Will: Why Determinism Doesn't Matter* (Oxford: Oxford University Press)

Wasserman, Ryan 2017, *Paradoxes of Time Travel* (Oxford: Oxford University Press)