

Able to Do the Impossible

JACK SPENCER

Massachusetts Institute of Technology
jackspen@mit.edu

According to a widely held principle—the poss-ability principle—an agent, S, is able to ϕ only if it is metaphysically possible for S to ϕ . I argue against the poss-ability principle by developing a novel class of counterexamples. I then argue that the consequences of rejecting the poss-ability principle are interesting and far-reaching.

There is a principle, widely accepted and often taken for granted, that connects abilities and metaphysical possibilities. I call it the *poss-ability principle*:

If S is able to ϕ , then it is metaphysically possible for S to ϕ .

The poss-ability principle states that each of an agent's abilities must be witnessed by a possible exercising thereof; that is, if S is able to ϕ , then there must be some possible world in which S (exercises her ability and) ϕ -s.¹

As we shall see, quite a lot turns on whether the poss-ability principle is true. The poss-ability principle is a premiss in some of the main arguments for incompatibilism, a consequence of the leading analyses of ability, and often an unstated but critical assumption working in the background. The thesis that an object is visible (or knowable) only if there is some possible world in which the object is seen (or known) relies, implicitly, on the poss-ability principle, as does the widely held thesis that 'ought' implies 'it is metaphysically possible that'.

In this paper I argue against the poss-ability principle by introducing and developing a novel class of counterexamples. I claim that an agent might be able to do what it is metaphysically impossible for her to do. In fact, somewhat more surprisingly, I claim that an agent might be able to do what it is metaphysically impossible to do *tout court*.

¹ Possible worlds are maximally complete possibilities. I assume that incomplete metaphysical possibilities are always contained within complete metaphysical possibilities. So: if there is some possibility in which S ϕ -s, then there is some possible world in which S ϕ -s.

I am not, to be sure, the first philosopher to allege that there are counterexamples to the poss-ability principle. Descartes famously claimed that God's abilities were unconstrained by metaphysical possibility;² the early Leibniz, though less radical than Descartes, also thought that God had abilities that could not be exercised;³ and more recently there has been some discussion of whether the poss-ability principle is true of agents who are situated unusually vis-à-vis time (for example, backwards time-travellers⁴ or agents whose four-dimensional bodies form a closed causal loop⁵). I think that considering deities and temporally unusual agents can be instructive, since it helps to mitigate the knee-jerk impulse in favour of the poss-ability principle, but I will for the most part be putting these sorts of extraordinary agents aside. I am interested in a class of counterexamples that arise, not owing to supernatural powers or strange temporal predicaments, but owing to another reason altogether. If the counterexamples that I develop in this paper are successful, then even ordinary human beings might be counterexamples to the poss-ability principle. In fact, I will go on to claim that one of the agents who serves as a counterexample to the poss-ability principle is *you*.

This paper is divided into two sections. First, I argue that we should reject the poss-ability principle, and then I argue that the consequences of rejecting the poss-ability principle are interesting and far-reaching.

1. The case against the poss-ability principle

I begin by identifying a class of cases—*G-cases*—that have an interesting and unexplored structure. I think that *G-cases* are straightforward

² Descartes seems to think that God is able to will the falsity of any truth, even necessary truths, even logical truths; he writes in a 1630 letter to Mersenne that God was 'just as free to make it not true that the radii of a circle are unequal as he was free not to create the world' (AT 1:152, CSMK 25). For further discussion, see Conee (1991), Frankfurt (1977), and Kaufman (2002).

³ For discussion, see Adams (1994) and Lin (2012).

⁴ Fara (2010) and Lewis (1976), among others, argue that backwards time-travellers are counterexamples to the poss-ability principle; they think that although it is metaphysically impossible for an agent to (permanently) kill her younger self, still a backwards time-traveller might be able to do so. For discussion, see Skow (MS), Vihvelin (1996), and Vranas (2010). As I say below in note 13, I think that there is reason to doubt that time-travellers are able to (permanently) kill their younger selves.

⁵ For an interesting discussion of agents whose four-dimensional bodies form a closed causal loop, see Kment (MS).

counterexamples to the poss-ability principle. In a G-case, G is able to ϕ , I claim, even though it is metaphysically impossible for G to ϕ .

My argument against the poss-ability principle does not rely entirely on an appeal to intuition, however. It is also undergirded by a bit of theory. There is clearly *some* interesting and intimate relationship between abilities and metaphysical possibilities, and, after having introduced G-cases, I will put forward what I see as the correct account of this relationship. I claim that the relationship between abilities and metaphysical possibilities is not primarily metaphysical (as it is standardly thought to be), but primarily epistemological. I think that once we appreciate the epistemological relationship between abilities and metaphysical possibilities, (1) we will see what is right about the poss-ability principle, since there is something right about it; (2) we will see what is wrong about the poss-ability principle, since there is something wrong about it; and (3) we will come to expect counterexamples to the poss-ability principle of exactly the G-case variety.

There are many ways to construct G-cases. The simplest take place in deterministic universes and rely on the compatibilistic assumption that agents might be able to do what they are determined not to do. But there are also G-cases that take place in indeterministic universes. Ultimately I think that compatibilists *and* incompatibilists should agree that the poss-ability principle is false.

1.1 *The structure of a G-case*

More important than any particular G-case is the structure common to them all. I thus begin with the structure and then offer a particular G-case that nicely exemplifies the structure.

In a G-case, one possible world stands apart from the rest of the worlds vis-à-vis the prospects of ϕ -ing. (Strictly speaking, it need not be a single possible world; it could be a class, even an infinite class, of worlds, but the point is easiest to make if we imagine a single world being special.) If ϕ -ing occurs anywhere in modal space, it occurs exclusively at the one possible world. Agents in the one possible world might be able to ϕ , but agents in any other possible world clearly and demonstrably are unable to ϕ . Since there are unexercised abilities, the question arises whether an agent might have the unexercised ability *to* ϕ . And I think that an agent might. Thus the blueprint for a G-case: G has the unexercised ability to ϕ in the one possible world, and lacks the ability to ϕ in every other possible world. There is one possible world in which G is able to ϕ , but there are no possible worlds in which G (exercises her ability and) ϕ -s.

In §§1.3 and 2.6, we will look at other G-cases. For now, I want to focus on a particularly simple example:

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 nature. 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. The proposition that $h \wedge l$ does not exceed G's cognitive wherewithal, either in length or in complexity, and there are no special obstacles preventing G from forming the belief.

Compatibilists, one and all, will agree that G is able to *believe* that $h \wedge l$. I make a stronger claim. I claim that G is able to *truly believe*—or better still—that G is able to *know* or *learn* that $h \wedge l$. G has a *factive* ability.

What are factive abilities?

Factive abilities are a subclass of object-dependent abilities. Object-dependent abilities are abilities to perform object-dependent actions, such as *kissing the Blarney Stone* or *seeing the Statue of Liberty*. Object-dependent actions and abilities are common and familiar. Most of us are able to see the Statue of Liberty, for example, but blind people are not. My sister has an object-dependent ability that I lack: unlike me, she is able to get her dog, Pogo, to roll over.

Factive abilities are object-dependent abilities that depend on facts. Just as the ability to see the Statue of Liberty depends on the existence of the Statue of Liberty, so the ability to know that p depends on it being true that p (or in the language of existence: depends on the existence of the fact that p). Like object-dependent abilities generally, factive abilities are common and familiar. Just think of all of the things that you are able know, learn, discover, and remember!

Factive abilities divide modal space. There are the worlds at which the fact obtains, and there are the worlds at which the fact does not obtain. The worlds at which the fact obtains are the worlds that stand out vis-à-vis the prospects of performing the factive action.

Take $h \wedge l$, for instance. Agents in $h \wedge l$ -worlds might know that $h \wedge l$, but agents elsewhere in modal space clearly do not know that $h \wedge l$ —knowledge is factive. Similarly, agents in $h \wedge l$ -worlds might have the ability to know that $h \wedge l$, but agents elsewhere in modal space clearly lack such an ability—only truths are knowable.

Since there are unexercised abilities, the question arises whether an agent might have the unexercised ability *to know that $h \wedge l$* . And I think that an agent might. In fact, in the vignette above, 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. (If G was unable know that $h \wedge l$, then G should not have been enrolled in the physics class to begin with!)

If G is able to know that $h \wedge l$, then we have a counterexample to the *poss-ability principle*, for there is no possible world in which G knows that $h \wedge l$. A world at which G knows that $h \wedge l$ would have to be a world at which $h \wedge l$ is true. But because the laws of nature are deterministic, every world at which $h \wedge l$ is true is a duplicate of the actual world, and thus every world at which $h \wedge l$ is true is a world at which G fails to come to know that $h \wedge l$. So, while there are worlds at which G is *able* to know that $h \wedge l$, there are no worlds at which G (exercises her ability and) knows that $h \wedge l$.

Of course, *Simple G* is a strange, somewhat artificial example, and such examples should not be asked to pull much philosophical weight. Taken by itself, *Simple G* would be a fairly unconvincing case against the *poss-ability principle*. But G-cases are only a part of my argument. The other part is a non-standard proposal about the relationship between abilities and metaphysical possibilities, and the argument gathers force only when the two parts act in tandem. I thus want to switch gears and say a bit about why G-cases arise. To that end, we need to broach the epistemology of ability.

1.2 Why do G-cases arise?

Abilities are not easy to detect. It is, in general, much easier to detect the actions that spring from an agent's abilities than the agent's abilities themselves. For that reason, in our epistemology of ability, we use inference to the best explanation, inferring from facts about what agents *do* to conclusions about what agents are *able* to do. The best evidence that an agent is able to ϕ is the agent's repeatedly attempting to ϕ and repeatedly succeeding—Michael Jordan's attempts to slam-dunk a basketball are repeatedly successful, so, by inference to the best explanation, Michael Jordan is able to slam-dunk a basketball. The

best evidence that an agent is unable to ϕ is the agent's repeatedly attempting to ϕ and repeatedly failing—my attempts to slam-dunk a basketball are repeatedly unsuccessful, so, by inference to the best explanation, I am unable to slam-dunk a basketball.⁶

Keep in mind the order of explanation. It would be wrong to think that the outcomes of our attempts are explanatorily prior to our abilities (as if first someone successfully tries to slam-dunk a basketball and only thereafter and therefore is able to jump so high!). Successful or unsuccessful attempts are not what *make* agents able or unable. Rather, it is the other way around. Our abilities explain the outcomes of our attempts. Michael Jordan's attempts to slam-dunk a basketball are successes, and mine are not, *because* Michael Jordan is able to slam-dunk a basketball, and I am not.

But abilities are hard to detect. You cannot just look at someone and thereby ascertain what she is, and is not, able to do. So the epistemology of ability runs in the opposite direction of the metaphysics. The metaphysics follows the order of explanation: we use what the agent is, and is not, able to do to explain the outcomes of the agent's attempts. The epistemology follows the order of indication: the outcomes of the agent's attempts indicate what the agent is, and is not, able to do. Successful attempts indicate ability; failed attempts indicate inability.

The method of determining what agents are able to do by looking to the outcomes of their attempts would suffer from a poverty of stimulus if we restricted our attention to present or actual attempts, so we take into consideration non-present and non-actual attempts as well. To keep track of things, we index abilities to times and worlds. (After all, what an agent is able to do changes over time and across possibilities.) The question that we want to answer is whether the agent is presently and actually able to ϕ . Our strategy for answering the question is to look at the agent's attempts to ϕ —both present and non-present attempts; both actual and non-actual attempts—and extrapolate from the outcomes of the attempts to our conclusion.

Of course, not all of an agent's attempts are on an evidential par. Some attempts are epistemically relevant, but others are not. To dramatize the point, suppose that in the future I sprout wings.

⁶ *Terminological note:* I use the word 'attempt' in a broad sense to mean circumstance in which the would-be ability might reveal itself in action. Attempts, in my sense, need not be effortful or purposeful. Walking outside might be an attempt, in my sense, to sneeze. Reluctantly attending history class might be an attempt, in my sense, to learn that James Madison was the fourth president of the United States.

Making the most of my new-found appendages, I repeatedly attempt to fly and repeatedly succeed. Such attempts are, indeed, repeated and successful attempts by me to fly. But everyone recognizes that my future (winged) attempts to fly do not provide good evidence about what I (being unwinged) am presently and actually able to do. My (winged) attempts to fly are not, as I will put it, *representative* attempts.

What are the representative attempts? Hold fixed the facts that together determine whether S is presently and actually able to ϕ , and consider the attempts by S to ϕ done at times and worlds at which those facts obtain—*those* are the representative attempts.⁷ The attempts by S to ϕ that are representative of whether S is presently and actually able to ϕ are those attempts done by versions of S that are neither enhanced nor diminished (as compared to the present and actual version of S), in circumstances that are neither more nor less favourable to S ϕ -ing (as compared to the present and actual circumstances).

The epistemology of ability goes via the outcomes of the representative attempts. The outcomes of the unrepresentative attempts carry little or no information; we can ignore them.⁸ It is the outcomes of the representative attempts that reveal what an agent is, and is not, able to do. Successful representative attempts indicate ability; failed representative attempts indicate inability.

There is a point to make here about the relationship between abilities and counterfactuals. Counterfactuals, as I see them, are epistemological expeditors: they streamline the epistemology of ability by providing a quick and, for the most part, reliable way of identifying representative attempts.

We want to know whether an agent is presently and actually able to ϕ . So consider the closest possible worlds in which the agent attempts to ϕ . What are the outcomes of the agent's *closest* attempts? Are they successes or failures? Counterfactuals draw our attention to the outcomes of the agent's closest attempts. And there is, indeed, a robust correlation: for the most part, an agent is able to ϕ just if

⁷ Controversies about abilities thus re-emerge as controversies about representative attempts. For example, compatibilists and incompatibilists will disagree about whether representative attempts by S to ϕ must occur in worlds with the same past and laws as the actual world.

⁸ Unrepresentative attempts are not always epistemically irrelevant. If S ϕ -s in less favourable circumstances, that may be good evidence that S is presently and actually able to ϕ . Thanks to Eden Lin for discussion on this point.

her closest attempts are successes. I believe, however, that this correlation is indirect. Abilities and close successes are correlated, not because of some direct link between them, but rather because, for the most part, close attempts are also representative attempts. An agent's closest attempts are typically done by versions of the agent that are neither enhanced nor diminished, in circumstances that are neither more nor less favourable. Thus, close attempts, which are often easier to locate than representative attempts, can serve as useful proxies for representative attempts.

The indirectness of the correlation between abilities and close successes explains why the counterfactual test, although often reliable, occasionally misfires. When the closest attempts are representative attempts, the counterfactual test works fine. But when the closest attempts are unrepresentative, the test makes false predictions. There are two sorts of problem.

Problem 1: *The outcomes of the closest attempts might differ from the outcomes of the representative attempts.* Consider a finkish agent, for instance, me. I have puny biceps, and I am hopelessly unable to do ten chin-ups. But suppose that I have a fairy angel who hates to see me fail. If I tried to do ten chin-ups, my fairy angel would intervene, temporarily beef up my biceps, and my attempt would therefore be a success. Let us stipulate that the closest worlds in which I attempt to do ten chin-ups are worlds in which my fairy angel intervenes and supplements my strength. Good for me, I suppose, for counterfactually doing so many chin-ups! But my (beefed-up) attempts to do ten chin-ups are clearly unrepresentative of what I (being puny) am presently and actually able to do. It may be true that if I tried to do ten chin-ups, I would succeed. It may be true that if I tried to do ten chin-ups, I would magically and temporarily gain the ability to do ten chin-ups. But, despite all of that, I am not presently and actually able to do ten chin-ups. My closest attempts are successes, but my closest attempts are not representative.

There are representative attempts by me to do ten chin-ups, of course; they are attempts done by *unenanced* versions of me, that is, versions of me that have my same puny biceps, and they are, for the most part, failed attempts. As I said, I am not able to do ten chin-ups.

If there are representative attempts, then we can mitigate, and even eliminate, the difference between the outcomes of the closest attempts and the outcomes of the representative attempts by supplementing the antecedent of the counterfactual. We began with the first counterfactual test: if S attempted to ϕ , then S would (or might) succeed. We do

better by using the second counterfactual test: if S remained intrinsically very much the same and attempted to ϕ , then S would (or might) succeed. The second counterfactual test is better than the first, because while it is frequently the case that the closest attempts by S to ϕ are representative, it is even more frequently the case that the closest attempts by a not-too-intrinsically-dissimilar version of S are representative. For example, if the closest worlds in which I try to do ten chin-ups are worlds in which my fairy angel supplements my strength, then the first counterfactual test delivers the wrong results, but the second counterfactual test might still get it right.

The second counterfactual test is still imperfect, however, and for the same basic reason: the closest attempts by a not-too-intrinsically-dissimilar version of the agent are not necessarily representative of what the agent is presently and actually able to do. Imagine that my fairy angel operates not by changing my intrinsic nature but by altering my circumstances. When I try to do ten chin-ups, my fairy angel leaves my puny biceps alone, but produces a propitious gust of wind. My (wind-assisted) attempts are no more representative than are my (beefed-up) attempts.

We can continue in this fashion, supplementing the antecedent bit by bit, doing an ever better job of ensuring that the closest, antecedent-satisfying attempts are representative attempts, and thus improving the counterfactual test. For example, we can do better still by using the third counterfactual test: if S remained intrinsically very much the same, and the circumstances remained very much the same, and S attempted to ϕ , then S would (or might) succeed. If we are willing to help ourselves to the notion of representativeness, then we can put an end to this Chisholming and cut straight to the chase.⁹ Consider the fourth counterfactual test: if S made a representative attempt to ϕ , then S would (or might) succeed.¹⁰

The fourth counterfactual test ensures that if there are representative attempts, the closest attempts are among them. This is a dramatic improvement: it eliminates any potential difference between the outcomes of the closest attempts and the outcomes of the representative

⁹ Whether it is legitimate to appeal to representativeness depends on our ambitions. If we seek a reductive analysis of ability, it may be illegitimate, since representativeness should probably itself be understood in terms of ability. I seek a reliable test for abilities, so my appeal to representativeness is not illegitimate.

¹⁰ The fourth counterfactual test resembles Lehrer's (1976) semantics for 'can'.

attempts. In my opinion, the fourth counterfactual test is (qua test of ability) as good as a counterfactual test can be.

But even the fourth counterfactual test is imperfect. The problem is not a potential mismatch between the outcomes of the closest attempts and the outcomes of the representative attempts. Rather, the problem is that there might not be any representative attempts at all.

Problem 2: There might not be any representative attempts. There are at least two ways in which it could turn out that there are no representative attempts. First, the agent might be unable to try. Consider a snail. There are no representative attempts by the snail to prove the Riemann hypothesis. To imagine the snail even attempting to prove the Riemann hypothesis requires imagining a significantly enhanced version of the snail. The same might go for me in a coma. It might be impossible for someone in a coma to attempt to run. If so, and if I am indeed in a coma, then there will be no representative attempts by me to run.

Second, it might be that all of the agent's attempts are done in circumstances that are, as compared to the present and actual circumstances, considerably less favourable. This brings me back, finally, to G and her factive abilities.

It may be helpful to start with a temporal analogy. F is considering whether to visit his grandmother. After some deliberation, he decides not to. As of Monday night, let us suppose, F has never attempted to visit his grandmother, and let us suppose, moreover, that F's grandmother dies on Monday night. If, on Tuesday, F attempts to visit his grandmother, then his attempt will be a failure; his grandmother is not there to be visited. But F's attempt on Tuesday to visit his grandmother is not representative of whether he was, on Monday, able to visit his grandmother, for the circumstances on Tuesday are considerably less favourable to F visiting his grandmother than were the circumstances on Monday.

I think that what goes for F in the temporal case goes equally for G in the modal case. To keep things simple, suppose that G has not, and will not, make an attempt to know that $h \wedge l$, where an attempt would be something like asking someone or looking it up online.¹¹ There are,

¹¹ Since the simplest G-cases are ones in which G makes no attempt to know that $h \wedge l$, one might wonder whether G is able to attempt to know that $h \wedge l$. We can see that G is able to attempt to know that $h \wedge l$ by looking at slightly more complicated G-cases, in which G makes one or more representative attempts:

Mishearing G: G asks her professor about the laws of nature and initial conditions. Her professor tells her, ' $h \wedge l$ ', but G mishears, and thus comes to believe a nearly true but false proposition.

of course, possible worlds in which G attempts to know that $h \wedge l$, but all of them are worlds in which $h \wedge l$ is false, and worlds in which $h \wedge l$ is false are considerably less favourable to G knowing that $h \wedge l$ than are worlds in which $h \wedge l$ is true. Thus, modal space contains many attempts by G to know that $h \wedge l$, and all of them are failures; but none of the attempts are representative attempts, and therefore none of the attempts are epistemically helpful. We learn little about whether F was, on Monday, able to visit his grandmother by learning that he fails to visit his grandmother on Tuesday, and we learn little about whether G is presently and actually able to know that $h \wedge l$ by learning that G fails to come to know that $h \wedge l$ in worlds in which $h \wedge l$ is false. In general, failures to perform an object-dependent action in circumstances in which the object is *absent* tell us little or nothing about whether the agent has the object-dependent ability in circumstances in which the object is *present*.

There is, I believe, a true principle, in the vicinity of the *poss-ability* principle, which connects abilities, attempts, and actions. I call it the *revealing principle*:¹²

If S is able to ϕ , and there are enough representative attempts by S to ϕ , then at least one of the representative attempts is a success.

(*Question*: Why *enough* representative attempts, rather than *any*? *Answer*: Because there are fluky successes and fluky failures. An agent who is able to ϕ might attempt to ϕ , and fail, not because of inability, but simply because of bad luck. *Question*: How many attempts are *enough*? *Answer*: Enough is a matter both of number and diversity. To ensure that the outcomes are not unduly biased by the circumstances, the attempts must occur in qualitatively diverse circumstances. And to avoid the problem of very bad luck, the attempts must occur in different possible worlds. A single representative attempt can end in fluky failure, so, by a stroke of very bad luck, an agent who is able to ϕ might make any number of representative attempts, and *all* of

In *Mishearing* G, G is able to attempt to know that $h \wedge l$, attempts to know that $h \wedge l$, but fails to come to know that $h \wedge l$, as her attempt ends in fluky failure. Thanks to Ned Hall for discussion on this point.

¹² I am actually inclined towards a stronger principle, the long-run principle: if there are enough representative attempts by S to ϕ , then S is able to ϕ iff the long-run proportion of S's representative successes to representative attempts exceeds some minimal threshold. (Different abilities require different thresholds. I am not able to guess the outcome of a coin flip, although I guess right about half of the time; on the other hand, I am able to make a three-point basket, even though I shoot less than 20% from three-point range.) The long-run principle entails the revealing principle.

her attempts might end in fluky failure. But bad luck cannot persist across worlds. If the agent is able to ϕ , and her representative attempts are scattered across sufficiently many worlds, then it is not possible that *all* of her representative attempts end in fluky failure.)

I accept the revealing principle.¹³ If an agent has an ability, then, given enough representative attempts, the ability will reveal itself in action. But to get from the revealing principle to the poss-ability principle, we need an additional premiss. Specifically, we need the premiss that for any S and any ϕ , modal space always contains enough representative attempts by S to ϕ . And it is *this* premiss that I deny. Modal space often contains, but does not *always* contain, enough representative attempts. The revealing principle thus leaves a blind spot, and I claim that within this blind spot lie cases of both ability and inability. Modal space does not contain enough representative attempts by the snail to prove the Riemann hypothesis, and the snail is unable to prove the Riemann hypothesis. On the other hand, modal space does not contain enough representative attempts by G to know that $h \wedge l$, and G is able to know that $h \wedge l$.¹⁴

On this conception of the relationship between abilities and metaphysical possibilities—the epistemological conception—other possible worlds are like additional rooms in a laboratory: their primary purpose is to supply us with additional representative attempts.¹⁵ If we

¹³ Given the revealing principle, we have reason to doubt that time-travellers are able to (permanently) kill their younger selves. For though there may be enough representative attempts by the agent to (permanently) kill her younger self, if auto-infanticide is an impossible crime, none of the attempts are successful.

¹⁴ When modal space fails to contain enough representative attempts, we can sometimes circumvent the problem by going vicarious. Take G. The problem is that none of G's attempts to know that $h \wedge l$ are done in circumstances that are as favourable to G knowing that $h \wedge l$ as are the present and actual circumstances. But there is a second best. We can consider one or more agents who are enough like G and who do indeed make attempts in the favourable circumstances, for example G's classmates. The outcomes of the classmates' attempts are vicariously representative of whether G is presently and actually able to know that $h \wedge l$; they represent what would or might have happened if, per impossible, G had attempted to know that $h \wedge l$ in the favourable circumstances that presently and actually obtain (e.g. had G actually attended the physics lecture that her classmates actually attended). The vicarious attempts are successes—upon attending the lecture, G's classmates come to know that $h \wedge l$ —and this suggests that G too is able to know that $h \wedge l$. Going vicarious does not solve the problem entirely, however, as we shall see in §1.3. Even if we supplement the representative attempts with vicarious attempts, still modal space might fail to contain enough representative attempts.

¹⁵ My epistemological conception of the relationship between abilities and metaphysical possibilities is influenced by Kment's epistemological conception of the relationship between causation and metaphysical possibilities; see Kment (2010, 2014).

want to figure out whether S is able to ϕ , then we could bring her into the lab, put her in a variety of circumstances, and have her repeatedly attempt to ϕ . Extrapolating from the outcomes of her attempts, we could draw an informed conclusion about whether she is, or is not, able to ϕ . Of course, it is not always possible to bring S to the lab for testing, so we face a question: what are we to do when there are not enough actual attempts by S to ϕ to serve as the basis for our extrapolation? Answer: use non-actual attempts. The outcomes of S's non-actual representative attempts give us the same sort of information as do (or would) the outcomes of her actual representative attempts.

The epistemological conception explains both why there are counterexamples to the poss-ability principle and why the counterexamples are rare, strange, and somewhat artificial. Counterexamples to the poss-ability principle are necessarily unexercised abilities. According to the epistemological conception, there are necessarily unexercised abilities for the same reason that there are actually unexercised abilities. The revealing principle is true whether we are considering one world or many. If a set of worlds contains enough representative attempts, then at least one of the representative attempts contained therein is a success. There are actually unexercised abilities because, while S may be able to ϕ , the actual world might fail to contain enough representative attempts. There are necessarily unexercised abilities because, while S may be able to ϕ , the whole of modal space might fail to contain enough representative attempts.

Counterexamples to the poss-ability principle are rare, strange, and somewhat artificial because, in order to construct a case in which modal space fails to contain enough representative attempts, we need to exploit some feature of modal space itself. In the vast majority of cases, modal space contains enough representative attempts—which is why, in the vast majority of cases, agents' abilities are possibly exercised.

1.3 Constructing other G-cases

Here are five interesting varieties of G-cases.

Close calls. Despite it being metaphysically impossible for G to ϕ , G may get arbitrarily close to ϕ -ing. As G gets closer to ϕ -ing, the intuition that she is able to ϕ grows stronger. Here, for example, is a case that exploits the Gettier condition on knowledge:

Gettier G: Curious about the initial conditions and the laws of nature, G decides to ask her mother, who does indeed know that $h \wedge l$. By a curious twist of fate, however, aliens have randomly chosen one human being to

abduct for a single night, and G's mother is the unlucky Earthling. For the duration of the abduction, G's mother is replaced by an alien facsimile, who knows neither that h nor that l . G, who justifiably takes the facsimile to be her mother, asks the facsimile, 'What are the initial conditions and the laws of nature?' The facsimile, trying not to blow its cover, takes a wild guess and improbably gets it right, ' $h \wedge l$ '. On the basis of the facsimile's bad testimony, G comes to the true and justified belief that $h \wedge l$, but G does not know that $h \wedge l$, as G fails the Gettier condition on knowledge. And because G never receives any further evidence, and never changes the basis of her belief, G never comes to know that $h \wedge l$.

G does not, and hence necessarily does not, know that $h \wedge l$, but G seems to be able to know that $h \wedge l$, nevertheless. It is hard to believe that an agent might have the ability to form the true and justified belief that $h \wedge l$, yet lack the ability to know that $h \wedge l$.

Actions that are clearly actions. Perhaps knowing is not an action. There are, however, G-cases that clearly involve actions. Take, for example, the action of *teaching*, where teaching is understood factively:¹⁶

Teacherly G: F and G are both professors at Tech University. Their offices are adjacent. Students often come by to ask about the initial conditions and the laws of nature. Both F and G know that $h \wedge l$. Students flip a coin to decide which professor to ask—heads, F, tails, G. Perchance, the coins have always landed heads. Thus, while F has taught many students that $h \wedge l$, G has never taught anyone that $h \wedge l$.

It seems to me that G, no less than F, is able to teach students that $h \wedge l$; her ability to teach students that $h \wedge l$ is part of what makes her such a good professor, after all. But G's ability to teach students that $h \wedge l$ is unexercised, and hence necessarily unexercised.

Other facts. To construct a factive G-case we need a proposition, the truth of which necessitates that G does not perform some particular factive action. I have focused on $h \wedge l$, but only for reasons of convenience. Other G-cases involve other facts.¹⁷ One interesting class of G-cases, for example, involves facts concerning actuality.¹⁸ Consider,

¹⁶ Thanks to Kieran Setiya for discussion on this point.

¹⁷ Any proposition that specifies a pocket of determinism, even a pocket contained in an otherwise indeterministic universe, will provide an example. Chalmers (2011) proposes (perhaps unwittingly) an interesting case, which I discuss in §2.5. A different sort of example can be found in Leibniz. If the actual world is the best of all possible worlds, then an agent who does not exercise her ability to know that the world is as good as it could be therefore necessarily does not exercise her ability.

¹⁸ Kearns (2011) also discusses facts concerning actuality, arguing that agents might be responsible for impossibilities.

to take the most extreme example, the proposition *that the actual world is actual*. This is a maximally contingent proposition, a truth that would be false if the world were different in any way at all. I know that the actual world is actual, as do you. But not everyone knows that the actual world is actual; many have never even considered the matter. The question arises, then, whether an agent might have the unexercised ability to know that the actual world is actual. And I think so:

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.

I think that G, in *Actual G*, like most college students, has the unexercised (and therefore necessarily unexercised) ability to know that the actual world is actual.¹⁹

Let me pause here to address incompatibilists, in particular. Incompatibilists should not be persuaded by *Simple G*, *Gettier G*, or *Teacherly G*, as all of these cases rely on the assumption that an agent might be able to do what she is determined not to do. But I think incompatibilists should find other G-cases persuasive, among them *Actual G*. We can imagine that the actual world is indeterministic, that agents are free from manipulation and coercion, that agents have agent-causal powers and so on. G, in *Actual G*, does not, and hence necessarily does not, come to know that the actual world is actual, but it would be wrong, I think, to say that G is somehow determined not

¹⁹ Some have objected to *Actual G*. They doubt whether we would have been able to think about the actual world had the world been otherwise, and thus suggest that the ability to entertain the proposition that the actual world is actual cannot be unexercised. I disagree. I think that agents might have the unexercised ability to entertain the proposition that the actual world is actual. But a more concessive response can be given. Consider a hybrid of *Actual G* and *Gettier G*, on which G truly believes, but fails to know, that the actual world is actual. In such a case, G is both able to entertain, and able to know, that the actual world is actual, even though it is metaphysically impossible for G to know that the actual world is actual.

A helpful referee raises a different objection. The proposition that the actual world is actual appears to involve a singular thought about the actual world. What is needed for the counterexample to work is a maximally contingent proposition, not the tautology that the world that is actualized is actualized, so 'the actual world' must designate rigidly. Of course, we can know the actual world by description, as the world that is actually actualized. But can we entertain singular thoughts about the actual world? The question of whether we can entertain singular thoughts about the actual world is vexed and interesting, but I think that we can finesse the matter. Let 'the actual world' be a rigidified description, which rigidly designates the world that is actually actualized. The resultant proposition is then maximally contingent, even if it is not singular. The same dialectic then ensues. If G could have the unexercised ability to know this maximally contingent proposition, then we have a counterexample to the *poss-ability* principle, even by the lights of incompatibilism.

to come to know.²⁰ It is *up to G*, I think, in the relevant sense, whether she comes to know that the actual world is actual, even though it is metaphysically necessary that she will not.

Non-factive cases. Whether there are non-factive G-cases depends on what objects there are. All G-cases involve object-dependent abilities, where the existence of the object necessitates that G does not perform the object-dependent action. If objects are sparse and easy to recombine modally, then there might not be any non-factive G-cases. But some philosophers (myself included) believe that objects are plentiful. For instance, consider a large rock in Siberia. Someone inclined by what Karen Bennett (2004) calls ‘bazillion thing-ism’²¹ thinks that there is a modally fragile object, which is just like the rock, the same size as the rock, the same colour as the rock, co-located with the rock, but different from the rock, because, unlike the rock, it has all of its properties essentially. Suppose that nobody ever sees the large rock, and hence that nobody ever sees this modally fragile object. Is this modally fragile object visible? Might an agent have the ability to see it?

I think that if this modally fragile object exists, then clearly it is a visible object. It is as big as a rock, as brown as a rock, as conspicuous as a rock, and reflects photons just like a rock—indeed, necessarily so! Moreover, agents might have the ability to see it. The fact that this modally fragile object is unseen (and hence necessarily unseen) is not due to a collective lack of ability; it is due, rather, to the fact that modal space affords us so few opportunities to see it.

It is controversial whether these sorts of modally fragile object exist, and I will not assume that they do. Rather, I will continue to focus on factive abilities.

Lonely cases. In a Lonely G-case, G is the only being in all of modal space who is rightly positioned to perform a certain factive action:

Lonely G: Suppose that $h \wedge l$ is true, and hence that determinism is true. G, the only intelligent being in the universe, has lived for millions of years, and in that time developed all of modern science and mathematics. As it happens, G never comes to believe that $h \wedge l$, but this is an historical accident. (Perhaps G never considers the question, or perhaps G makes an uncharacteristic blunder in her calculations and thus comes to a false belief.) We can supply G with all the requisite technology (perhaps G need

²⁰ Being determined not to ϕ requires more than a proposition whose truth necessitates that the agent does not ϕ , else unexercised abilities would be impossible!

²¹ Proponents of bazillion thing-ism include Eklund (2008), Fine (1999), Hawthorne (2006), Johnston (2006), Leslie (2011), and Yablo (1987).

only type the question into her supercomputer) and make G as able-minded as we like.

The following principle is even weaker than the poss-ability principle:

if S is able to ϕ , then it is metaphysically possible to ϕ .

Lonely G is, however, a counterexample to even this very weak principle. G, in *Lonely G*, is able to do something that not a single being in all of modal space manages to do.

1.4 Summary

Before we reject the poss-ability principle, we need not just a class of alleged counterexamples but also a satisfying explanation for why the poss-ability principle has held such appeal. I claim that by coupling G-cases with the revealing principle, we have what we need. The poss-ability principle has seemed true because counterexamples are rare. Agents' abilities are almost always possibly exercised. Counterexamples appear only when we turn our attention to the cases for which modal space fails to contain enough representative attempts, namely, G-cases.

2. Consequences of rejecting the poss-ability principle

The revealing principle is only slightly weaker than the poss-ability principle. Nevertheless, the consequences of rejecting the poss-ability principle and retreating to the revealing principle are interesting and far-reaching.

2.1 Abilities

The first consequence belongs to metaphysics. An analysis of ability is a specification of what it is for an agent to have an ability. The leading analyses of ability take ability to be analysed in terms of metaphysical possibility; they are, in that sense, *modal* analyses. If we reject the poss-ability principle, then modal analyses of ability fail.

There are two main sorts of modal analysis of ability: counterfactual analyses and restricted possibility analyses. Let me start with counterfactual analyses. According to a counterfactual analysis: S is able to ϕ iff S would (or might) ϕ if S ψ -ed. The most famous counterfactual analysis is the so-called conditional analysis, due to G. E. Moore (1912), according to which S is able to ϕ iff S would ϕ if S tried to ϕ . On the standard Stalnaker-Lewis semantics for counterfactuals, 'S would ϕ if S tried to ϕ ' is true iff some or all of the closest worlds in

which S tries to ϕ are worlds in which S ϕ -s. Similarly, 'S would (or might) ϕ if S ψ -ed' is true iff some or all of the closest worlds in which S ψ -s are worlds in which S ϕ -s.

G-cases make trouble for counterfactual analyses in the obvious way. In a G-case, G is able to ϕ , even though there is no possible world in which G ϕ -s. In a Lonely G-case, G is able to ϕ , even though there is no possible world in which anyone ϕ -s. G-cases thus suggest that 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 . An agent might have an ability, even though there is no circumstance in which she would (or might) exercise the ability.

G-cases also make trouble for restricted possibility analyses. There is a sense of 'can'—often called the 'can' of ability—for which the following biconditional is uncontroversially true: S is able to ϕ iff S can (in the special sense) ϕ . The 'can' of ability denotes a particular grade of possibility, which I shall here label *agentive possibility*. We thus have three equivalent ways of speaking: S is able to ϕ iff S can (in the special sense) ϕ iff it is agentively possible for S to ϕ .

According to the restricted possibility analysis, agentive possibility is a restricted grade of metaphysical possibility, where x -possibility is a restricted grade of y -possibility iff to be x -possible is to be y -compossible with a certain set of facts. The paradigm is nomic possibility, which is a restricted grade of metaphysical possibility: to be nomicly possible is to be metaphysically compossible with the nomic facts. If agentive possibility is likewise a restricted grade of metaphysical possibility, then there is some (perhaps time-, individual- and context-relative) set of facts—we may call them the *agentive facts*—such that to be agentively possible is to be metaphysically compossible with the agentive facts. For S to be able to ϕ , according to the restricted possibility analysis, is for it to be metaphysically compossible with the agentive facts that S ϕ -s.

If any conception of ability deserves to be called the prevailing view, it is the restricted possibility analysis.²² But if the poss-ability principle

²² See, for example, Lewis (1976) and van Inwagen (1983). Mele says, 'Philosophers happy to talk in terms of possible worlds will say that an agent at a world W is [able] to A at t if and only if she (or a counterpart) A-s at t in some relevant possible worlds ... One way to see the disagreement about incompatibilists and compatibilists about determinism and being able to do otherwise is as a disagreement about what worlds are relevant. According to incompatibilists, all and only worlds with the same past and natural laws as W are relevant; they hold the past and the laws fixed. Compatibilists disagree' (Mele 2003, p. 451). Vetter treats the restricted possibility of analysis as given, and 'propose[s] that we extend the restricted possibility of abilities to dispositions' (Vetter 2014, p. 133).

is false, then so too is the restricted possibility analysis. After all, no matter what the agentive facts are, it is not metaphysically compossible with *them* that G knows that $h \wedge l$!²³

2.2 The ability to do otherwise

The second consequence also belongs to metaphysics. One of the most important arguments for incompatibilism is the Consequence Argument.²⁴ If the poss-ability principle is false, however, then the Consequence Argument is unsound.

The Consequence Argument, as it is standardly presented, involves a sentential operator, ‘N’, and two rules of inference that govern its application, Rule (α) and Rule (β). According to Rule (α): $\Box p \vdash Np$. (In English: if it is metaphysically necessary that p , then nobody is able to render it false that p .²⁵) According to Rule (β): $Np, N(p \rightarrow q) \vdash Nq$.²⁶ (In English: if nobody is able to render it false that p , and nobody is able to render it false that p materially implies that q , then nobody is able to render it false that q .) Whereas Rule (β) is famously controversial, many regard Rule (α) as a virtual

²³ Objection: Even if existing modal analyses of abilities entail the poss-ability principle should thus be rejected, couldn’t we formulate a new modal analysis that did not entail the poss-ability principle? Reply: Yes. One option is a counterpart-counterpart theory, according to which S is able to ϕ iff in the right circumstances a counterpart of S would (or might) perform an action that is a counterpart of ϕ -ing. Given the right sort of counterpart relation—if, for example, believing that $h \wedge l$ is a counterpart of knowing that $h \wedge l$ —the counterpart-counterpart theory may evade the problem of G-cases. But we should still accept that G-cases are counterexamples to the poss-ability principle, and we should still accept the epistemological conception. If by accepting the counterpart-counterpart theory we evade the problem of G-cases, we do so only by increasing the pool of representative attempts. The counterpart-counterpart theory is extensionally adequate (if it is) only because (i) the revealing principle is true, and (ii) since attempts by counterparts of S to do actions that are counterparts of ϕ -ing now count as representative of whether S is presently and actually able to ϕ , modal space always contains enough representative attempts. Thanks to Sarah-Jane Leslie and Steve Yablo for discussion on this point.

²⁴ See, for example, Ginet (1966) and van Inwagen (1983).

²⁵ I gloss N in terms of ability, since I am concerned with interpretations of N on which $N(S \text{ does not } \phi)$ entails that S is not able to ϕ . As a helpful referee points out, however, one could adopt alternative understandings of N. For instance, in the quoted passage below, van Inwagen glosses N in terms of choice: Np iff nobody has a choice about whether p . I think that van Inwagen, himself, even when he glosses N in terms of choice, thinks that $N(S \text{ does not } \phi)$ entails that S is not able to ϕ , since he thinks that freedom is the ability to do otherwise. But one could sever the connection between abilities and choices. The result would be an interesting conception of choice, on which, though some agent is able to render it false that p , nobody has a choice about whether p .

²⁶ Here and throughout, ‘ \rightarrow ’ and ‘ \leftrightarrow ’ are the material conditional and biconditional, respectively.

truism. Kadri Vihvelin (2011, §5) says, ‘Rule Alpha is uncontroversial’; Alicia Finch and Ted A. Warfield (1998, p. 517) say, ‘Rule Alpha is surely unobjectionable’. Peter van Inwagen himself says,

I do not see how anyone could reject Rule (α). If (α) is invalid, then it could be true that someone has a choice about what is necessarily true. Hardly anyone besides Descartes has been willing to concede such a capacity to God. No one, so far as I know, has ever suggested that human beings could have a choice about what is necessarily true. (van Inwagen 1983, p. 96)

I, however, do reject Rule (α). I also reject van Inwagen’s intimation that it would take a creature of godly powers to invalidate Rule (α). I think that Rule (α) is invalid, even when restricted to human beings. G-cases are counterexamples. It is necessarily false that G knows that $h \wedge l$, but G is able to know that $h \wedge l$. Hence, for any interpretation of ‘N’ on which $N(S \text{ does not } \phi)$ implies that S is unable to ϕ , Np does not follow from $\Box p$. Moreover, if $N(p \rightarrow q)$ follows from $\Box q$, as presumably it does, then G-cases are also counterexamples to Rule (β).²⁷

The failure of the poss-ability principle implies, then, that the Consequence Argument, at least as it is presently formulated, is unsound.²⁸ As menacing as this conclusion may appear, I suggest that incompatibilists see it less as a menace than as a challenge. In my view, everyone should reject the poss-ability principle, compatibilists and incompatibilists alike. And I suspect that the Consequence Argument can be formulated in a way that does not rely on the poss-ability principle. But the failure of the poss-ability principle does present a challenge: it forces incompatibilists to say more about how, exactly, determinism poses a threat to freedom. (I here understand freedom as the ability to do otherwise.)

According to a long and storied tradition, the threat of determinism is none other than the threat of necessity. On this way of seeing things, freedom requires that it be contingent that we act as we do;²⁹ determinism implies that it is necessary that we act as we do; and the

²⁷ There are other ways to formulate the Consequence Argument. In Finch and Warfield’s (1998) presentation, for instance, the Consequence Argument relies on a single rule of inference, namely, Rule (β -2): $Np, \Box(p \rightarrow q) \vdash Nq$. (Rule (β -2) also features in Lamb’s (1977) formulation of the Consequence Argument.) G-cases are counterexamples to Rule (β -2).

²⁸ I focus on the Consequence Argument, but the poss-ability principle is a premiss in many arguments that leverage metaphysical necessity against freedom. See, for example, Holliday’s (2012) Action-Type Argument for incompatibilism, and Pike’s (1965) argument for the incompatibility of freedom and foreknowledge.

²⁹ Frankfurt (1969) argues that freedom does not require that it be contingent that we act as we do. Fara (2008) and Vihvelin (2013) argue against Frankfurt. I am convinced by Fara and Vihvelin.

fundamental question in the compatibility debate is whether the contingency required by freedom is compatible with the necessity implied by determinism. The Consequence Argument is an attempt to lay bare the incompatibilists' view that the necessity implied by determinism excludes the contingency required by freedom. According to the Consequence Argument: (1) if determinism is true, then, holding fixed the past and the laws, it is metaphysically impossible for agents to act otherwise; (2) nobody has any choice about the past and the laws, so the facts about the past and the laws should be included among the agentive facts; and (3) the contingency that freedom requires is metaphysical contingency holding fixed the agentive facts. From (1), (2) and (3), incompatibilism follows.

There is, of course, a great deal of controversy about (2).³⁰ But what the failure of the *poss-ability* principle shows us is that, even putting such controversies aside, (3) is false. The Consequence Argument, at least as it is presently formulated, relies on a mistaken conception of the contingency that freedom requires. The contingency that freedom requires is not metaphysical contingency; hence it is not metaphysical contingency holding fixed some set of facts. Thus the challenge for incompatibilists: why, exactly, does determinism pose a threat to freedom?³¹ The threat of determinism cannot simply be the threat of metaphysical necessity (holding fixed some facts), for S might be able to ϕ , even though it is metaphysically necessary that S does not ϕ .

For compatibilists, the failure of the *poss-ability* principle is interesting, because it brings into focus the question of whether metaphysical contingency matters at all. Does the freedom enjoyed by agents in deterministic universes crucially depend on the past or the laws being metaphysically contingent? Many compatibilists have thought that agents in deterministic universes act freely only if (and only because) the past or the laws are metaphysically contingent.³² But if we reject the *poss-ability* principle, then there is room for the opposing view, that agents might act freely, even in the absence of metaphysical contingency—which, in so far as I am inclined toward compatibilism, is

³⁰ See, for example, Holliday (2012), Huemer (2000), and Lewis (1981)

³¹ This question, I think, can be answered, though I suspect that different incompatibilists will answer it differently. A fan of agent causation (see, for example, O'Connor 2000) might think that the ability to do otherwise is the ability to agent-cause yourself to do otherwise, and that the causal necessity associated with determinism excludes the possibility of agent-causation. So, while the ability to ϕ might be compatible with it being metaphysically impossible to ϕ , the ability to ϕ is not compatible with being determined not to ϕ .

³² See, for example, Lewis (1981), Fischer (2006a), and Vihvelin (2013).

the sort of view that I would want to defend. Such a view might be labelled an ‘actual sequence’ conception of the ability to do otherwise.³³ It says that determinism might be true, the past and the laws might be metaphysically necessary, and still agents might act freely.

2.3 ‘Can’

The third consequence belongs to semantics. In semantics, the ‘can’ of ability is standardly treated as a possibility modal.³⁴ Oversimplifying, if the ‘can’ of ability is a possibility modal, then there is a space of worlds, and a way of selecting a subset thereof, such that $\llbracket \text{can } \phi \rrbracket$ is true iff there is at least one world among the selected subset at which $\llbracket \phi \rrbracket$ is true. If the poss-ability principle is false, then one of two consequences follows. Either we must abandon the project of treating the ‘can’ of ability as a possibility modal³⁵ or the space of worlds that we use for semantics must include worlds that are not metaphysically possible.³⁶

2.4 ‘Ought’ and moral responsibility

The fourth consequence belongs to metaethics. There are several alleged connections between moral notions and metaphysical modality. I believe, however, that these connections are indirect; they are mediated by a direct connection between moral notions and ability. And since the poss-ability principle is false, these alleged connections admit of counterexamples.

Consider the following G-case, in which G’s failure to ϕ is wilful, deliberate, and morally weighty:

Wilful G: Suppose that $h \wedge l$ is true, and hence that determinism is true. G receives a one-question true-or-false quiz. The question: ‘True or

³³ Wolf (1990) seems to have this sort of ‘actual sequence’ view of the ability to do otherwise. For a defence of an ‘actual sequence’ conception of moral responsibility, see Fischer (2006b).

³⁴ See, for example, Krazter (1977, 1981).

³⁵ See, for example, Vetter’s (2013) semantics for ‘can’, which does not invoke possible worlds.

³⁶ Note that the usual problems associated with moving from a possible worlds semantics to a possible-and-impossible worlds semantics (see, for example, Berto 2013 and Stalnaker 1996) are exacerbated in the present context, because we are interested in factive abilities. For example, if, at w , the impossible world at which G knows that $h \wedge l$, it is both true and false that G believes that $h \wedge l$, then it would seem to follow that G is able to know a proposition that she does not believe, since there is a relevant (albeit impossible) world in which she does. But G is not able to know propositions that she does not believe.

false: $h \wedge l$? As it happens, G's father is up for the PTA Father of the Year Award, which he both yearns for and deserves to win. G knows that her father will lose the award if she fails the quiz. Out of nothing but wilful and bitter teenage rebellion, G intentionally fails the true-or-false quiz. G, who knows full well that $h \wedge l$, incorrectly answers that $h \wedge l$ is false.

Correctly answering that $h \wedge l$ is true is a factive action; it is an action that most of G's classmates perform, an action that, in my view, G is able to perform, yet an action that it is metaphysically impossible for G to perform. Here, where ability and metaphysical possibility part ways, moral notions seem to me to travel together with ability.

I claim that G ought (in the moral sense) to correctly answer that $h \wedge l$ is true, and therefore that G does as she ought not to do.³⁷ There may be a true version of 'ought' implies 'can'. If so, it is the principle that 'ought' implies 'it is agentively possible that': ought implies *ability*. The principle that 'ought' implies 'it is metaphysically possible that' is false; *Wilful G* is a counterexample. It is metaphysically impossible for G to correctly answer that $h \wedge l$ is true, but G ought (in the moral sense) to correctly answer that $h \wedge l$ is true.

Similarly, I claim that G is morally responsible and blameworthy for failing to correctly answer that $h \wedge l$ is true (or equivalently: G is morally responsible and blameworthy for incorrectly answering that $h \wedge l$ is false). Typically, if an agent ought to ϕ , but cannot, the agent is thereby excused for failing to ϕ . But the fact that it is metaphysically impossible for G to correctly answer that $h \wedge l$ is true does not seem to excuse G for failing to do so. There might be a true version of the principle of possible action.³⁸ If so, it is the principle of agentively possible action: an agent is morally responsible for failing to ϕ only if it is agentively possible that the agent ϕ . The principle of metaphysically possible action is false; *Wilful G* is a counterexample. It is metaphysically impossible for G to correctly answer that $h \wedge l$ is true, but G

³⁷ *Objection*: It is not true that G ought to correctly answer that $h \wedge l$ is true. What is true is that G ought not to have the bitter thoughts and desires that lead her to incorrectly answer that $h \wedge l$ is false, but it is metaphysically possible for G to lack those bitter thoughts and desires. *Reply*: Yes, G ought not to have those bitter thoughts and desires, but, in addition, G ought not to do as she does. Consider a psychological duplicate of G, G*, who similarly knows that $h \wedge l$, who similarly knows about her father's desire for the PTA Father of the Year Award, who has the same bitter thoughts and desires as does G, but who, for whatever reason, correctly answers that $h \wedge l$ is true, perhaps unintentionally. It stands to reason that G has done more wrong than G* has. G has more to regret, more to atone for, than does G*. Thanks to Caspar Hare and Steve Yablo for discussion on this point.

³⁸ The principle of possible action was introduced by van Inwagen (1978).

is morally responsible and blameworthy for failing to correctly answer that $h \wedge l$ is true.³⁹

2.5 *The blankables and, in particular, knowability*

The fifth and final consequence belongs to philosophical logic. The *blankables* are a subclass of words in English that end ‘-able’ or ‘-ible’; they include the words ‘visible’, ‘perceptible’, ‘discoverable’, ‘thinkable’, ‘provable’, ‘fallible’, and ‘conquerable’.⁴⁰ The blankables are to be analysed in terms of ability. A visible object is an object that can, in the ability sense, be seen. A knowable object is an object that can, in the ability sense, be known.

Like ability, the blankables are often analysed in terms of metaphysical possibility. The claim that x is visible is regimented as the claim that it is possible that x be seen: ‘ $\diamond Sx$ ’. The claim that p is knowable is regimented as the claim that it is possible that p be known: ‘ $\diamond Kp$ ’.

If the poss-ability principle were true, then the modal regimentation of the blankables might be legitimate. But the poss-ability principle is false, and the modal regimentation of the blankables is not just illegitimate, but also a source of philosophical confusion.

Start with *exercisability*. Some have objected to my negative thesis—that G-cases are counterexamples to the poss-ability principle—on the grounds that abilities must be exercisable. But this objection misses its mark. I fully agree that abilities must be exercisable. The crucial question is whether we should understand exercisability in terms of ability or in terms of metaphysical possibility. I say that something might be exercisable even though it is necessarily unexercised. G’s ability to know that $h \wedge l$ is a good example.

In general, the distinction between being necessarily unblanked and being unblankable strikes me as neglected but important. I suspect

³⁹ Whether there are obligations that are impossible to satisfy bears on a number of questions in deontic logic. For example, some have attempted to understand obligation as a restricted grade of metaphysical necessity. Where d is the claim that all obligations are met, obligation is then understood as compossibility with d . It is unclear whether this ‘reduction of deontic logic to alethic modal logic’, as Anderson (1958) calls it, can succeed, however, if there are impossible obligations. For more on the Andersonian-Kangerian reduction, see Kanger (1957), Lokhorst (2006), and Mares (1992).

⁴⁰ There is another class of words that end ‘-able’ or ‘-ible’, which include the words, ‘admirable’, ‘desirable’, ‘valuable’, and ‘enviable’. These words, at least on their most natural readings, are of a kind with ‘trustworthy’ and ‘newsworthy’; they are to be analysed, not in terms of ability, but in terms of fittingness or worthiness. To be admirable, for instance, is to be a fitting object of admiration.

that some perceptible things are necessarily unperceived (I gave a possible example in §1.3); I suspect that some tangible things are necessarily untouched; and I suspect that some knowable things are necessarily unknown. In the remainder of the paper, I want to focus on knowability. The fact that something can be knowable yet necessarily unknown can help to shed light, I think, on two paradoxes of knowability.

Fitch's paradox of knowability. Anti-realists of a Dummettian bent hold that there is a knowability constraint on truth: all truths are knowable.⁴¹ There is a famous argument against the knowability constraint, first published by Frederic Fitch (1963). The argument begins by regimenting knowability as ' $\diamond K$ ', and proceeds as follows:

- | | |
|---|---|
| (1) $p \rightarrow \diamond Kp$. | premiss, the knowability constraint |
| (2) $q \wedge \neg Kq$ | premiss, some truth is unknown |
| (3) $(q \wedge \neg Kq) \rightarrow \diamond K(q \wedge \neg Kq)$ | from (2), by the knowability constraint |
| (4) $\diamond K(q \wedge \neg Kq)$ | from (2) and (3), by sentential logic |
| (5) $\diamond(Kq \wedge K\neg Kq)$ | from (4), by $\Box(K(p \wedge q) \rightarrow (Kp \wedge Kq))$ |
| (6) $\diamond(Kq \wedge \neg Kq)$ | from (5), by the factivity of knowledge |

But (6) is false; contradictions are not possibly true. Yet the other premisses are true, and the inferences are valid, so the knowability constraint stands refuted—or so says Fitch's argument.

I think that we can expose a flaw. We should respond to Fitch in the form of a dilemma. If ' $\diamond Kp$ ' means that it is possible that p be known, then the argument may be sound, and its conclusion may be true, but the argument says nothing about the knowability constraint. The argument is then a refutation of the 'possibly known' constraint on truth: all truths are possibly known. One could propose a 'possibly known' constraint on truth, I suppose, but that was not the anti-realists' proposal. The anti-realists proposed a knowability constraint on truth, and the knowability constraint does not imply the 'possibly known' constraint. A proposition might be knowable yet necessarily unknown. I am suggesting that $K(q \wedge \neg Kq)$ may be one such example.⁴²

⁴¹ Dummett wavers on what, exactly, the knowability constraint is supposed to come to; see Dummett (2001, 2009). For a good recent discussion, see Hand (2014).

⁴² Fara (2010) also suggests that $K(q \wedge \neg Kq)$ might be knowable but necessarily unknown. Fara does not, however, discuss G-cases, the epistemological conception, or the bearing of the failure of the poss-ability principle on the compatibility debate or on metaethics.

On the other horn of the dilemma, if ‘ $\diamond Kp$ ’ is stipulated to mean that p is *knowable*, then the move from (4) to (5) is invalid. The fact that a conjunction is knowable does not imply that there is a possibility in which each of the conjuncts is known.

I myself do not accept the knowability constraint on truth, but my advice to those who do is get clearer about the relationship (and distinction) between being knowable and being possibly known. Fitch’s argument does not obviously give us reason to abandon the knowability constraint on truth.

Let me briefly switch from offence to defence. Some who have seen or heard this paper have thought that the argument that I run from the failure of the poss-ability principle to the potential knowability of Fitch propositions should be run *modus tollens* against me. They treat the unknowability of Fitch propositions (i.e. propositions of the form $p \wedge \neg Kp$) as a fixed point, and believe that any theory that implies the knowability of Fitch propositions thus stands refuted. Does rejecting the poss-ability principle necessarily entail that Fitch propositions are knowable?

One argument suggests that it does. Suppose that G is able to know that $h \wedge l$. The truth of $h \wedge l$ necessitates that G does not come to believe that $h \wedge l$, so G (or $G+$, a suitably enhanced version of G) ought to be able to know that $\Box((h \wedge l) \rightarrow \neg K(h \wedge l))$. By a plausible principle of closure, then, G (or $G+$) ought to be able to know that $h \wedge l \wedge \neg K(h \wedge l)$. But this is a Fitch proposition, hence the allegedly threatening conclusion: G -cases are counterexamples to the poss-ability principle *only if* Fitch propositions are knowable.

I respond to this objection by taking issue with it at three points. First, the argument that Fitch propositions are always unknowable remains to be seen. The standard argument relies on the poss-ability principle, but I reject the poss-ability principle. At least in the present context, another argument is needed. Second, necessitation is not always epistemically transparent. Although $\neg K(h \wedge l)$ may be necessitated by $h \wedge l$, it need not be scrutable on the basis of $h \wedge l$, so there is no reason to think that G (or $G+$) ought to be able to know that $\Box((h \wedge l) \rightarrow \neg K(h \wedge l))$. Thus, even if the threatening argument is sound—which I doubt very much—there would still be *some* G -cases that did not lead to the knowability of Fitch propositions, and so the main claims of the paper would stand. But third, and most importantly, if we start from the premiss that Fitch propositions are always unknowable, then we can show that a number of plausible-sounding closure principles fail. Knowability is not closed under

conjunction introduction; nor it is closed under known implication.⁴³ And although it is hard to find examples that are completely uncontroversial, if Fitch propositions are always unknowable, then I am inclined to think that knowability also fails to be closed under known necessary implication.⁴⁴ I am thus unconvinced by this Fitch-y objection. Perhaps G is able to know a Fitch proposition; and if not, if Fitch propositions are always unknowable, for instance, then G might still be able to know both that $h \wedge l$ and that $\Box((h \wedge l) \rightarrow \neg K(h \wedge l))$. G-cases thus would still be counterexamples to the poss-ability principle.⁴⁵

⁴³ Here is a counterexample to the closure of knowability under conjunction introduction, assuming that Fitch propositions are always unknowable. Let p be a mathematical truth. Suppose that S is the best mathematician in the history of the world, and knows as much. With an hour of concerted effort, S could prove that p . I assume that, if S is able to prove that p , S is able to know that p . So S is able to know that p . S knows, however, that she will never come to know whether p . She knows that it would take her an hour of concerted effort to determine whether p , and she knows that she will never devote an hour of concerted effort to determining whether p . S also knows that nobody else knows whether p , as she knows that she alone possesses the mathematical prowess needed to determine whether p . So S knows that p is unknown. Knowledge implies knowability, so S is able to know that $\neg Kp$. But if Fitch propositions are always unknowable, then S is not able to know that $p \wedge \neg Kp$, so knowability is not closed under conjunction introduction.

A similar argument shows that knowability is not closed under known implication. Imagine a list containing one thousand mathematical statements: p_1, \dots, p_{1000} . S knows that exactly one statement on the list is true. Suppose that p_i is the lone truth. S is able to prove that p_i and hence S is able to know that p_i . Moreover, for each p_j , $p_i \neq p_j$, S is able to prove that p_j is false. For every statement on the list, p_m , S knows that she will never come to know whether p_m ; for she knows that it would take her an hour of concerted effort to determine whether p_m , and she knows that she will never devote such time to determining whether p_m . S also knows that nobody else knows whether p_m , as nobody else possesses the requisite mathematical prowess. Performing a simple deduction, S comes to know, for each p_m , that $p_m \rightarrow (p_m \wedge \neg Kp_m)$. In particular, S knows that $p_i \rightarrow (p_i \wedge \neg Kp_i)$. But if Fitch propositions are always unknowable, then S is not able to know that $p_i \wedge \neg Kp_i$, so knowability is not closed under known implication.

⁴⁴ I am tempted by a hypothesis, suggested to me by Nilanjan Das, that the only non-trivial operations under which knowability is closed are single-premiss operations. Consider a logically valid deduction from some premisses, p_1, p_2, \dots, p_n to q . If each premiss, p_i , can be known without q being known, then, since knowability is not closed under conjunction introduction, we can construct a model in which each of the premisses is knowable but the conclusion is not, since there is no accessible world in which the premisses are known together. This give us model-theoretic reason to think that knowability should fail to be closed even under known necessary implication.

⁴⁵ Thanks to Nilanjan Das, Katie Elliott, Wes Holliday, Matthew McGrath, Bernhard Salow, Kieran Setiya, James Shaw, Jonathan Vogel, and Ian Wells for discussions of knowability and closure.

Chalmers's paradox of knowability. A second paradox of knowability has recently been put forward by David Chalmers:

It is widely believed that for all p , or at least all entertainable p , it is knowable a priori that (p iff actually p). It is even more widely believed that for all such p , it is knowable that (p iff actually p). There is a simple argument against these claims from four antecedently plausible premisses ... [Let] 'A', 'E', 'K', '□', '◇' stand for 'Actually', 'Someone entertains', 'Someone knows', 'Necessarily', and 'Possibly', while '→' and '↔' are the material conditional and biconditional. In addition, q is any (entertainable and expressible) proposition that no-one actually entertains, while r is $\neg E q$, the proposition that no-one entertains q .

- (1) $A r$
- (2) $A r \rightarrow \Box A r$
- (3) $\Box(K(r \leftrightarrow A r) \rightarrow (r \leftrightarrow A r))$
- (4) $\Box(r \rightarrow \neg K(r \leftrightarrow A r))$
- (5) $\neg \Diamond K(r \leftrightarrow A r)$

... The conclusion follows from the premisses by classical logic and weak modal logic **K** ... From (3) and (4), one can derive $\Box(K(r \leftrightarrow A r) \rightarrow ((r \leftrightarrow A r) \wedge \neg r))$. From (1) and (2), one can derive $\Box(K(r \leftrightarrow A r) \rightarrow A r)$. From these two claims, one can derive $\Box(K(r \leftrightarrow A r) \rightarrow (r \leftrightarrow \neg r))$, from which the conclusion follows. (Chalmers 2011, pp 411–12)

The intuitive argument goes like this. Every world is either an r -world or a $\neg r$ -world. There is no r -world in which someone knows that $r \leftrightarrow A r$, because there is no r -world in which someone entertains that $r \leftrightarrow A r$. One must entertain that $r \leftrightarrow A r$ in order to know that $r \leftrightarrow A r$, and one must entertain that q in order to entertain that $r \leftrightarrow A r$, but the r -worlds are precisely the worlds in which no one entertains that q . At the same time, there is no $\neg r$ -world in which someone knows that $r \leftrightarrow A r$, as $r \leftrightarrow A r$ is false at $\neg r$ -worlds.

I think that Chalmers's argument can be resisted. As before, we should respond in the form of a dilemma. If '◇' and '□' denote metaphysical possibility and necessity, then the argument might be valid and its conclusion might be true, but the paradox is falsely advertised. Its conclusion, then, is that it is metaphysically impossible for someone to know that $r \leftrightarrow A r$. I accept this conclusion. But the fact that $r \leftrightarrow A r$ is necessarily unknown does not imply that $r \leftrightarrow A r$

is unknowable. There are knowable truths that are necessarily unknown. I gave one example in *Lonely G*, and I think that $r \leftrightarrow Ar$ is another example.

Chalmers's argument has the paradoxical conclusion that it advertises only if ' \diamond ' and ' \square ' denote agentive possibility and necessity. But if ' \diamond ' and ' \square ' denote agentive possibility and necessity, then premiss (2) is false. Although Ar might imply that it is *metaphysically* necessary that Ar , it does not imply that it is *agentively* necessary that Ar . (Put another way, the fact that no one actually entertains that q does not imply that no one is able to actually entertain that q .) We are supposed to accept (2) as a consequence of a more general principle, namely, $\forall p(Ap \rightarrow \square Ap)$. But if ' \square ' denotes agentive necessity, then $\forall p(Ap \rightarrow \square Ap)$ is the principle of actualized fatalism; it says that agents are able to *actually do* only what they actually do. And just as the principle of fatalism (i.e. $\forall p(p \rightarrow \square p)$) stands refuted by unexercised abilities, so the principle of actualized fatalism stands refuted by unexercised actualized abilities. If I reach a fork in the road, then I am able to *actually go left* and able to *actually go right*, even though I actually go only one way or the other. In the same sort of way, the fact that I do not actually entertain that q does not imply that I am not able to actually entertain that q ; my ability might be unexercised (and hence necessarily unexercised).⁴⁶

We should, in the end, then, reject Chalmers's argument. Those who believe that, for all entertainable p , it is knowable that (p iff actually p) may continue doing so.

But Chalmers's argument is still very interesting, I think. Chalmers has drawn attention to a particularly interesting factive ability—the

⁴⁶ Chalmers (2011, pp. 417–18) considers an objection along these lines: 'One ... strategy appeals to agentive possibility ... holding that a proposition is knowable when it is agentively possible for someone to know it, while denying that agentive possibility entails metaphysical possibility. This position faces an obvious challenge, however, in that the original argument might be reformulated in terms of agentive possibility, and the case for the four key premisses remains strong when read this way.' I respond, as I do in the text, that the case for (2) is not strong when read in terms of agentive possibility.

In personal communication Chalmers has asked me whether we could introduce an actuality-like operator, \mathcal{A} , which, still reading ' \square ' as agentive necessity, makes premiss (2) true. I answer: perhaps we could, and it would be interesting if we could, but introducing this operator does not save the paradox. The widespread belief that all entertainable truths of the form (p iff actually p) are knowable concerns the familiar notion of actuality, which does not imply agentive necessity. If there is an operator, \mathcal{A} , which stands to agentive necessity as the familiar notion of actuality stands to metaphysical necessity, then Chalmers has proved that there are entertainable but unknowable truths of the form (p iff $\mathcal{A}p$). But \mathcal{A} is unfamiliar, so the degree to which we should be surprised by this conclusion is unclear.

ability to know that $r \leftrightarrow Ar$ —and proved that this ability is necessarily unexercised. Although necessarily unexercised, this ability is nevertheless widely possessed. Just about anyone who is able to entertain that $r \leftrightarrow Ar$ is able to know that $r \leftrightarrow Ar$. You and I and others like us are able to know that $r \leftrightarrow Ar$, for example—which is interesting. For the fact that you and I and others like us are able to know that $r \leftrightarrow Ar$ implies that, like G, in *Lonely G*, we are able to do something that not a single being in all of modal space manages to do. You and I and others like us literally are able to do the impossible.⁴⁷

References

- Adams, Robert Merrihew 1994: *Leibniz: Determinist, Theist, Idealist*. Oxford: Oxford University Press.
- Anderson, Alan Ross 1958: ‘A Reduction of Deontic Logic to Alethic Modal Logic’. *Mind*, 67, pp. 100–3.
- Bennett, Karen 2004: ‘Spatio-temporal Coincidence and the Grounding Problem’. *Philosophical Studies*, 118, pp. 339–71.
- Berto, Francesco 2013: ‘Impossible Worlds’. In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy (Winter 2013 Edition)*. <http://plato.stanford.edu/archives/win2013/entries/impossible-worlds/>.
- Brand, Myles, and Douglas N. Walton (eds.) 1976: *Action Theory*. Dordrecht: Reidel.
- Chalmers, David J. 2011: ‘Actuality and Knowability’. *Analysis*, 71, pp. 411–19.
- Conee, Earl 1991: ‘The Possibility of Power Beyond Possibility’. *Philosophical Perspectives*, 5, pp. 447–73.
- Descartes, René 1897: *Oeuvres de Descartes*. Edited by C. Adam and P. Tannery. Paris: Cerf.
- 1991: *The Philosophical Writings of Descartes, Volume III: The Correspondence*. Translated by John Cottingham,

⁴⁷ Thanks to Selim Berker, David Chalmers, Sam Cumming, Nilanjan Das, Shamik Dasgupta, Katie Elliott, Nate Gadd, Gabriel Greenberg, Ned Hall, Caspar Hare, Liz Harman, Pamela Hieronymi, Wes Holliday, Mark Johnston, Boris Kment, Sarah-Jane Leslie, Eden Lin, John Maier, Matthew McGrath, Calvin Normore, Joe Rachiele, Gideon Rosen, Bernhard Salow, James Shaw, Kieran Setiya, Jonathan Vogel, Ian Wells, and Steve Yablo; to *Mind*’s editor and referees; and to audiences at Princeton (2011, 2012), University of Pittsburgh (2013), MIT (2013), UCLA (2013), University of Missouri (2013), University of Oklahoma (2013), and Harvard (2014).

- Robert Stoothoff, Dugald Murdoch, and Anthony Kenny.
Cambridge: Cambridge University Press.
- Dummett, Michael 2001: 'Victor's Error'. *Analysis*, 61, pp. 1–2.
- 2009: 'Fitch's Paradox of Knowability'. In Salerno 2009, pp. 51–2.
- Eikmeyer, Hans-Jürgen, and Hannes Riemer (eds.) 1981: *Words, Worlds, and Contexts: New Approaches in Word Semantics*. Berlin: de Gruyter.
- Eklund, Matti 2008: 'The Picture of Reality as an Amorphous Lump'. In *Sider, Hawthorne and Zimmerman* 2008, pp. 382–96.
- Fara, Michael 2008: 'Masked Abilities and Compatibilism'. *Mind*, 117, pp. 843–65.
- 2010: 'Knowability and the Capacity to Know'. *Synthese*, 173, pp. 53–73.
- Finch, Alicia, and Ted A. Warfield 1998: 'The Mind Argument and Libertarianism'. *Mind*, 117, pp. 843–65.
- Fine, Kit 1999: 'Things and their Parts'. *Midwest Studies in Philosophy*, 23, pp. 61–74.
- Fischer, John Martin 1994: *The Metaphysics of Free Will*. Malden, MA: Blackwell.
- 2006a: *My Way: Essays on Moral Responsibility*. New York: Oxford University Press.
- 2006b: 'A Framework for Moral Responsibility'. In *Fischer 2006a*, pp. 1–37.
- Fitch, Frederic 1963: 'A Logical Analysis of Some Value Concepts'. *Journal of Symbolic Logic*, 28, pp. 135–42.
- Frankfurt, Harry G. 1969: 'Alternative Possibilities and Moral Responsibility'. *Journal of Philosophy*, 66, pp. 829–39.
- 1977: 'Descartes on the Creation of the Eternal Truths'. *Philosophical Review*, 86, pp. 36–57.
- Ginet, Carl 1966: 'Might We Have No Choice?' In *Lehrer* 1966, pp. 87–104.
- Hand, Michael 2014: 'Antirealism and Truths Never Known'. *Journal of Philosophy*, 111, pp. 113–34.
- Hawthorne, John 2006: *Metaphysical Essays*. Oxford: Oxford University Press.
- Hilpinen, Risto (ed.) 1971: *Deontic Logic: Introductory and Systematic Readings*. Dordrecht: Springer.
- Holliday, Wesley H. 2012: 'Freedom and the Fixity of the Past'. *Philosophical Review*, 121, pp. 179–207.

- Huemer, Michael 2000: 'Van Inwagen's Consequence Argument'. *Philosophical Review*, 109, pp. 525–44.
- Johnston, Mark 2006: 'Hylomorphism'. *Journal of Philosophy*, 103, pp. 652–98.
- Kanger, Stig 1957: 'New Foundations for Ethical Theory'. In *Hilpinen 1971*, pp. 36–58.
- Kaufman, Dan 2002: 'Descartes's Creation Doctrine and Modality'. *Australasian Journal of Philosophy*, 80, pp. 24–41.
- Kearns, Stephen 2011: 'Responsibility for Necessities'. *Philosophical Studies*, 155, pp. 307–24.
- Kment, Boris 2010: 'Causation: Determination and Difference-Making'. *Noûs*, 44, pp. 80–111.
- 2014: *Modality and Explanatory Reasoning*. Oxford: Oxford University Press.
- Kment, Boris MS: 'Freedom, Luck, and the Principle of Sufficient Reason'. Unpublished manuscript.
- Kratzer, Angelika 1977: 'What "Must" and "Can" Must and Can Mean'. *Linguistics and Philosophy*, 1, pp. 337–55.
- 1981: 'The Notional Category of Modality'. In *Eikmeyere and Rieser 1981*, pp. 38–74.
- Lamb, James 1977: 'On a Proof of Incompatibilism'. *Philosophical Review*, 86, pp. 20–35.
- Lehrer, Keith (ed.) 1966: *Freedom and Determinism*. New York: Random House.
- 1976: ' "Can" in Theory and Practice: A Possible World Analysis'. In *Brand and Walton 1976*, pp. 241–70.
- Leslie, Sarah-Jane 2011: 'Essence, Plenitude, and Paradox'. *Philosophical Perspectives*, 25, pp. 277–96.
- Lewis, David 1976: 'The Paradoxes of Time Travel'. *American Philosophical Quarterly*, 13, pp. 145–52.
- 1981: 'Are We Free to Break the Laws?' *Theoria*, 47, pp. 113–21.
- Lin, Martin 2012: 'Rationalism and Necessitarianism'. *Noûs*, 46(3), pp. 418–48.
- Lokhorst, Gert-Jan C. 2006: 'Andersonian Deontic Logic, Propositional Quantification, and Mally'. *Notre Dame Journal of Philosophical Logic*, 47, pp. 385–95.
- Mares, Edwin J. 1992: 'Andersonian Deontic Logic'. *Theoria*, 58, pp. 3–30.
- Mele, Alfred R. 2002: 'Agents' Abilities'. *Noûs*, 37, pp. 447–70.
- Moore, G. E. 1912: *Ethics*. London: Williams and Norgate.

- O'Connor, Timothy 2000: *Persons and Causes: The Metaphysics of Free Will*. Oxford: Oxford University Press.
- Pike, Nelson 1965: 'Divine Omniscience and Voluntary Action'. *Philosophical Review*, 74, pp. 27–46.
- Salerno, Joe (ed.) 2009: *New Essays on the Knowability Paradox*. Oxford: Oxford University Press.
- Sider, Theodore, John Hawthorne, and Dean Zimmerman (eds.) 2008: *Contemporary Debates in Metaphysics*. Malden, MA: Blackwell.
- Skow, Bradford MS: 'Notes on the Grandfather Paradox'.
- Stalanker, Robert 1996: 'Impossibilities'. *Philosophical Topics*, 24, pp. 193–204.
- van Inwagen, Peter 1978: 'Ability and Responsibility'. *Philosophical Review*, 87, pp. 201–24.
- van Inwagen 1983: *An Essay on Free Will*. Oxford: Oxford University Press.
- Vetter, Barbara 2013: '“Can” without Possible Worlds: Semantics for Anti-Humeans'. *Philosophers' Imprint*, 13(16), pp. 1–27.
- 2014: 'Dispositions without Conditionals'. *Mind*, 123, pp. 129–56.
- Vihvelin, Kadri 1996: 'What Time Travelers Cannot Do'. *Philosophical Studies*, 81, pp. 315–30.
- 2011: 'Arguments for Incompatibilism'. In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*. <http://plato.stanford.edu/archives/spr2011/entries/incompatibilism-arguments/>.
- 2013: *Causes, Laws, and Free Will: Why Determinism Doesn't Matter*. Oxford: Oxford University Press.
- Vranas, Peter B. M. 2010: 'What Time Travelers May Be Able to Do'. *Philosophical Studies*, 150, pp. 115–21.
- Wolf, Susan 1990: *Freedom Within Reason*. Oxford: Oxford University Press.
- Yablo, Stephen 1987: 'Identity, Essence, and Indiscernibility'. *Journal of Philosophy*, 84, pp. 293–314.