Oracles and Time Machines

**Introduction**

Suppose determinism is true. Now suppose you find out that someone is an oracle of a certain kind. She has complete knowledge of the initial conditions of the universe and the laws of nature, and the ability to accurately predict facts about what will happen by deducing her predictions from this information. Now, suppose the oracle tells you that she knows with absolute certainty that you will order coffee in the next five minutes. The oracle’s prediction doesn’t initially strike you as very impressive. After all, you are with the oracle in a coffee shop, and you are visibly tired and in need of coffee. Still, she’s so confident she promises to give you a thousand dollars if her prediction is incorrect. Eager both to make money and to refute the oracle, you plan not to order coffee in the next five minutes. You look at your watch, determined to wait until five minutes pass before ordering coffee—make that six minutes just to be safe. Even better: you plan to wait until six minutes has passed and then satisfy your caffeine fix by ordering tea instead. “I’ll show you, oracle!” you think to yourself. All you have to do is not order coffee in the next five minutes. What could be easier than that?

This case generates the so-called “paradox of predictability.” On the one hand, the oracle’s predictions carry a lot of weight. After all, given that determinism is true and this is what allows the oracle to make accurate predictions, it doesn’t seem possible for anything you do to falsify what she says. She predicts you will order coffee in the next five minutes. It seems you will. On the other hand, you will try not to order coffee in the next five minutes, and it is extremely easy for you not to order coffee in the next five minutes. It seems you will *not*. Although this case is not genuinely *paradoxical*, there is something puzzling afoot. It both seems that you will and it seems that you will not order coffee. Which is true?

Michael Scriven (1965) presents a seminal version of this sort of case. Some theorists, including Stefaan Cuypers and Stefan Rummens (2009) and Jenann Ismael (2016: 170-181), think the lesson to be learned is that it is impossible for there to be an infallible predictor who (a) makes predictions about what you will do, and (b) shares those predictions with you. Ismael reaches this conclusion because she thinks you would be able to “thumb your nose” at any purportedly infallible predictor. That is, you would be able to falsify the predictor’s prediction. If anyone claims to know with absolute certainty that you will order coffee in the next five minutes, you can easily prove them wrong. Ismael infers two things from this. First, she infers that the laws must be a certain way to resolve the “paradox”—viz. the laws should be thought of in a Lewisian way, as global patterns in local matters of fact. Second, she infers that people have a certain kind of freedom that is compatible with determinism.[[1]](#footnote-1) These inferences are related. The view of laws is posited as a general solution to the paradox of predictability that allows us to recognize that classical determinism doesn’t pose a threat to freedom.

We will try to show two things here. First, even if adopting a specific view of laws of nature allows for a solution to the paradox as framed by determinism, it does not allow for a general solution to the paradox of predictability, since the paradox can be framed without relying on determinism. Second, it is wrong to conclude from the paradox that it is impossible for there to be an infallible predictor who shares many predictions with people about their behavior.[[2]](#footnote-2) We will explain the real lesson to learn from the story of the oracle in the coffee shop: not that it is impossible for an infallible predictor to share many predictions with people about their behaviorbut merely that it is *unlikely* that this will ever happen. In order to explain this, we will rely on Katrina Elliott’s (2018) discussion of time-travel to the past. As we will explain, theorists should think of infallible predictors in much the same way that Elliott thinks of time machines.

**Non-Deterministic Predictions**

Ismael may be right that adopting a Lewisian understanding of the laws allows one to resolve the paradox of predictability in some cases where determinism plays an important role. This, however, does not provide a general solution to the paradox of predictability. One reason why is that equally puzzling situations can arise even if determinism is false. Let us quickly describe two cases to show this.

Suppose indeterminism is true. Now suppose God tells you that God knows with absolute certainty that you will order coffee in the next five minutes. On the one hand, God’s predictions carry a lot of weight. It’s God. God knows everything. God knows you’ll order coffee in the next five minutes. It seems like you will. On the other hand, you really want to take the opportunity to spoil God’s prediction, and it is extremely easy for you not to order coffee in the next five minutes. It seems you will *not*.

Take another scenario in which you find out that someone is a time traveler from the future. The time traveler tells you that she knows with absolute certainty that you will order coffee in the next five minutes. She saw you do it already. On the one hand, the time traveler’s predictions carry a lot of weight. She was there. She saw you order coffee. It seems like you will. On the other hand, you really want to take the opportunity to spoil the time traveler’s prediction, and it is extremely easy for you not to order coffee in the next five minutes. It seems you will *not*.

These scenarios show the paradox of predictability doesn’t require the truth of determinism to arise. Determinism’s being true is just one conceivable way one could come to meet an infallible predictor. Since the underlying problem could conceivably arise in an indeterministic world, we should want a more general solution than the adoption of a certain view about the laws. In order to address what we think is the real lesson to be learned from the case of the oracle in the coffee shop, we will further discuss time travel.

**Time Travel**

Imagine a time traveler, Frank. Frank completely forgets about his sister’s birthday, and she doesn’t hear from him on her birthday. She receives many phone calls, text messages, and emails but none from Frank. (Nor does she receive any communication given by a third party on behalf of Frank.) The next day Frank realizes his error and immediately calls his sister to wish her a belated happy birthday. After the conversation, Frank remembers something important: he has a time machine! So, he uses the time machine to travel to yesterday, his sister’s birthday. Upon arrival he tries to use his cell phone to call her to wish her a happy birthday.

What will happen next? Well, we’ll let you finish the story. But here’s something that won’t happen, provided this mundane time-travel story stays logically consistent: Frank won’t successfully wish his sister Happy Birthday on her birthday. For, as stipulated above, she doesn’t hear from Frank on that day. Something will happen, then, to thwart Frank’s attempt to reach out to her. Maybe his phone won’t work. Maybe he’ll borrow another phone that also doesn’t work. Maybe he’ll go to the local library to use a computer to email his sister but on the way slip on a banana peel and be rushed to a hospital. Maybe he’ll type the email and attempt to send it, but it will fail to send because of a technological issue he doesn’t even notice. No matter how hard Frank tries, and no matter how many attempts he makes—via phone, email, text, or telegram—Frank will not successfully wish his sister Happy Birthday on this day.

This story might seem contradictory. On the one hand, here Frank is on his sister’s birthday with the knowledge that it is her birthday and a desire to wish her a happy birthday. There are so many ways for him to reach out to her on this day, and they all are easy. It is easy to call her with his cell phone. It is easy to borrow somebody else’s phone. It is easy to email her. Given how easy it is to wish her Happy Birthday it seems Frank *can* wish his sister Happy Birthday on the right day. On the other hand, given that Frank will not wish his sister Happy Birthday on this day, every way in which he tries to wish her Happy Birthday will fail. This consideration makes it seem that, alas, he *cannot* wish her Happy Birthday on the right day. So, it seems that Frank both can and cannot wish her Happy Birthday on the right day. And this might seem like a contradiction. One might infer from this that time travel to the past is metaphysically impossible.

That would be a mistake. David Lewis shows that there is no contradiction. To say someone can do something means that their doing it is “compossible with certain facts.” (Lewis, 1976: 149). Relative to the fact (let us suppose) that you ski frequently and are in good shape, you *can* ski. Relative to the fact that you are currently reading a philosophy paper and are nowhere near a snowy mountain, you *cannot* ski. The claim that you can ski and the claim that you cannot ski are both true, because each claim is made in relation to different facts.[[3]](#footnote-3) Likewise, relative to the fact that Frank has access to the marvels of modern technology he can wish his sister Happy Birthday on her birthday. Relative to the fact that his sister never hears from him on this day, he cannot wish her Happy Birthday on her birthday. No contradiction. There might be some completely independent reason why time travel to the past is metaphysically impossible, or at least inconsistent with our physical laws, but one should not infer from the story of Frank alone that time travel to the past leads to any contradiction or is metaphysically impossible.

Instead, one should conclude, as Katrina Elliott (2018) does, that frequent time travel to the past is unlikely. Elliott’s argument is a variant of an argument given by Paul Horwich (1987). Here is, roughly, her variant. She starts with this proposition: if people were to frequently travel to the past, many time travelers would try to change the past by doing things that are easy. This proposition is supported by trivial empirical facts, including facts about human psychology and facts about how easy it is to accomplish certain tasks. Some time travelers might try to wish their siblings Happy Birthday on days their siblings never hear from them. Wishing a sibling Happy Birthday is easy. Other time travelers might seek out desirable real estate they have never purchased and then go in the past to try to buy it at cheap prices. Buying cheap real estate is easy. Others might try to save a pedestrian who dies in a car accident by going back in time and pushing the victim out of harm’s way. Pushing someone out of harm’s way (especially when you know a car is coming) is easy.

Any attempt to change the past, however, will fail. Doing something in the past that didn’t happen is contradictory. It doesn’t matter how easy it is to wish your sister Happy Birthday, buy cheap real estate, or push someone out of harm’s way. If something didn’t happen, then it didn’t happen. If you try to wish your sister Happy Birthday on a day she doesn’t hear from you, you will fail. If you try to buy cheap real estate when you didn’t buy it you will fail. If you try to prevent someone from dying when they died, you will fail. Time travelers who try to change the past by doing easy tasks will inevitably fail at easy tasks.

Given ordinary facts about human psychology, then, there are two main possibilities: (a) there will be frequent time travel to the past, and many time travelers will try to change the past by trying to do easy tasks; all of those who do so will fail, or (b) there will not be frequent time travel to the past. Elliott recognizes that the first possibility is comparatively unlikely. It is unlikely that there will be a string of consequences preventing all the relevant time travelers from accomplishing easy tasks. It is comparatively likely that, instead, there will be no time travel to the past. We have a reason—albeit a defeasible one—to think that there will not be frequent time travel to the past. Note that Elliot’s reasoning does not rely on any claims about whether determinism is true or false. Nor does it rely on any claims about the nature of the laws. Relying only on basic claims about human psychology and about how easy it is do certain tasks, she concludes that it is unlikely that infallible predictors will share many predictions with people about their behavior.

Elliott’s argument tells us nothing about what prevents there from being frequent time travel to the past. The explanation might be that time travel is too expensive (as Horwich thought for independent reasons), or that nobody figures out how to build a time machine, or that it’s physically impossible to time travel, or that some institution bans time travel, etc. Elliott’s reasoning tells us only that frequent time travel to the past is unlikely without telling us why.

**Infallible Predictors**

We will now highlight two ways time machines and infallible predictors are analogous. First, we should not infer from the story of Frank that time travel to the past is metaphysically impossible. There is nothing contradictory about Frank traveling to the past, trying to wish his sister Happy Birthday by trying to do easy tasks, and then failing. Moreover, there is nothing contradictory about many time travelers trying to change the past by doing easy tasks and all of them failing. Analogously, *pace* Ismael, we should not infer from our story of the oracle in the coffee shop that infallible predictors who share predictions about many people’s behavior are metaphysically impossible. There is nothing contradictory about an infallible predictor predicting you will order coffee in the next five minutes and then you trying but failing to prove the predictor wrong. That is, there is nothing contradictory about you trying really hard to not order coffee in the next five minutes but still somehow ordering coffee in the next five minutes. Moreover, there is nothing contradictory about many people trying to refute an infallible predictor by trying to do easy tasks and all of them failing. The story of Frank does not show that time travel is impossible; the story of the oracle in the coffee shop does not show that infallible predictors who share predictions with many people about their behavior are impossible.

Second, just as the case of Frank reveals that frequent time travel to the past is unlikely, the case of the oracle in the coffee shop reveals that an infallible predictor sharing predictions with many people about their behavior is unlikely. Here is why. We start with this proposition: if there were an infallible predictor that shared predictions with many people about their behavior, then many people would try to falsify some of the predictor’s predictions by trying to do things that are easy. This proposition is supported by trivial empirical facts, including facts about human psychology and facts about how easy it is to accomplish certain tasks. Some people might try not to order coffee when an infallible predictor predicts they will. Not ordering coffee is easy. Some people might try to say “hi” when an infallible predictor predicts they will say “hello” instead. Saying “hi” instead of “hello” is easy. Some people might try to jump when the predictor says they will sit still. Jumping is easy.

It is human nature to try to falsify predictions of any purportedly infallible predictor. Many people would want to show that they have free will—or that they are unpredictable creatures, more nuanced than any predictor or algorithm can imagine. If nothing else, many people would try to falsify the predictions of a purportedly infallible predictor out of mere curiosity: a desire to see what will happen when they try to refute the predictor.

There are exceptions. Imagine a demon who makes predictions about people’s behavior and threatens them not to even try to falsify those predictions. If such a demon predicted you would do something, you might think twice before thumbing your nose at the demon. Another exception, of course, would arise if a predictor predicts that you will do something that is very beneficial to you, like win the lottery or not injure yourself. For these sorts of predictions there is naturally no desire to thumb one’s nose at the predictor. The current argument, then, is restricted to predictions about mundane actions, like ordering coffee and saying “hello”, when there is no threat from the predictor.

Any attempt to falsify the prediction of an infallible predictor will fail. Infallible predictors make only correct predictions. If an infallible predictor predicts you will do something then you will do it. Not doing something you do is contradictory. It doesn’t matter how easy not ordering coffee is or how easy saying “hi” is. If you try not to order a cup of coffee when you do order a cup of coffee you will fail. If you try to say “hi” when you instead say “hello” you will fail. People who try to falsify predictions made by an infallible predictor about their behavior will inevitably fail at easy tasks.

Given ordinary facts about human psychology, then, there are two main possibilities: (a) an infallible predictor will share predictions with many people about their behavior, many of those people will try to falsify the predictor’s prediction by trying to do easy tasks; all of them will fail, or (b) there will not be an infallible predictor that shares predictions with many people about their behavior. The first possibility is comparatively unlikely. It is unlikely that there will be a string of consequences preventing all the relevant wannabe falsifiers from accomplishing easy tasks. It is comparatively likely that, instead, that there will not be an infallible predictor that shares predictions with many people about their behavior. We have a reason—albeit a defeasible one—to think that there will not be an infallible predictor that shares predictions with many people about their behavior. Note that our reasoning does not rely on any claims about whether determinism is true or false. Nor does it rely on any claims about the nature of the laws. Relying only on basic claims about human psychology and about how easy it is do certain tasks, we conclude that it is unlikely that infallible predictors will share many predictions with people about their behavior.

Just as Elliott does not tell us why time travel is unlikely—only that it is unlikely—our reasoning is silent about why it is unlikely infallible predictors will share many predictions with people about their behavior. It might be that it is too expensive to build a machine that infallibly predicts human behavior. It might be that there are infallible predictors, but they would rather keep their predictions to themselves. It might be that a government body will ban the existence of infallible predictors. Our reasoning tells us only that it is unlikely that infallible predictors will share many predictions with people about their behavior, without telling us why.

References

Cuypers, Stefaan E., and Stefan Rummens. (2009). “Determinism and the Paradox of Predictability.” *Erkenntnis* 72, 2: 233-249.

Elliott, Katrina. “How to Know that Time Travel is Unlikely Without Knowing Why.” *Pacific*

*Philosophical Quarterly*.

Horwich, Paul. (1987). *Asymmetries in Time: Problems in the Philosophy of Science.* Cambridge: MIT Press.

Ismael, J. T. (2016). *How Physics Makes Us Free*.Oxford.

Lewis, David. (1976.) “The Paradoxes of Time Travel.” *American Philosophical Quarterly* 13: 2. 145-152

Scriven, Michael. (1965). “An essential unpredictability in human behaviour.” In B. B. Wolman and E. Nagel (Eds.), *Scientific psychology: Principles and approaches* (411-425). New York: Basic Books.

1. Note that Ismael does not conclude that it is impossible for there to be an infallible predictor *tout court*. She remains neutral about whether it is possible for there to exist outside of our universe a being who makes infallible predictions about us and doesn’t share any of those predictions with us. She thinks merely that the paradox of predictability shows that it is impossible for there to be an infallible predictor who exists inside of our universe and shares with us their predictions about us. [↑](#footnote-ref-1)
2. We take no stand here on Ismael’s conditional claim that *if* it is impossible for there to be an infallible predictor of the sort in question, then people have a certain kind of freedom that is compatible with determinism. [↑](#footnote-ref-2)
3. Although of course it’s not true in relation to any fact that you both can and cannot ski. [↑](#footnote-ref-3)