

Is Free Will Compatible with Lewis's Account of Time Travel?

Jaime Conchillo¹ and Kyle Van Oosterum[#]

¹Colegio Aldoveo, Alcobendas, Spain

[#]Advisor

ABSTRACT

Time travel may be possible but whether time travellers are free to do what they want isn't so clear. It is generally thought that free will isn't compatible with the possibility of backwards time travel because of the things that time travellers can't do. Since they can't create logical contradictions, it seems they can't do anything to change the past. What does this say about a time traveller's free will and free will more generally? This paper aims to look through the different views of philosophers to come to a conclusion. To prove that free will is compatible with Lewis's account of time travel we will go through incompatibilism and compatibilism, which hold different views about whether causally determined things (like the past) can be compatible with the existence of free will. I will show why a compatibilist view is defensible when talking about free will in the context of time travel. Through this paper, I will start by looking at some models of time travel. Then I will examine how the logic of time travel commits us to an unsatisfying view of free will. Afterward, I will discuss distinct models of free will and how those may help us understand the problem and then conclude with discussion of a common objection to Lewis's account of time travel and free will.

Introduction

The common way of looking at free will is to think that it consists in the ability of doing otherwise or to decide what we do, but this conception is not true. Free will has different interpretations for each person but in this paper, I will take "free will" to mean "the ability and possibility to act at one's discretion and with one's own intentions" and I will go more deeply in this topic in the next paragraph. Lewis's conception of time travel is that it is logically possible because all the logical contradictions that it creates are solved because you don't have the possibility of doing them. Unsurprisingly, we can't change the past. The problems that come up with free will in time travel are the lack of capacity to do most actions we usually do. This is what the view of incompatibilism argues, namely, our free will is not compatible with anything that is causally determined such as the fact that the past cannot be changed. It is interesting to see how these familiar topics about free will apply to time travel and give us a totally different perspective on the issue. We need to understand that there is a difference between possibility and ability. Possibility is something that may happen or be the case. Ability is having the skill to do something. To be free you don't only need the possibility but also the ability. So, for example, in this paper we assume that we can say that someone has the possibility of being free even though he doesn't have the ability of doing something different to what he does. This assumption comes from the consideration that even if determinism is true and your actions are consequences of the past and the laws of nature you are still free. This kind of philosophical view is called compatibilism, which holds that free will (depending on the definition one adopts) is compatible with causal determinism.

It is famously defended by Frankfurt (1969) who puts an example of two assassins: Imagine an assassin has to kill somebody with someone else present. He is afraid of his partner freaking out and leaving him alone in the moment of the crime, so he decides to hypnotize his partner to make sure that instead of freaking out he will still help the assassin with the crime. In the moment of the assassination, everything

goes as planned, and they kill the victim. Was the partner really free? If we think about it, he may be because if he wanted to kill the victim, he would act that way without needing to be hypnotized. But the other possibility is that he freaked out and didn't want to kill the victim. In that case he would have done that even though he didn't want to. In the first case we can say that he is free despite the fact that he couldn't do anything otherwise. But some people may argue that he wasn't free because he couldn't make a different choice even if he didn't want to kill the victim. As Frankfurt says: "A person may do something in circumstances that leave him no alternative to doing it, without these circumstances actually moving him or leading him to do it -without them playing any role, indeed, in bringing it about that he does what he does." (Frankfurt, 1969: 830). So, if we believe what Frankfurt says it is easy to say that if he chose to kill the victim then he was making a free decision and if he didn't choose it, then he did not act freely. So, if one agrees with this reasoning, then one subscribes to compatibilist reasoning which holds free will and determinism are compatible meaning that you are free but can't do otherwise.

Turning now to time travel, we need to understand that this paper assumes that time travel is logically possible which Lewis (1976) defended. When Lewis talks about the possibility of time travel, he speaks about why it is possible to do it. What Lewis is trying to reject is the following argument:

1. It is not possible to create logical contradictions.
2. It is impossible to travel in time without producing logical contradictions.
3. Therefore, it is not possible to travel in time.

This argument is valid, but Lewis says it is not sound because there exists the possibility of traveling in time without making a logical contradiction. For Lewis, the reason we can't make logical contradictions is that there is always something that stops you from making them. What that "something" is will be explored in the last section of this paper.

Lewis decides to challenge the above argument by saying that you can travel in time without making logical contradictions. So even if I have the possibility of killing my grandfather, I won't be able to do it because if I kill my grandfather I wouldn't be born and therefore, I won't be able to go back in time and kill my grandfather. If I don't kill him, I will be born to kill him, but the reason I can't kill my grandfather is that I would make a logical contradiction. To understand this Lewis makes a distinction with what is possible and what is *compossible*. The concept of compossibility allows us to distinguish the different senses of what we "can" do relative to a set of circumstances or facts. "To say that something can happen means that it's happening is compossible with certain facts. Which facts? That is determined, but sometimes not determined well enough, by context. An ape can't speak a human language? Say, Finnish? But I can. Facts about the anatomy and operation of the ape's larynx and nervous system are not compossible with his speaking Finnish. The corresponding facts about my larynx and nervous system are compossible with my speaking Finnish. But don't take me along to Helsinki as your interpreter: I can't speak Finnish." (Lewis, 1976: 150). So, it isn't possible to make a logical contradiction and not compossible because you can't do it. In time travel we can say it is compossible with a narrower set of facts to kill your grandfather perhaps because you're a good shot, you have a gun, you're aiming the gun at him. But it is not compossible with a broader set of facts (i.e. I would not be born if my grandfather were killed before he had my father) to do things like killing your grandfather in the past because that creates logical contradictions.

In this paper, we aim to explore and reconcile free will with Lewis' conception of time travel. In order to do so, the remainder of the paper is structured as follows. In section 1, we discuss models of time. In section 2, we will talk about the limitations of free will in time travel. In section 3, we argue the possibilities of there being free will in time travel. Last of all, in section 4, we discuss some worries for incompatibilism.

Now that we understand free will and Lewis's conception of time travel we can discuss what the implications of time travel on free will are. We are assuming time travel is possible, but it seems time travellers are limited and can't do the same things they normally do. I would like to argue why I believe that free will is compatible with time travel but is limited by the sort of events that make logical contradictions.

Section 1: Models of time

There are three ways of looking at time. These three models are discussed by Tognazzini (Tognazzini, 2017). The first model is the single timeline model which consists in looking at the four-dimensional time box from left to right and seeing there is a single united timeline so when I travel back in time, I am part of the past. Imagine a guy called Jones decides to travel back in time and say 'hi' to his mother when she was young. If he does it and everything continues, his mother will grow up to give birth to him, and he will grow up to go back in time, but that had already happened in the past, so he has inevitably two pasts: his future which forms part of the past and his personal past. The distinction between personal past and external past is that personal past is the past of someone (Lewis 1976). So, for example, his personal past may be in the external future because he is a time traveller. On the other hand, the external future in the single timeline model is linear because it is always the same. For example, a time traveller from the future may go to the external past, but it is his personal future.

The second model is the branching timeline model. In this model, when something happens again it is like a different moment. For example the first timeline is the one in which Jones is not able to say hi to his mother and in the branch, Jones goes back and says hi so we can say this is not really time travel because he is not actually going to the past, he is creating another branch in time.

The last model is called the hypertime model, which entails another dimension of time in which apart from the original moments on the four-dimensional box there is also hypermoments in this second dimension of time. The hypertime means that the four-dimensional box is constant change and if you look from outside the box you can only see what the timeline looks like in that exact hypermoment. So, time is in constant change. In one hypermoment we would see that Jones didn't say hi to his mother, and in the next one we would see it has happened.

In this paper we are going to be looking at the first model because the implications are the same as in the other two but it is a simpler way of looking at it.

Section 2: Free will in time travel

When we talk about free will in time travel, there are two problems that arise. The first one consists in the impossibility of doing certain kinds of acts. The following argument shows precisely why with a famous example:

1. It is not possible to make a logical contradiction.
2. Jones killing his grandfather in the past would create a logical contradiction.
3. Hence, it's impossible for Jones to kill his grandfather.

This argument here is valid, and we can assume it is also sound. As I said before if Jones kills his grandfather, he wouldn't have been born in the first place, so he wouldn't be able to kill his grandfather, yet he would have been born wanting to kill his grandfather. We can then assume that what we are usually able to do is restricted by time travel in different ways. This raises interesting questions about free will which people believe necessarily has to do with what we are able to do.

So, it is impossible not to conclude that free will is limited in time travel because we would be in a kind of bubble where we can't do anything that changes the past. But some people will argue that not being able to do something doesn't mean that you are not free. For example, you are not able to fly, and we can still reasonably say that you are free. The only difference here is that you're limited in all your actions and that it is not compossible with a set of facts about the logic of time travel that you are able to do much, just as Gorillas can't speak Finnish. However, if we maintain our initial description of "free will" we can't say that we are not less free with this conception of time travel, but we might still have the possibility to do as we want. So maybe free will is just limited, after all. This means I can still do what I want even though there aren't many things I can do freely or the ones that I can freely do are limited. A critic of this view might say, "well this really doesn't look like free will, or the kind of free will we want", but to them we'd simply reply: "this might be the only kind of free will we have." In any case, it seems that it would be wrong to insist otherwise.

The second problem that arises is a more complex one and requires further background explanation so I will put forth a Tognazzini-like example to illustrate the problem.

Imagine Jones has a friend called Kim with whom he can travel back in time. Jones is with Kim. They're walking down the street when someone tries to steal from them and is going to shoot Jones. Right before he shoots, someone in an alley kills the robber. So, Jones gets the gun off the robber and decides to travel back in time with Kim. They go back a few minutes ago and appear in an alley where they see Jones of the past who is going to be shot by the robber. Jones of the present decides to shoot the robber and kill him. So, if we order the set of events, we get the following:

1. At T1, Jones is threatened by the robber.
2. At T2, the robber is killed by Jones from T3
3. At T3, Jones goes back in time and kills the robber who is going to kill the one from T2.

This is a simple time travel story; Jones saves himself without him knowing it was him. Then he finds out the truth and saves himself. The problem that arises here is whether the action of saving himself was free or not. Due to the curious situation, if he didn't kill the robber he wouldn't have survived and hence wouldn't be alive even to make the choice of saving himself and traveling back in time.

Of course, this discussion is only possible when assuming that time travel is possible, so the discussion is not only about free will in time travel but also about the consequences for free will due to time travel. (Tognazzi, 2017)

Depending on the conception of free will we defend, we could ask whether Jones could have done otherwise or whether he could have acted according to his own intentions. The two moments in which we can discuss whether Jones had the possibility of doing otherwise is in T2 and T3. We could ask the following:

1. At T2 could Jones not saved himself?
2. At T3 could Jones not travelled back in time?

In T2 we can see that Jones couldn't have saved himself because that would have caused a logical contradiction because if Jones refrained from saving himself then he wouldn't be alive to make that decision, and we know that is not possible. Therefore, it is necessary for Jones to save himself. At T3 we have a similar problem if Jones doesn't travel back in time, so Jones really has no other choice than going back in time. So, Jones's free will looks to be in danger because he can't do otherwise than travelling back in time.

Section 3: Hope for free will.

We mentioned the main problems that free will faces when we introduce the possibility of time travel. So, the first problem of limited actions in the past can be solved by arguing that the impossibility of doing some things doesn't mean we are not free. Some may argue that we can't fly and that doesn't mean our free will is in risk, but in this case, we have a much wider number of restrictions that make it impossible to do pretty much anything. So, if we value the possibility of our actions, we can say we are not free. A more visual example would be a man who has a car accident and is totally immobilized forever. The only thing he can do is see, but he can neither talk nor listen. Would you say that this man is free? Most people would say he is free even though he can't move. We could say that he is mentally "free" and can think as he wants. So, here we are making a distinction between freedom of actions and freedom of intentions or ideas. If we consider that freedom of intentions exists then we can see even though you can't do anything you still have your mind and that's enough to be free. You can argue that causal determinism does affect the mind too, but we can assume that our way of thinking is more than just our neurons and, therefore, we can still have freedom in our minds and intentions (of course, not everyone will accept this assumption. Unfortunately, I do not have the space here to defend it but simply note the philosophical benefits of accepting this assumption).

A second distinction we can add to the discussion to defend the compatibility of free will with time travel is to claim that one can be free but not totally free; there might be different levels of freedom as opposed to either being free or not. So, we could talk about a man in a prison. He still has a level of free will which is not the same as the liberty that another individual has. A case in which we can see there is no free will, as we define it, is in *1984* by George Orwell. In this book, we can see a dystopic future in which

people can't even think differently to the way they are told to. So, this would mean that they lack most of the free will of actions and ideas.

If we suppose this argumentation is true then the first problem that free will faces about the incapacity of doing anything in time travel is solved but what about the second one?

In the second case, we see a causal loop in which a lot of issues are happening at the same time. We need to introduce the terms of soft facts and hard facts. So, a hard fact is a consequence of the past, so it is impossible to change it because causes have their consequences. On the other hand, soft facts are affected by what a person does in the future and therefore can be changed because they aren't a consequence of the past. Let's put another Tognazzini example:

"Suppose it was true a million years ago that I will have a cup of coffee tomorrow morning. If I'm nevertheless able to refrain, then it must be possible for me to refrain. What would the world be like if I were to refrain? (...) The first question seems like it has a relatively straightforward answer: if I were to refrain from having that cup of coffee, then something that was in fact true would be false. That thought is perhaps initially worrying, but once we make the distinction between hard and soft facts, the worry dissolves. If I were to refrain from having that cup of coffee, only some soft fact about the past would be false." (Tognazzini, 2017; 686).

If we put the case of Jones into this logic, it seems more worrisome. This is because if Jones refrains from going back in time in T3 then he would have died in T2, and therefore, he couldn't have refrained from going back in time in T3. The problem here is that Jones depends on the action he is going to do in his personal future but is in the external past. If we think about it, Jones even has two pasts in the same moment. The first one is Jones of T2 and the other one is Jones from T3 which has travelled back in time.

The main point of the question is seen in the Principle of the Fixity of the Hard Past (Tognazzini, 2017). This principle states that:

If, in order for S to do A, some hard fact about the past would have to be false, then S is not able to do A.

We can't change hard facts, so we can't change the past. Tognazzini points out that the Principle of the Fixity of the Hard Past does not account for soft facts. He then discusses two variations on that principle, where it is difficult to pinpoint what distinction or argument to show those principles are defective and therefore create room for a plausible compatibility of free will with time travel. These principles state that:

(KV) S is able to do A only if, had S tried to do A, S would or at least might have succeeded.

(JS) S is able to refrain from doing A only if, had S not done A, S would have done something else instead. (Neil Tognazzini, 2017: 13).

Both these principles look valid and sound, so we need to argue why we can believe in some definition of free will even though right now some set of circumstances may seem fixed or determined. (JS) gives two possibilities of doing otherwise. The first one is that S didn't do A, so we would be in the same position, in which time travel is not compatible with free will because if Jones didn't do A, then Jones wouldn't be able to do otherwise. The second possibility is that S doesn't exist to do A but this means that S couldn't have done otherwise or done anything at all. This is a hard challenge to the compatibilist view we are defending because it becomes more unclear what kind of free will, if any, the time traveler possesses.

Section 4: A worry for incompatibilism

Lewis talks about some commonplace reasons that stop us from making logical contradictions when we travel to the past, but if everything we've said above is true it seems like incompatibilism about free will and time travel is the right position to hold. However, this incompatibilist view takes us to a problem not previously mentioned which was put very well by Sider (2002: 122): "Once the inability of the time traveler to kill her former self is admitted, one wonders what prevents her from doing so". Sider went so far as to say there are some seemingly "strange shackles" on the time traveller that the incompatibilist needs to explain. David Lewis mentions the impossibility of killing yourself, for example, but doesn't say what is

what stops us from doing it. Smith's (2017) response to Sider is an argument which says maybe there has never been the possibility of doing that, it is just our perception which scams us into thinking that it is possible to kill ourselves and do logically contradictory things in the past. Smith (2017) brings up an interesting example that illustrates this trick on our reasoning that makes us think we need an explanation for this "strange shackles" worry. This is the following scenario:

"You have become accustomed to having an attic room and have completely forgotten that your house lacks one... Yet – from your point of view – whenever you try to enter it, something always stops you: either something distracts you from your attempt (e.g., the telephone rings or someone knocks at the front door), or you slip on a banana peel. Let's suppose you find this very perplexing and mysterious. When you try to enter your kitchen, bathroom, or any other room, there is no problem – and yet when you try to enter the attic, you always fail. We – considering the case from the outside – can furthermore see that you will invariably fail: you will never succeed in entering the attic." This allows Smith to give another solution to the problem of the "commonplace reasons" which Lewis had been too unclear about.

However, we might think the time travel situation just is different. What Tognazzini says about these "strange shackles" and why they are so "strange" is that the agent is their own shackles, yet we are claiming they are not free. Even though every action depends on Jones and is seemingly up to him, the incompatibilist wants to claim that Jones is not free, which may sound odd. It is in this sense that the shackles are "strange" and that the attic example really doesn't seem analogous to the case of Jones. Perhaps this moves the burden onto the incompatibilist to explain why we cannot say this is free will.

If we look at Jones' case, he depends on his own past and that entails he needs to go back in time to T3, so he doesn't depend on anyone else besides himself. So, this gives us a different conception of free will, one in which the agent doesn't depend on anyone besides his personal thoughts and mental state despite not being able to do otherwise.

I believe that the "strange shackles" are just an impossibility that occur because of the incapacity of making logical contradictions. Therefore, just as you can't fly you can't kill your grandfather in the past. This conception is assuming that we are just not able to do that, but this does not mean that we are not able to do many other things. As Tognazzini says, we are the reason these strange shackles occur. Incompatibilists will say that this means there is no free will at all but if we follow our conception of free will we can say that in both cases, the Jones and time travel ones, we are free. It is really important to realize that these problems depend on what we mean by "free will" and that's why my view which builds on Tognazzini's is one opportunity to plausibly claim free will and backwards time travel are compatible.

Conclusion

In conclusion, I have argued that free will is compatible with Lewis's conception of time travel. To demonstrate that free will doesn't require the ability of doing otherwise and therefore free will is compatible with time travel I would like to recall the case explicated in the introduction.

As I said at the start of this paper, free will can be thought of as the ability of acting at one's discretion or intentions. We know that in the case discussed above this ability is possible as the assassin's partner had the possibility of acting as he wanted (in at least one of the cases). The main problem we face is the wrong interpretation of what free will is. If we make the distinction between free will of intentions and free will of actions, we can say that everyone has the opportunity of thinking or trying to do what they want even if they can't do it later. It is also important to bear in mind the level of free will we have, depending on how much we can act at our own discretion. If we were to neglect talk of levels of free will then we would argue that no one is free because everyone has different limitations whether it is laws or physical inabilities which stop us from doing as we want. But, as I have tried to show earlier, this is not a necessarily plausible conclusion to hold. We can assume that you can be more or less free depending on the ability of doing as you want, but we should obviously accept no one is totally free. However, this doesn't mean we are not free at all. So, following this reasoning we can say that free will doesn't require the ability of doing otherwise because no one will have the ability of doing otherwise *completely*, and it could still make sense to say we are free. If we connect this to time travel, we can assume that there is a lower level

of free will than in our daily life because in time travel, we can't do what we can't do in our daily life plus some other actions that create logical contradictions. Therefore, if Lewis' account is right, we shouldn't be necessarily worried that we have no free will. We should just think that we are limited in some ways, but that does not mean we cannot meaningfully say we are "free".

Acknowledgments

I would like to thank my advisor for helping me with this project.

References

- Frankfurt, H. (1969). Alternate possibilities and moral responsibility, *The Journal of Philosophy*, Vol. 66, No. 23: 829-839.
- Lewis, D. (1976). The paradoxes of time travel, *American Philosophical Quarterly*, Vol. 13, No. 2: 145-152.
- Sider, T. (2002). Time Travel, Coincidences, and Counterfactuals. *Philosophical Studies* Vol. 110: 115-38.
- Smith, N.J.J. (2017) I'd do anything to change the past, but I can't "do that". *American Philosophical Quarterly*, Vol. 54, No. 2: 153-168.
- Tognazzini (2017) Free will and Time Travel. In Griffith, Levy, and Timpe (eds.), *The Routledge Companion to Free Will*, (pp. 680-691). Routledge.