

Why Would Time Travellers Try to Kill their Younger Selves?

Nicholas J.J. Smith<sup>1</sup>

In this note I raise a new problem for backwards time travel, and make some first suggestions as to how it might be solved. I call it the *motivation problem*. It is not a logical or a metaphysical problem, but a psychological one. It does not impact upon the possibility, or even the likelihood, of backwards time travel. Yet it is deeply puzzling, and we will have no idea what time travel would actually be *like* until we explore it. Thus, where other problems for backward time travel assume that we know what time travel would be like, and argue that we cannot have it, this new problem gives us no reason to think that we cannot have time travel, but argues that we have much less idea than we usually suppose about what it would really be like to travel back in time.

Consider the auto-infanticidal time traveller in the act of attempting to kill her younger self. What is going through her mind? What is her motivation? Simple, you might say: whatever motivation anyone ever has for committing suicide—except that in the time traveller's case, she thinks she can get a head-start on ending her life: she can end it years ago. Perhaps she wishes she had never been born: her life was so awful, she wishes to end it before it progressed very far. But things are not so simple at all. The problem is that the older time traveller *knows exactly what she will do next*: for she remembers seeing it happen (from the perspective of the younger self) when she was young! Suppose she is pointing a gun at her younger self. She will remember this very scene (it is not the sort of thing one would be likely to forget): as a child, she was looking down the barrel of the gun; her assailant pulled the trigger and...the gun jammed with a click. She now pulls the trigger and...the gun jams with a click. She remembers what happened next: her assailant banged the gun on the edge of the table and raised it again. She now bangs the gun on the edge of the table and raises it again. She remembers what happened next: there was a power cut and the lights went out. The lights now go out. She remembers what happened next: she crawled under the table and hid, but then the lights came back on and her assailant looked under the table and pointed the gun at her again. She now looks under the table and points the gun at her younger self again. (She always wondered how her assailant knew exactly where to look, as soon as the lights came back on.) She remembers what happened next: a Mr Whippy truck drove by and her assailant was momentarily distracted by the wavering strains of Greensleaves, and she managed to grab the gun. Just then a Mr Whippy truck drives by and she is momentarily distracted by the wavering strains of Greensleaves, and her younger self grabs her gun. She remembers what happened next: she fired the gun, and shot her assailant in the shoulder, but then dropped the gun in shock, and the assailant picked it up again with her good hand. She now hears a bang and realizes she has been shot in the shoulder, but then notices her younger self has dropped the gun, and picks it up with her good hand. And so on...The time traveller thus knows exactly what she is going to do, *before she does it*. This is extremely odd!

---

<sup>1</sup> Department of Philosophy, Main Quadrangle A14, University of Sydney, NSW 2006, Australia.  
<njjsmith@mail.usyd.edu.au>

The problem is not limited to cases of attempted auto-infanticide. The time traveller may just go back and have a pleasant conversation with her younger self. She might tell her younger self about the future and about what it feels like to travel in time. But the problem is, whatever she says will sound oddly familiar: for she will remember hearing these words when she was young. Imagine having a conversation with someone, and suddenly becoming aware that you knew exactly what you were about to say. Not so odd, you might think: actors must feel like this, for example. But the situation is not at all like that of an actor. An actor always has the option of saying something else, or she can fall silent, or leave the stage: sure, she will ruin the show, but she *can* do these things—she *can* depart from the script—*if she really wants to*. But if the time traveller is about to fall silent, she remembers that too; if she is about to run screaming from the room because she is totally freaking out, she remembers that, too. There is no escape: *whatever* she does next, she knows what she will do, before she does it.

OK, so we have a weird situation here. What exactly is the problem, and how much of a problem is it? The first thing to note is that the problem is *not* that the time traveller lacks free will, lacks control over her actions, or is fated to act in particular ways. To see this, imagine that you are sitting watching someone, and you know exactly what he will do next. You know this not because you have any control over what he will do, but because you have seen the future. (You have simply *seen* it, *as it happens*—you are just an observer. Perhaps a time traveller brought back a video tape of the person you are watching, and you have seen the tape.) Suppose the person you are watching is deciding what to eat for lunch. You know that he ends up choosing baked beans on toast. Does this mean his choice was fated, or was not free? Not at all. You simply knew the outcome of his free choice, before he made it. The fact that you know now what you chose for breakfast this morning does not mean you did not choose freely; no more does the fact that you know now what your friend *will* choose mean that his choice will not be a free one. He certainly *could* choose otherwise: and if he was going to, something else would be on the tape you saw.

The situation is the same, as far as freedom is concerned, with the time traveller: the fact that she knows what she will do before she does it does not in any way mean that she is not free to choose otherwise. There is a difference between the two cases, however. The difference is one of psychological mechanisms. Your knowledge of what your friend will choose for lunch does not figure in the causal processes leading up to that choice. In the time traveller's case, however, the knowledge of the outcome is not isolated from the process leading to that outcome: what one knows plays a role in one's decisions about what to do next. There is thus feedback between the outcome—the time traveller's decision about what to do next—and the causal processes that produce it.<sup>2</sup> This is what makes the time travel case particularly odd. But the oddity has nothing to do with fatalism, or lack of free will on the time traveller's part.

---

<sup>2</sup> For example, knowing that the bottom step is loose—because you have seen someone else step on it and have a nasty fall—you (a non-time traveller) try not to step on it now. The time traveller, knowing that the step is loose—because she remembers, as a child, seeing her older self step on it and fall—also tries not to step on it now. Unlike you, the time traveller thereby tries not to do the very thing whose memory causes her to act as she does now.

The second point to note is that this problem does not constitute an argument against the possibility or even the likelihood of backwards time travel. It is not a logical or a metaphysical problem; it is a *psychological* problem. Of course we could stop right there, and say that that is our answer: the problem is merely psychological; all the problem shows is that time travel would be weird, and who doubted that? But it would be much more satisfying if we could say a little more than this. The challenge is to tell time travel stories in which time travellers interact with their younger selves—and thus remember what they are going to do, as or before they do it—and which are psychologically possible, which *get the motivation right*. Without such stories, we really have no idea what auto-infanticide attempts—or more benign attempts at self-interaction—would be like. Despite the vast number of pages that have been devoted to discussions of auto-infanticide, we actually have no idea at all why time travellers would try to kill their younger selves: that is, we have no idea what could be going through their heads as they make their attempts.<sup>3</sup>

A third point is that of course time travellers can *avoid* the whole problem by not interacting with their younger selves, or by only interacting with them as infants, when their younger selves are too young to remember the encounter, or by getting amnesia after the encounter,<sup>4</sup> or by misremembering or only partially remembering the encounter (as we sometimes do forget or misremember things, even when we do not have amnesia), or by disguising their identities so that their younger selves do not realize who they are (although this would be of limited use, as in a long encounter, it is quite likely that the older self would eventually realize “Hey, *I* am that strange person who visited me when I was young, and *this* is that visit!”), or by keeping a distance, so that the younger selves cannot see the details of what the older selves are doing (the older self might, for example, fire a silenced gun at her younger self from the top of a tall building, or might discretely brush past the younger self and slip a tasteless poison into her drink), and so on. However this point does not in any way diminish the interest of the question of what would motivate a time traveller who *was* interacting with her younger self at an age at which the younger self would remember the encounter, and who did not have amnesia etc.—and *that* is the question in which we are interested. We can *ignore* this question by only considering other sorts of time traveller, but that is very different from *answering* the question, which is what we would like to do. After all, there is every reason to suppose that if time travel were possible, then time travellers would *sometimes* interact with their not-too-young selves, without previously having suffered amnesia, etc.

So what would they be thinking if they did do this? Let us look one by one at various psychologically possible options for self-interaction on the part of time travellers. It

---

<sup>3</sup> To clarify: distinguish what we might call ‘ends-motivation’ from ‘means-motivation’. There is no problem here concerning ends-motivation—that is, concerning why the time traveller might desire that her younger self die young. The time traveller might simply have had a miserable life, and wish that she had never lived it. The problem concerns means-motivation: it concerns what is going through the time traveller’s head as she attempts to kill her younger self. Thus our question is not ‘Why would a time traveller wish her younger self dead?’, but ‘Why would a time traveller try to perform *certain specific actions* (pulling the trigger, throwing the grenade), given that she remembers seeing her older self perform these very actions?’

<sup>4</sup> After, that is, from the point of view of the younger self; before, from the point of view of the older self.

would seem that it is psychologically impossible for someone to both know exactly what she is going to do next, and also deliberate in the normal way about what to do next: this is what sets our problem.<sup>5</sup> Hence it would seem that either the time traveller would not be deliberating in the normal way, or she would not be remembering the encounter in the normal way. In some (not all) of the following scenarios, however, the time traveller does both. All of the following scenarios are weird in various ways; but none of them is psychologically *impossible*. This is what matters when it comes to making progress on the motivation problem. We want to see if there are ways in which time travellers could interact with their younger selves that do not involve a complete breakdown of all known psychological mechanisms. For if there are not, then we really have no idea of what time travel—or at least, time travel involving (amnesia-free etc.) self-interaction—would be like.

(1) Suppose the time traveller clearly remembers being visited by her older self, and clearly remembers everything that happened. She now encounters her younger self, and begins to say and do the very things she remembers so well. One possibility is that she acts as if under duress. Imagine that there is a figure who has, perhaps through bullying, or by means of psychological pressure, attained an ascendancy over you: whatever she tells you to do, you do it, without seriously considering doing otherwise. Now suppose this person maps out your afternoon for you: she says you will go to the bank, produce a gun (which she gives you), threaten the teller with the words ‘do as I say or you won’t live to regret it’, demand \$10,000 in \$100 bills etc. You simply go and do what she has told you to do, without deliberating. Now the time traveller might feel similarly compelled: knowing what she will do, she feels she *has* to do it, and she just does it, without deliberating, in much the way the bullied person robs the bank.

There is, however, an important difference between the two cases. The time travel case as just described is psychologically possible, but metaphysically odd. The time traveller acts as she does because she remembers doing so and feels that she has no choice in the matter. But in fact she does have a choice. She remembers acting this way because...this is the way she acts! Unlike in the bank robbery case, there is no-one compelling the time traveller to act as she does. But from her point of view she *feels* compelled, she feels as though she is acting under duress. So her actions in a way ‘come from nowhere’: they get done, but the time traveller does not *decide* (in the normal way) to do them, and—unlike in the bank robbery case—no-one else decides that they should be done either.

This sort of oddity is familiar in time travel scenarios. Sometimes *information* comes from nowhere: for example, the time traveller who reads a physics book and then goes back and explains relativity theory to the adolescent Einstein. Sometimes *objects* come from nowhere: for example, the time traveller who steals a time machine from the local museum, travels back in time, and donates the time machine to the local museum. In the present case, we have *actions* coming from nowhere, in the sense that no-one decides, in the usual way, to perform them (or decides that they should be performed), and yet they are performed nonetheless.

---

<sup>5</sup> By ‘deliberating in the normal way’ I mean weighing up a range of options, all of which seem open to one, and then choosing which of these actions to perform.

(2) Again, suppose that the time traveller remembers being visited by her older self, and clearly remembers everything that happened. Instead of acting as if under duress, she might convince herself that the things she knows she will do are the very things she would actually like to do. The psychological mechanisms involved here would be similar to those involved when a situation is suggested to one, but one convinces oneself that it is what one really wanted anyway. For example, suppose someone—perhaps a parent, or therapist, or coach, or careers guidance counselor—has achieved a more subtle sort of ascendancy over you than the sort considered above: whenever she suggests that you do something, you do not feel as if you have to do it; rather, you invariably feel that that is the very thing you would like to do, and you do it—exactly as she described—but all the while feeling as if it is your free choice to do so. The time traveller, knowing what she will do next, may similarly feel, not that she has to act in that way, but that—happily—that is the very way she would like to act in any case.

In this scenario, as in case (1) above, the time traveller's actions 'come from nowhere'. She wants to do what she does because her memory tells her that that is what she shall do, and in the way just discussed, she goes along with this: but her memory only tells her that because that is what she will in fact do. Unlike in the parent/therapist/etc. case, in the time travel case there is no second party making the suggestions.

(3) The time traveller might deliberate in the usual way and have all the normal feelings of agency that we usually have when deciding what to do and doing it, *and* she might at the same time clearly remember the encounter in which she is engaged, if she in some way fails to connect her memory of the encounter with her current experience of that encounter. There are several ways in which this might happen.

(3a) One way is that the time traveller might simply not realize that this is the very encounter she remembers so well. You may have been somewhere—say, Napoli Centrale railway station—and remember it quite well; you may then end up there again, but fail to realize where you are (you may for example think you are in Napoli Mergellina station). Even though the station looks just the way you remember it, you simply do not 'click' as to where you are. This sort of thing might happen in the time travel case. The time traveller remembers the encounter with her older self like the back of her hand, and now she freely plays out that scene (this time as the older self), without making the connection and realizing the perfect match between what she is doing and what she remembers. Of course, the longer the encounter, the less plausible this scenario will be. Similarly, if you have extremely detailed memories of Napoli Centrale station—you know the shops and bars, the porters and guards—then it will be quite bizarre (although not impossible) if, when you now see shops and bars, porters and guards which exactly fit your memory, you nevertheless fail to realize that you are once again in Napoli Centrale station.

(3b) A second way is if the time traveller is in *denial*. Suppose that something very traumatic happened to you when you were young: perhaps you were the subject of repeated murder attempts. You might come to believe that your memories of these events are really nightmares, and that the events never really happened. The time traveller, having experienced auto-infanticide attempts in her youth, might similarly

be in denial. Even though she does in fact remember the attempts (from the point of view of the younger self), she might believe that her memories are just nightmares, and thus not be subject to the problem of knowing exactly what she will do next as she attempts to kill her younger self. This is the sort of scenario we find in the film *Twelve Monkeys* (and also in the film which inspired it, *La Jetée*): as a boy, Bruce Willis's character Cole in fact witnessed the death of his time-travelling older self; but he does not realize this, thinking instead that the mental images he has of this event are the products of his dreams. Thus, even when he is doing exactly what the person in his 'dreams' did, he is not subject to the problem of knowing exactly what he will do before he does it: because he does not realize that his 'dreams' are in fact an accurate depiction of the very actions he is currently performing.

(3c) A third way is by having a split personality. One part of the time traveller's mind remembers and another part deliberates—and the deliberating part ends up deciding to do exactly what the remembering part remembers doing.

(4) The time traveller might believe that she will do something different from what she remembers doing: she might believe this because she falls for the *second-time-around fallacy* [Smith 1997] and thinks that she can change the past (when really this is logically impossible); or because she believes in parallel universes, and thinks that she has traveled to such a universe, rather than simply to the past of her own universe. Her attitude to her memories is: that is what I did "the first time around", or in the universe I left; "this time around", or in this new universe, I shall act quite differently.

Note however that while she *tries* to do something different from what she remembers doing, she always ends up doing exactly what she remembers doing; and her repeated failures to do something other than what she remembers doing might start to place some strain on the belief that she really will do otherwise, if the encounter with her younger self is a protracted one.

There are, then, various things which might be going through a time traveller's mind as she encounters her younger self. No doubt there are also other options which I have not considered. Although time travellers involved in such encounters might be guilty of various deviations from perfect rationality, such encounters could at least occur without causing total psychological meltdown on the part of time travellers. Nevertheless, such encounters would *not* be straightforward. None of the above options fits the standard image of auto-infanticide attempts, which is of someone facing up to her younger self armed with a gun and murderous intentions, and squeezing the trigger, only to have her gun jam, and so on and on—but with her murderous intentions remaining intact all the while, through all the foiling coincidences. *This* picture—according to which auto-infanticide attempts are just like any other murder attempts, with the only proviso being that they must fail—cannot be the correct one. The psychology of self-interaction is essentially different from that of interaction with others—because the former, but not the latter, involves the problem of agents knowing what they will decide to do, before they decide to do it.

I have given no reason for thinking that backwards time travel is impossible, or even improbable. Nevertheless, it is very odd. Even the sort of time travel scenario that has received the most attention in the literature—the sort involving attempts at auto-infanticide—turns out to be strange in quite unexpected ways when we examine it

closely. Nothing that I have said implies that time travellers *could not* interact with their younger selves: but such interactions would be quite different from what has usually been imagined.<sup>6</sup>

#### Reference

Nicholas J.J. Smith [1997] Bananas Enough for Time Travel? *British Journal for the Philosophy of Science* 48, 363–89.

---

<sup>6</sup> I am very grateful to Catherine Vidler for many valuable discussions and suggestions. Thanks also to Cody Gilmore and the anonymous referees for helpful comments.