

Intentions and Personal Identity*

[AUTHOR]

12th October 2005

This paper explores the relationship between intentions and personal identity. Specifically I aim show that there are ways in which intentions and personal identity are related that may be problematic for a particular theory of personal identity - for a *neo-Lockean* theory of personal identity. I will do this by presenting two arguments, an argument from ‘belief-dependence’, and an argument from ‘extrinsic causation’. Each aims to show that the neo-Lockean theory is objectionably circular. Whether these arguments in the end succeed in this aim depends in part upon some difficult methodological issues about the permissibility of circularity in philosophical theory, issues that unfortunately lie beyond the scope of the present paper. What I hope will be achieved in the scope of the present paper is at least a fresh sense of the relevance to the personal identity debate of some interesting issues in the Philosophy of Mind.

1 Introduction

What is a neo-Lockean theory of personal identity? A neo-Lockean theory - of the kind defended by Shoemaker, Parfit, Lewis, Perry and others - claims that personal identity over time consists in *psychological continuity*¹. Psychological

*[ACKNOWLEDGMENTS]

¹For some classic statements see (Parfit, 1975) (Lewis, 1976) (Perry, 1976) (Shoemaker, 1984) (Parfit, 1987) (Shoemaker, 1997)

continuity is in turn understood to consist in overlapping chains of direct psychological connections. Direct psychological connections are connections such as the connection, cited by Locke, between an action or experience at one time, and its recall at a later time. But neo-Lockeans generalize Locke's alleged memory-theory of personal identity to include other connections, such as the connections that hold when a belief or a desire is retained, and the connection that holds between an intention and its later execution. The theory is generalized to include a larger range of cross-temporal psychological relations because the theorists see no reason to privilege memory in particular; and indeed, given the plausibility of the idea that a person could survive memory loss, there seems to be a good theoretical reason *not* to privilege memory in particular.

Why think that the persistence of persons is a psychological matter at all? I think that the appeal of the approach has basically two sources: first, the standard notion of a person is that of a self-conscious, psychologically-endowed entity. Thus it can seem to stand to reason that the persistence of such an entity should be in some way a psychological matter. Additionally, the psychological approach accords well with our judgments about a wide range of cases. Particularly significant is the brain-transplant case, which is in effect a contemporary updating of Locke's Prince and the Cobbler story. Most of us are inclined to judge about this case that the person goes with the transplanted brain, and does so precisely because the brain is the seat of psychological continuity.

This is not the place to assess these motives for adopting the neo-Lockean theory. For present purposes it is important to further clarify the *content* of the theory. In particular, I want to ask what is the force of the claim that personal identity over time *consists* in psychological continuity? It should be uncontroversial that psychological continuity is at least in some way generally associated with the persistence of persons: throughout the course of our lives we typically remember our earlier experiences, retain various psychological characteristics, and execute our own earlier plans and intentions. But the neo-Lockean

is making the stronger ‘constitutive’ claim that this psychological continuity is *what it is* for a person to persist. What does that sort of claim involve?

An analogy may help here. In standard observers and in standard conditions, red objects tend to bring about experiences with a certain distinctive phenomenal character. This by itself should be uncontroversial. What is much more controversial is the ‘dispositionalist’ view that an object’s being red *consists* in this tendency to bring about certain experiences. And one reason why this dispositionalist view is controversial is that it can appear to be circular, and in more than one respect. For example, it is arguable that the only way to specify the distinctive phenomenal character in question is by reference to the property that the experience presents, namely the colour red. Additionally it may be that there is no way to specify ‘standard conditions’ except as precisely those revelatory of colours such as red.

These are challenges that the dispositionalist about colour ought to face up to. It cannot be claimed that colours consist in a kind of disposition unless the disposition in question can be shown not simply to presuppose colour. The general idea here - and one that should be familiar, although I won’t have the space to defend it fully - is that it is the mark of a constitutive account of a kind of thing that it offers a *non-circular* account of that kind of thing. If an account of *F*s simply helps itself to the notion of *F*s then while it may be saying something non-trivial about *F*s it will not be telling us what *F*s *are* most fundamentally. It will not be offering a constitutive account.²

²It should be noted that the *necessity* of an account does not by itself secure a constitutive account. To adapt an example from (Fine, 1994): it is necessarily true that *x* is Socrates if and only if *x* is a member of the singleton set (Socrates). But surely this does not in any way define the nature or essence of Socrates. I conjecture that this is so in part because it would be a circular definition of Socrates: the singleton set (Socrates) presupposes Socrates. Another example: If justification is simply defined to be whatever is needed in addition to true belief to get knowledge, then it will be a necessary truth that knowledge is justified true belief. But this will not thereby explain what knowledge *is*, since the justification component

For the purposes of this paper I will make the plausible assumption that the neo-Lockean account aims to tell us what personal identity consists in, and is therefore subject to a similar non-circularity constraint: Psychological continuity should not presuppose personal identity. As noted above, psychological continuity involves as a component the direct psychological connection between intention and action. Therefore it is a condition on the neo-Lockean position that the direct psychological connection between intention and action does not presuppose personal identity. The rest of this paper will be asking whether this condition can be satisfied.

A useful way into the issues is a comparison of intentions with memory. Joseph Butler objected to Locke's theory that memory presupposes and so cannot constitute personal identity. In the next section 2. I will distinguish two things that Butler may have had in mind here: *causal-explanatory presupposition*; and *representational presupposition*. It is *prima facie* plausible that intention no less than memory involves both these kinds of presupposition. The aim of section 3. is to assess the charge that intention representationally presupposes personal identity. I will show that the standard way of defusing the charge for the case of memory is difficult to carry over to the case of intention. I will explain why the *belief-dependence* of intention is the distinguishing factor.

In section 4. I go on to discuss the charge of causal-explanatory presupposition. The possibility of 'fission' cases provides a powerful response to this charge in the case of memory. Again, I will suggest that the causal link between intention and action is more complex, and that there is a way in which personal identity can be presupposed as a factor in this link, unparalled in the memory case. I sketch an argument to this effect, appealing to some broadly 'externalist' ideas about mental causation.

has been defined in terms of knowledge. The account is circular. See (Williamson, 2000).

2 Memory and Intention

I cannot perceive the entire universe at once. In fact, at any one time I can perceive only rather a small sector of the universe. This kind of limitation on our perceptual powers is what gives memory much of its significance. Memory allows us to retain knowledge of things we may not now be in a position to perceive. It also frees us from having to reason afresh about things we have reasoned about in the past. Through memory we can retain the knowledge acquired as the result of those earlier inferences.

There are limitations on our practical reasoning too. The process of assessing reasons for action is not instantaneous: deliberation about what to do takes time. But when the opportunity for action arises we do not always have the time to assess properly the merits of a course of action. In the time it would take to deliberate properly the opportunity for action may have passed. This limitation is part of what gives *intention* its significance. By forming firm prior intentions we can effectively transmit over time or *preserve* the results of earlier deliberation. This means that deliberation can be carried out in advance, when things are less pressing and information more readily accessible. Relatedly, general ‘standing’ intentions or *policies* free us from repetitive deliberation. I do not every morning have to meditate on the advantages and disadvantages of cleaning my teeth. The adoption in advance of a policy of cleaning my teeth saves me a great deal of time. Just as memory frees me from having to return to something already perceived, so intention frees me from having to return to practical matters already considered. Memory and intention are thus broadly analogous: they each mitigate our limitations at a time, by making resources available across time³.

Given that memory and intention are broadly analogous, a sensible question

³The significance of resource-limitation is a central theme in (Bratman, 1987). Note that my concern throughout is with prior future-directed intentions; I will not discuss alleged present-directed intentions, or “intention-in-action”.

to ask is whether anything paralleling Butler's objection that memory presupposes personal identity holds of intention. To answer this question we need to know what was Butler's objection. He wrote "one should really think it self-evident, that consciousness [ie.memory] presupposes and cannot constitute personal identity, any more than knowledge can constitute truth, which it presupposes". It is not entirely clear what what he had in mind here - but I think that we can distinguish at least two forms of circularity or presupposition, Butlerian in their inspiration at least.

The form of memory that Locke had in mind was conscious, experiential, sensory memory - the sort of memory that presents past experiences and deeds 'from the inside' as it is sometimes put. But what is it for a past event to be presented in this way - from the inside? One might think that it involves, at least, that the episode is presented as having happened to *me*. To say that I remember from the inside ϕ ing is just an elliptical way of saying that I remember from the inside *myself* ϕ ing. It is simply built into the content of the memory that it was me - the person who now remembers - who was involved in the event remembered. In this sense personal identity over time is built into the representational content of the memory. Experiential memory represents personal identity over time. We can call this the *representational presupposition* of personal identity by memory.

I think that a similar form of presupposition is discernible in intention. Just as one's remembering from the inside ϕ ing is arguably just the same thing as remembering from the inside *oneself* ϕ ing, so one's intending to ϕ is arguably just the same thing as one's intending *oneself* to ϕ . In this sense the identity of the intender and the eventual agent of the action is built into the content of the intention. Intention representationally presupposes personal identity.

Perhaps there is a sense in which one can plan for someone else to do something; but such a plan must be executed indirectly, by directly executing intentions to do things oneself, things which will in turn cause the other to act in the

way planned. But what neo-Lockeans call the ‘direct psychological connection’ between intentions and their execution is a connection that seems to involve one’s intending oneself to ϕ . One’s identity with the agent of the intended action is part of the representational content of this sort of intention.

In addition to representational presupposition, I want to identify a second kind of presupposition that Butler may have had in mind, which I will call *causal-explanatory* presupposition. The direct psychological connection between an earlier experience and its later recall presupposes personal identity in this sense if the causal link between the experience and its later recall involves personal identity as a causal factor. And one might well think that personal identity is just such a causal factor. Simply, one might think that part of the reason why I now recall some earlier experience is that it was me who underwent the experience. My having been there at the time is part of the explanation of why I now have a memory state. If it had been someone else who underwent the experience then I would not now be in a memory state. Far from personal identity over time being partly constituted by the causal link between experience and memory, personal identity over time is itself a causal factor in the link between experience and memory. There is a causal-explanatory presupposition of personal identity by memory.

In a similar spirit, one might object that the link between intention and action presupposes personal identity as a causal factor: part of the reason that my intentions get executed - if they do - is that I stick around. My persisting into the future is part of the causal explanation of my intentions giving rise to action. If I did not persist into the future my intentions would not get executed. Intention is a form of control over the future that relies on the independent fact of personal identity over time. It gets things completely backwards to claim that part of why I persist into the future is that my intentions get executed. There seems to be a causal-explanatory presupposition of personal identity by the link between intention and action

I will assess these two presupposition objections, in the next two sections respectively.

3 Representational Presupposition

It was objected that experiential memory presupposes personal identity in the sense that this form of remembering ‘from the inside’ involves the representation of personal identity over time. Such memory presents earlier experiences and deeds as things that *I myself* underwent or did.

Now, I want to suggest on behalf of the neo-Lockean that this description of experiential memory is not mandatory. What does seem indisputable is that such memories inherit some of the character of the sensory experience from which they derive. In particular they inherit what is often called the ‘egocentric’ content of sensory experience. When I remember things ‘from the inside’ those things are remembered as being to the left, above, in front, and so on, in rather the same way in which things are currently perceived as being to the left, above, in front, and so on. From the point that things are remembered in these egocentric ways, it seems a short step to the claim that things are remembered as having been related to *me*, as the term ‘egocentric’ rather suggests.

But those familiar with discussions of imagination and the self will pause over this short step. There is a controversy about whether sensory imagining - imagining from the inside ϕ ing - does or does not fall short of imagining *oneself* ϕ ing. Many think that it does fall short of this first personal content. To imagine witnessing the creation of the Earth, or to imagine being Napoleon, does not obviously amount to imagining *myself* - a particular human being - witnessing the creation of the Earth, or being Napoleon⁴. I can if I wish accompany these episodes of sensory imagination, with a narrative about my travelling back in time to do these things. But I need not: any such narrative is in principle

⁴This is one conclusion of the brilliant (Williams, 1973)

separable from the basic sensory episode. Perhaps the egocentric content of an episode of sensory imagining needs to be specified in relation to a notional point of view in the imagined scene: things are imagined as to the left, and so on, in relation to this point of view. The idea is that one need not think of this imagined point of view as one's own.

I'm not endorsing the details here. I just want to suggest that the neo-Lockean might respond to the charge that experiential memory representationally presupposes personal identity by arguing that, as in the case of imagination, the egocentric representation of a remembered episode does not straightforwardly entail that one represents one's own involvement in that episode. The egocentricity might be accounted for in terms of relations to some more abstract point of view in the remembered scene. Although we automatically assume that this point of view was our own point of view, this is a separable belief. This belief does represent that the subject of the earlier experience was me; but it can be detached or separated from the memory episode proper. One might say that memory does not *intrinsically* representationally presuppose personal identity. Again the analogy with imagination may be helpful: when I imagine from the inside witnessing the creation of the Earth, I imagine certain events in relation to a point of view. Whether I accompany this episode with the supposition that the point of view is an alien cosmonaut's point of view, or my own time-traveller's point of view, is incidental. These are separable suppositions, which are not intrinsic to the sensory episode of imagining.

To illustrate the separability of the belief in personal identity from memory, one can imagine that one comes to believe that one has received memory-trace implants of the sort that Parfit and Shoemaker discuss. Or perhaps that one is the offshoot of some person's earlier 'fissioning'. In such cases one might cease to believe that one was oneself the subject of the experiences now being remembered - or 'quasi-remembered'. Yet the sensory memory state would remain stable in the face of this change of belief. The memory state is in this sense

belief-independent, in rather the same way that the Müller-Lyer arrow illusion continues to present a certain sensory appearance even when one acquires the belief that it is misleading.

So it is at least intelligible that the neo-Lockean can respond to the charge of representational circularity in the case of memory, by insisting that the egocentric character of sensory memory is not intrinsically first-personal. That it was me who underwent the experience recalled is something I automatically assume, but this belief is separable, or extrinsic, to the memory itself. The natural next question now is this: does that sort of response translate into a response to the charge of representational circularity in the case of intention? Recall the charge here was that the direct psychological connection between intention and action involves one's intending *oneself* to ϕ . Intention represents the agent of the future action as oneself.

Unfortunately for the neo-Lockean it is at this point that the analogy between intention and memory begins to break down. Intending to ϕ is not a sensory state in anything like the way that remembering ϕ ing or imagining ϕ ing are sensory states. Perhaps sometimes when one forms an intention to act one forms a sensory image - a sensory anticipation - of acting, but this does not seem to be the case generally. We can easily imagine someone who manifests intention formation simply by writing words on a to-do list, or by uttering the words 'I will do ...'. In fact it is plausible that no sensory state is necessary or sufficient for intending to ϕ . Wittgenstein noted that the occurrence of distinctive imagery is neither necessary nor sufficient for understanding words. For similar reasons, no sensory event is necessary or sufficient for intention formation: The occurrence of such an event cannot guarantee the presence of the dispositions to means-end reasoning, planning, and acting, which are distinctive of intention.

Why does this matter? The response to the representational circularity charge in the case of experiential memory relied on the fact that such memories have a sensory core, whose egocentric character can be specified in terms that

do not presuppose personal identity over time. The representation of personal identity over time is a belief that can be separated from this sensory state. The memory itself is a belief-independent sensory state. But if intentions are not sensory states like this, the representation of one's identity cannot be argued to be separable in the same way. In fact I want to argue that intentions are *belief-dependent*.

Although the details are controversial, it is generally agreed that there are *belief-constraints* upon intention. For example, it is extremely plausible that one cannot intend to ϕ if one believes that one's ϕ ing is impossible. Some make the stronger, although still entirely plausible, claim that one cannot intend to ϕ without believing that one will ϕ . In intending one takes to be true the proposition that one will act. In making the commitment to action distinctive of intention, one regards as *settled* the question of whether or not one will act. It is thus very far from clear that the belief that one will act can in any sense be 'separated' from one's intending to ϕ ⁵

This crucial contrast with memory can be illustrated with an example. Suppose I have formed the intention to go to the theatre tomorrow evening. Unfortunately, I subsequently come to believe that I will be executed at dawn tomorrow. Evidently, it will be impossible for me to go to the theatre tomorrow evening if I am executed at dawn tomorrow, for I will no longer even exist tomorrow evening if that happens. Would my intention persist independently of this change of belief? Surely not. I cannot continue to intend to ϕ at t if I come to believe that I will not exist at t and *a fortiori* that it will not be possible for me to ϕ at t . This contrasts with the fact, already noted, that a memory of ϕ ing will persist even if I come to believe that I could not have been the one who ϕ ed; in contrast an intention to ϕ will be destabililised - if not immediately destroyed

⁵David Velleman has argued that intentions are *identical* to a certain kind of belief that one will act. If he is right then there would obviously be no question of the belief being separable from the intention. See (Velleman, 2000, Introduction)

- if I come to believe that I will not be in a position to ϕ . So the representation of one's identity over time is not extrinsic to, or separable from, intention in the way it might be in the case of memory. The formation and persistence of intentions is belief-constrained, and it is constrained by belief with first person content.

I think it can be shown additionally that it is not merely the formation and persistence of intentions that involves first personal representation, but also their direct execution. I commented when I introduced the charge of representational circularity that although one could in some sense plan for someone else to do something, the connection between such a plan and that person's eventually acting would be indirect, and would be different from the direct causal connection between intentions - ordinary first person intentions - and action. However, Derek Parfit has disputed this: he makes the somewhat strange proposal that if one were about to undergo fission, one could form a kind of intention - what he calls 'quasi-intentions' - which were not first-personal; but which could nevertheless be executed with the same directness as ordinary intentions. It's illuminating to see where this proposal goes wrong, so I'll quote the relevant passage at length.

I could quasi-intend both that one resulting person roams the world, and that the other stays at home. What I quasi-intend will be done, not by me, but by the two resulting people. Normally, if I intend that someone else should do something, I cannot get him to do it simply by forming the intention. But if I was about to divide, it would be enough simply to form quasi-intentions. Both of the resulting people would inherit these quasi-intentions, and unless they changed their inherited minds, they would carry them out. Since they might change their minds, I could not be sure that they would do what I quasi-intended. But the same is true within my own life. Since I

may change my own mind, I cannot be sure that I will do what I now intend to do. But I have some ability to control my own future by forming firm intentions. If I was about to divide, I would have just as much ability, by forming quasi-intentions, to control the futures of the two resulting people. (Parfit, 1987, 261)

Suppose then that before fission I quasi-intend 'Righty' to roam and quasi-intend 'Lefty' to stay at home. The problem here is that after fission Righty and Lefty will each find themselves with *both* quasi-intentions. Each will have inherited the intention that Righty roam and the intention that Lefty stay at home. In order for these quasi-intentions to give rise to action as planned, Righty will need to realize that he himself is Righty, and Lefty will need to realise that he himself is Lefty. Until the agent grasps that they are identical with the relevant quasi-intended agent, the quasi-intention will not be acted upon.

This contradicts the supposition that the directness of execution of quasi-intention would be the same as the directness of execution of first-personal intentions. First-person intentions do not have to give rise to action via some 'intermediate grasp' that I am the intended agent of the action. The execution of any sort of intention generally involves an intermediate grasp of the intended *time* of action: If I form the intention to act at three o' clock, then the execution of the intention relies on my keeping track of the time, and realizing, when the time for action comes around, that three o' clock is *now*. Quasi-intentions would generate a parallel - and otherwise entirely unnecessary - requirement that the agent keep track of whether or not the quasi-intended agent is *me*. In contrast ordinary intentions with first-person content are executed without the need for any such intermediary realization⁶. This is one way in which the con-

⁶As Evans puts it, there is no need to for any 'skill or care (not to lose track of something) on the part of the subject.' (Evans, 1980, 237) The distinctive immediate motivational role of the first person was noted in (Perry, 1979)

nection between intention and action deserves to be called a *direct* psychological connection.

It is time to draw together the strands of this discussion of the representational presupposition of personal identity by intention. Since intentions are not belief-independent sensory states like experiential memories, the representation of personal identity over time is not separable. The formation and persistence of intention is constrained by first-person belief in a way unparalleled by memory. Furthermore, as we have just seen, the causal directness of the execution of intentions requires first personal intention. So it can be concluded that intention representationally presupposes personal identity. Since it is an adequacy condition on the neo-Lockean theory that the direct psychological connection between intention and action does not presuppose personal identity, this result would seem to be a problem for the theory.

In fact I cannot say for sure whether representational presupposition results in a vicious form of circularity. It is unclear whether explaining *F*s in terms of the representation of *F*s is as bad as analysing *F*s in terms of *F*s themselves. For example, an account of *x*'s being funny in terms of *x*'s being generally *judged* to be funny does seem to be completely uninformative in the way that an account of *x*'s being funny in terms of *x*'s being funny seems to be.

I take it that it would be more clearly serious if the direct causal connection between intention and action involved not merely the representation of personal identity, but personal identity itself. Since causal-explanatory presupposition threatens just this result, it is to causal-explanatory presupposition that I now turn.

4 Causal-Explanatory Presupposition

The charge of causal explanatory presupposition was this: my persisting into the future is part of the causal explanation of why my intentions give rise to

action. It is in part because I stick around that my intentions give rise to action. Personal identity over time is a causal factor. An account of personal identity in terms of the causal relation between intention and action would thus be circular, because the causal relation in question itself involves personal identity over time - and not merely the representation of personal identity over time. This form of presupposition looks more serious than representational presupposition.

The possibility of fission cases promises a powerful neo-Lockean response to this objection. The most realistic - though of course still imaginary - illustration of this possibility was provided by David Wiggins⁷. He imagined a complex operation in which my brain is divided into two and placed into the skulls of two debrained humans. Although the interpretation of this state of affairs is controversial, I will assume - along with most neo-Lockeans - that this operation results in my going out of existence, and the coming into existence of two distinct persons, each psychologically similar to me. If, before fission, I had intended myself to write a paper, it is plausible that my two fission descendants would each inherit an intention to (themselves) write a paper.

Suppose that one of these people - call him Righty - does end up writing a paper. The neo-Lockean will claim that the way in which my earlier intention gives rise to Righty's action is the same as the way in which my earlier intention would have given rise to my own action in an ordinary survival case. But since Righty is not identical to me, it cannot be correct to object that this causal relation involves personal identity as a factor. The causal link between intention and action does not involve personal identity as a factor because the very same causal link could obtain down one 'branch' of a fission case, where there is not personal identity over time.

Why think that the causal link will be the same? The natural thought here is that the appropriate kind of causal dependence, whatever it is, is wholly determined by matters neurophysiological. So it follows that if the brain is divided

⁷(Wiggins, 1967, 53)

then there must be a branching in this appropriate kind of causal dependence.

This kind of argument from fission is undeniably a powerful reply to the charge of causal-explanatory presupposition; but it does seem to make a significant assumption. It assumes that the appropriate psychological causation is wholly a matter of some *intrinsic*, presumably neurophysiological mechanism, supervenient on the brain. I think that for the case of experiential memory this assumption is quite in order. Such memories are merely the intrinsic marks left on us by earlier experience. And such marks can be transmitted in branching form, as the case of brain division suggests. But I want to question the assumption that the causal dependence of action upon intention is merely a matter of this kind of intrinsic mark transmission, and with it the correlative assumption that the causal dependence can take a branching form. In fact I think it's arguable that the causation of action by intention typically involves interaction with *extrinsic* environmental factors, and does so in a way that puts pressure on the response from fission.

It is sometimes commented that fission may create 'practical problems' for those involved. The implication is usually that these are not problems for the *theorist* of personal identity. However, since the causation of action by intention features in the neo-Lockean theory, practical problems may well turn out to be theoretical problems. To take an example: suppose that before fission I intended not merely to write a paper, but to get the paper published in a particular journal. Suppose that after fission Righty gets his paper published in this journal. Is the causal dependence of his action upon my earlier intention the same as would have been the causal dependence of my action upon my earlier intention had I survived in the ordinary way and got my own paper published in the journal?

I want to suggest that it is not. In the case of fission there was a non-negligible chance that the other fission offshoot - call him Lefty - would have got his paper published first. Since the journal is unlikely to accept two very

similar papers, this would have blocked Righty from executing the intention to get his paper published. Whereas had I survived in the ordinary way such undermining would not have threatened. So the occurrence of branching in this case *weakens* the dependence of action upon my earlier intention down each branch compared to the dependence that would obtain in the ordinary survival case. The degree of dependence down a branch of the fission case does *not* mimic the dependence in an ordinary survival case. Personal identity raises the probability of action following intention: in other words personal identity is a causal factor.

This can seem like a surprising result. For any survival case of intention-action, one might think that there could have been an additional branch; and that this branch would be completely extrinsic to the connection between intention and action. How could the existence or otherwise of such an extrinsic branch causally affect the connection between intention and action? Surely this would be a form of magic or ‘action-at-a-distance’.

The crucial point to make in response to the worry is this: the period between intention formation and action typically leaves time for environmental interaction to occur⁸. In the case considered there is time for the existence of one branch to disturb the efficacy of intention down the other branch. The existence of another branch is extrinsic or environmental with respect to the intrinsic pattern of neurophysiological states lying between intention and action; but it is relevant nonetheless.

It might be objected at this point that the case of intending to get one’s paper published is somehow a peculiar or exceptional case, invidiously selected. If that were so, then perhaps the neo-Lockean would be entitled to ignore the causal-explanatory presupposition that it involves.

⁸Tim Williamson’s defence of the claim that externally individuated states such as knowledge can be causally efficacious turns on a similar point about environmental interaction. See (Williamson, 2000, Chapters 2 and 3). See too (Yablo, 1997) and (Yablo, 2003)

However the intention is in fact entirely typical. There are countless quite homely cases of the same sort: I intend to find my wallet. Again, my fission products would each have a reduced chance of executing this intention, for they will tend to undermine each other by finding my wallet before the other.

This interference in the causation of action by intention is amplified by another feature of intentions: their role in structured plans. Suppose that I intend to find my wallet and intend to find my keys, both as preliminary parts of a larger plan to go and do my shopping. Since the occurrence of branching will tend to interfere with the execution of both preliminary intentions, the interference in the execution of the larger plan is multiplied, like compound interest. The greater the sequence of intentions whose execution is identity-sensitive, the more personal identity - the absence of branching - becomes causally crucial.

There is a lot more to be said about precisely which intentions display this form of causal sensitivity to personal identity. But I hope it is fairly obvious that personal identity becomes causally relevant roughly because these are intentions that we form on the assumption of our own numerically identical persistence into the future; these are intentions we would not form if fission was known to be an ordinary occurrence. But that does not detract from the importance of these intentions; on the contrary, the fact that these are intentions we form on the assumption of our own numerically identical persistence into the future means that they are intentions which very clearly exemplify our capacity - paraphrasing Locke - to consider ourselves, as ourselves, the same thinking - and acting - thing in different times and places. They are paradigmatic of personhood.

There is a range of broadly 'internalist' moves that might be suggested on behalf of the neo-Lockean here. By restricting their account to highly 'proximal' intentions to act in the very near future; or intentions to perform very basic bodily actions; or even to the connection between intentions and mere 'tryings', the neo-Lockean might avoid the causal-explanatory presupposition of personal identity, by minimizing the relevance of environmental interaction.

I don't have the space to assess these moves here. What I hope has been established is that the causation of action by intention cannot simply be assumed to be a matter of intrinsic mark-transmission, like the dependence of experiential memory on earlier experiences. This serves to corroborate one general upshot of the earlier discussion of representational presupposition: it should not be assumed that circularity challenges posed by intention can be met in the same way as circularity challenges posed by memory.

References

- Bratman, Michael. 1987. *Intention, Plans and Practical Reason*. Cambridge: Harvard University Press. Reissued 1999 by CSLI Publications.
- Evans, Gareth. 1980. *The Varieties of Reference*. Oxford: Blackwell.
- Fine, Kit. 1994. Essence and Modality. *Philosophical Perspectives*, 8, 1–16.
- Lewis, David. 1976. Survival and Identity. *Pages 17–40 of: Rorty, Amelie (ed), The Identities of Persons*. Berkeley: University of California.
- Parfit, Derek. 1975. Personal Identity. *Pages 199–233 of: Perry, John (ed), Personal Identity*. University of California Press.
- Parfit, Derek. 1987. *Reasons and Persons*. Second edn. Oxford: Clarendon.
- Perry, John. 1976. The Importance of Being Identical. *Pages 67–91 of: Rorty, Amelie (ed), The Identities of Persons*. Berkeley: University of California.
- Perry, John. 1979. The Problem of the Essential Indexical. *Nous*, 3–21.
- Shoemaker, Sydney. 1984. Personal Identity: A Materialist's Account. *In: Shoemaker, S., & Swinburne, R. (eds), Personal Identity*. Oxford: Blackwell.
- Shoemaker, Sydney. 1997. Self and Substance. *In: Philosophical Perspectives 11, Mind Causation and World*. Boston, MA: Blackwell.

Velleman, David. 2000. *The Possibility of Practical Reason*. Oxford: OUP.

Wiggins, David. 1967. *Identity and Spatio-Temporal Continuity*. Oxford: Blackwell.

Williams, Bernard. 1973. Imagination and the Self. *Pages 26–45 of: Problems of the Self*. Cambridge: CUP.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.

Yablo, Stephen. 1997. Wide Causation. *Philosophical Perspectives*, **11**, 251–281.

Yablo, Stephen. 2003. Causal Relevance. *Philosophical Issues*, **13**, 316–328.