



ORDER: GOD'S, MAN'S AND NATURE'S

A Sketch of How Agents Enter the World of Events Tim Mawson

Conference Paper: Where Free Will and Nature Meet

If we are drawn to Indeterminism, we shall probably regard the movements that characteristically accompany certain medical conditions, palsies and the like, as brought about by factors which include a certain amount of randomness and thus which are undetermined. But we hardly regard a person's arm flying up as a result of a nervous disorder of this sort as that person's raising his or her arm. We do not do so even if, coincidentally, by flying up as it does, it happens to fulfil an occurring intention of the agent in question, or even if - as Davidson's example of the nervy mountaineer has shown us - it happens to be caused by that occurring intention. What is necessary for a bodily movement to be an action is not that it be uncaused, but rather that it be caused and caused – I shall suggest – not in the right way by events but by the right thing, an agent, the causal potency of which transcends that of the events which he or she undergoes. That's the view I'm going to sketch this morning for your consideration and critical response.

Consider a situation in which I tell the truth after some reflection and despite knowing that a certain lie would in this case remain undiscovered and bring me greater advantages. It would be foolishness to deny that for as complete an explanation as can be given of my choice to tell the truth, one would need to mention prior events of a general sort - one would need to mention how it was that I got into the position where I faced the particular choice in the first place; and one would need to mention prior events of a more particular sort, for example, how it was that my character had been formed by my attending in my impressionable youth a very impressive series of

lectures on Kant's ethics. These together would have led to my capacities and my dispositions as to how to exercise them – that I am able to speak and that I have developed the virtue of truthfulness. Most obviously, one would need to mention particular events in the most recent period of my mental life - the period in which I was coming to my decision - e.g., let us suppose, that I reflected on the question 'What would Kant do?' This 'world of events' would, *in toto*, set the background to my choice and would, we may suggest, not leave what choice I made a matter of causal indifference; taken together these factors would incline me towards truthfulness; we may suppose even that they would make the chances of my telling the truth very much higher than fifty/fifty. My suggestion is that for a situation of genuine choice in addition to all this – unlike the case with non-agential indeterministic systems (say roulette wheels) – one would need to mention an agent – me and what I did, choose to tell the truth. Someone who pointed out merely that, given my upbringing, character, and mental life, it was quite probable that I'd end up telling the truth would *not* have offered as complete an explanation as could be offered for the happening that was my telling the truth; to offer as complete an explanation as could be offered for that sort of happening (indeed I shall suggest to offer an explanation of it qua action in contrast to mere event), he or she would have to add mention of the fact that I ended up providing my own bit of causal 'oomph' as I shall put it to the world, that I as an agent entered the world of events.

The mention of this additional thing, me, and the event in which this additional thing most immediately features as the cause, the truth's being spoken, is needed to make it clear that what was always quite likely to happen anyway (the truth's getting told) was in this particular instance neither uncaused by anything other than the events preceding it, nor causally necessitated by something exterior to me. This I am suggesting is what makes the difference between its being my action rather than an event I underwent or the action of someone else. If there had been no cause in addition to those that inclined me to write the truth, so that the fact that 'I' ended up writing the truth was similar to the fact that a particular number came up on a roulette wheel on a particular spin, then 'my' writing the truth would not have been genuinely an action that I performed, rather than merely an event that I underwent; I would not have written the truth; the truth would just have ended up being written, and written by 'me' only in the attenuated sense that its getting onto the page involved a causal chain which went through my body and mind. (Thus, in the text, I use scare quotation marks around 'me', 'I', and so in such cases.) Had that been the case, then of course no moral credit would have ended up coming to me for 'my' truthfulness (unless a

relevant self-forming action had taken place earlier, which self-forming action would itself, I suggest, have had to have fitted my model of agency). Authorship of this sort then is a necessary condition for moral responsibility, because without it our bodily movements cease to be our actions.

Harder to imagine than 'my' telling the truth's being uncaused by me in this fashion is a case where something external to me causally necessitated that 'I' tell the truth. But we can construct such a case by imagining a scenario where someone implants in my brain a 'Frankfurt'-style chip, such that he or she could then control without possibility of error (later randomness notwithstanding) what movements my body made, right down to the details of how my lips moved and thus what 'I' ended up saying. Let us in this case imagine someone who wishes to ensure simply that the truth gets told on this occasion and has implanted a chip with that end in view. He operates the chip and thus makes 'me' speak the truth. I meanwhile have decided to speak the truth and, as I observe my lips moving in accordance with my will, I take myself to be doing so. But – rather as in the case of a driving student who has not realized that the instructor has taken over driving the dual-control car in which both sit - I am in fact mistaken when I think of myself as having spoken the truth (in anything other than the attenuated sense that it is through my bodily movements that the truth has got told). I think that I've provided some causal oomph to the world and that it's in part as a result of that that the truth got told, but in actuality either I've failed to provide any oomph or the oomph that I have provided hasn't causally contributed to the truth being told (which of these will have happened depends on precisely how the chip is supposed to work). In either case, I am of course unaware of it; my body acts in accordance with my will but it is not following my will. Thus it is not, strictly speaking, me who has been truthful. Such would be a scenario where things would look very similar from both the outside and the inside to the way things look in more ordinary cases. But in such a scenario, unbeknownst perhaps to everyone but the person who implants and operates the chip, the fact that 'I' would have ended up doing that which it was always quite probable I'd end up doing – telling the truth - would not have been down to me; it would have been down to preceding events and this intervening brain surgeon. Again, I would fail to be morally responsible for the truth being written down on the form through its failing to be an action I was performing.¹

These two sorts of case – 'uncausedness' by anything but prior events (and thus 'uncausedness' by me [unless I caused these prior events with a self-forming action])

and causal necessitation by an exterior agent or body (and thus 'uncausedness' by me) however are different from the ordinary and in particular from our example as originally imagined. We suppose in our example that there was indeed something, or rather someone, which made the truth get told. This was someone who was above and beyond the factors which causally inclined me to do what I did and was someone who made that which may well have been likely to happen 'by chance' – i.e. on balance of probabilities - anyway happen not as a result of chance. We suppose that that someone was not an intervening brain surgeon exterior to me, but rather that that someone was me. And thus we need to make mention of this supposition, the supposition that I have added some agent-causal oomph, if we are to give as full an explanation as we think we can give of why I told the truth, indeed explain at all why it is that I told the truth rather than merely that the truth got told. When I chose to tell the truth, that decision was caused in part by my background, my character, and my being aware of the reasons that favoured it (certain events in my mind); these together might have brought it about that 'I' told the truth even had I —the agent—not chosen as I did; they might have inclined me in that direction and I not done anything about it, just passively letting the events unfold that way; or they might have inclined an intervention-prone brain surgeon to operate a chip that then necessitated that 'I' do as I did. But in fact neither of these things happened; the totality of events was not, as it happens, the only thing that brought about my telling the truth; I was in fact a part cause of its happening: I superadded to all of this causal commerce at the moment of choice. And thus the telling of the truth was an action that I performed.

So, to recap, I am suggesting that for me to be the author of my action of telling the truth, it is not necessary that nothing outside of myself would have proved causally sufficient for the truth to be told by 'me'. It could be the case that the inclining causes would have proved sufficient to bring about that the truth be told in conjunction with the right sort of intervention-prone brain-chip implanter (one who would have intervened if he'd detected the right sort of pre-waver brain state), but, were that to have in fact obtained, then the truth's being told, whilst it would be an event that would have been brought about by my body, would not have been me bringing about the event, would not have been me acting: it would have been brought about by the events that were these inclining causes and the action of the intervening brain surgeon. The answer to the question, 'Was anything outside of my control causally sufficient for a particular outcome, the truth's being told?' might then be 'Yes' even when we consider a situation in which that same outcome has in fact been brought about by me acting. We should not then seek to identify the nature of authorship by

supposing that where it is present the person's choosing is *necessary* for the particular outcome that it actually brought about; it might not have been necessary.² That I take as the lesson from Frankfurt-style cases. Rather than, the essence of ultimate authorship is that, where it is present, I in fact cause (even if other things would have caused had I not) whatever it is I do. For this to be the case, it has to be true that nothing outside of me causally necessitated that 'I' do whatever it is I do, but not that nothing outside of me causally inclined me to do whatever it is I do, nor even then – as we have just seen – that nothing outside of me would have proved sufficient to make 'me' do whatever it is I do if I hadn't done it myself. It just has to be the case that in fact one of the causes of my doing it was me. If I add my own causal oomph to the world of events in this way, the resulting movements of my body and the outcomes that they bring about are actions that I am performing using my body, rather than merely events that my body is undergoing.

Not all actions are free. If we act in ignorance of the nature of our actions or in rushed circumstances (like the doctor who inadvertently harms his patient with a medicine that would ordinarily prove beneficial) or if we are coerced (like the bank manager whose family are threatened with death if he does not assist in the robbing of his bank), we are not properly regarded as free in the sense necessary for moral responsibility even though we are not rendered by ignorance, rush, or coercion non-agents; these sorts of difficulties do not reduce our bodily movements to mere events. So, for an action to be free in the sense necessary for moral responsibility, not only does it need to satisfy the authorship condition which I have been sketching and thereby get to be an action, but it also needs to satisfy the condition that its agent will what he or she does under a morally salient description. The doctor did harm his patient, but when he administered the drug to his patient he did not will the action under this description; we know that for certain as we know he didn't even believe that description to be true of what he was doing. This being the case, we say of his action in harming his patient that it was not free in the sense necessary for moral responsibility. He performed the action that was his harming his patient, but – willing that action under the description 'healing my patient' - he non-freely performed the action. If it was also the case that he was not negligent in not knowing of the drug he administered that it would be harmful, we would therefore think he is not to be negatively morally assessed. The bank manager did know that the description 'robbing my bank' was true of his action in assisting the bank robbers, but he too was not acting freely when performed the action that was robbing his bank; he was not acting freely as he willed his action under an entirely different (and indeed again

morally laudable) description ('doing the least bad thing necessary to save the lives of my family'); thus again he non-freely performed an action and thus is not to be negatively morally assessed for it. He was coerced – that's the nature of coercion, getting someone to perform an action that one wills under one description which they'd not will it under by making it fit another description under which they'll then will it. But such cases are unusual and so a large number of our actions get to be free in the sense necessary and sufficient for moral responsibility.

On the account I have been developing then, one true substitution for 'x' in 'x causes y' is an agent whenever the outcome, y, is genuinely an action rather than merely an event. When it is a free action, the substitution is a well-informed, unrushed and uncoerced agent. It is profitable to ask at this stage whether agents are the sorts of things that themselves can feature as effects, as ys, in the causal relationship. There is one – and only one - context in which it appears that they can. It seems natural to say that parents cause the agents that are their offspring to come into existence and we may perhaps express this fact without its sounding too unnatural as their 'causing agents'. Be that as it may, once an agent is in existence, the only thing that other agents and events can cause with respect to him or her is for him or her to undergo events (and, if this is not considered an event, possibly to cease existing); once an agent has started existing, he or she cannot in any way himself or herself be caused, be the effect of other agents and events (although of course other agents and events can cause him or her to cause something else, through e.g. inclining him or her towards a particular outcome by presenting him or her with reasons to favour that outcome). That being the case, once we are considering an already-existing agent, we may say of him or her that nothing can cause him or her; he or she cannot himself or herself be the effect of anything preceding him or her. And that being the case, if he or she then goes on to cause some effect, that effect *must* be undetermined by preceding events - because it was caused by him or her and he or she must be undetermined by preceding events. It is this then that explains most directly why it is that genuine actions cannot be determined: genuine actions require agent-causal oomph and the first effect of the agent providing agent-causal oomph has as its cause— obviously – a pre-existing agent, a pre-existing agent being the sort of thing that by its nature cannot itself be the effect of any cause.³

As we have seen, whilst there can be actions that are not free (if they are done in ignorance, a rush, or coerced), an alteration being made to the world in such a way

that an agent who would otherwise have stood at the start of a chain that led to some causal oomph entering the event-causal chain is bypassed and a similar causal oomph is provided by mere events or some other agent, would not lead to a non-free action being performed by that first agent, but rather to his or her performing no action at all (though the fact that it is not his or her action might in principle be obscure even to that first now-former agent). Were Determinism true then, we wouldn't really perform any actions at all; all we would actually do is participate in events. Agents 'disappear' from a Deterministic universe, although of course creatures in all other respects like them might take themselves to be agents in such a universe, being in a situation analogous to the driving student in the dual-control car who hasn't realised that his instructor has been driving it for the entirety of the lesson. Were Indeterminism true, we might still not perform any actions (agents 'disappear' from some Indeterministic universes too). Again we'd just perhaps mistakenly think that we did when all that was really happening was events causing others in an indeterministic fashion. It is in this latter connection that the need to posit agent causation (in addition to indeterministic event causation) may be made most obvious by considering what has come to be known as the problem of 'luck', a problem which – I suggest – needs the non-disappearing agent that I have been sketching for its solution. First, let us state the problem; then we can see how a theory of agency that posits in-eliminable agent causation can solve it in a way that the theory that posits that agent causation may be reduced to event causation cannot.

If Libertarianism is right, then we could – right up until the last moment – chose differently from the way that we actually do in situations of genuine choice, however unlikely our doing so may have been rendered in particular cases by events that preceded that moment, e.g. in cases where we realized that we had overwhelming reason to choose the way that we did. One way of presenting this fact is by imagining another possible world in which our correlates *do* choose differently from the way that we actually choose. So, in the actual world, 'Timmy One' as we may call him chose to come along to this conference rather than – as it suddenly occurred to him he could do yesterday evening – give up on trying to improve his thinking on difficult topics altogether and live off his wife's earnings for the rest of his life. In a near possible world, everything up until the moment at which I came along here remaining the same then (that's plausibly – though not unarguably – what makes it so near), my correlate - 'Timmy Two', we may call him – chose to give up on being a philosopher. No events were different in these universes prior to these choices being made differently by Timmy One and Timmy Two, so, someone might say, it is just a matter

of luck that the actual world contains philosopher Timmy One, rather than wastrel Timmy Two and surely I then cannot take any credit for its doing so.

I concede that this 'luck objection' would indeed be pressing were there merely to be event causation, because it is indeed true that there would then be nothing other than events to which one might turn to avoid the charge of luck and – in that the events preceding the choice were, *ex hypothesi*, exactly the same in the two worlds – so indeed it would have been true that we just got lucky in getting Timmy One, a feature of the world which is then hardly to my credit. But that is definitively *not* the case on the model that I have been sketching. On this model, there is something in addition to events to which we may turn: me. In the actual world, I chose to come along, rather than give up on trying to improve my thinking on philosophical topics. It is simply not true that this was me 'getting lucky'; it was me 'making my own luck', if you will, in that it was me choosing to come along here rather than go along with the fleeting desire to give up on all strenuous intellectual activity. But this is not really luck at all. Unlike on the pure event causation model, there is something to which we can apportion the credit for the actual world going the way it does (rather than the way it does in the world in which Timmy Two lives); that thing is me. It's down to me that the actual world ends up being a world which contains philosopher Timmy One, rather than wastrel Timmy Two.

This then is the view and at least the beginnings of some arguments in its favour. We have in addition to these arguments, I suggest, direct experience of its truth, of our providing the agent-causal oomph of which it speaks. Obviously, we most clearly experience the causal oomph we provide qua agents in cases in which other causal influences are as absent as they can be. So, in a moment, I'm going to call upon you to imagine with me a case where other causal influences are stripped away as much as possible and, then, whilst an in-eliminable residue of event-causation will remain, we shall, I suggest, be able to see the contribution made in addition by agent causation. I shall suggest that if you reflect on yourself as you were when you last made a choice between options which you regarded as genuinely indifferent and ask yourself what it was that in the end made you do whatever it was you actually ended up doing, you will most probably recall an experience of your giving the world an oomph, not simply that of the event that was your deciding to do whatever it was you did decide following on from the oomph provided by the thought of yours that immediately preceded your decision. Of course, this sort of self-reflection can be

deceptive. Perhaps what you mistake for agent-causal oomph is only event-causal oomph bubbling up from your sub-conscious. We know that in some cases this sort of self-reflection is certainly deceptive in just this way. If we asked the person who had had a stage magician 'force' a particular card on him or her whether he or she felt himself or herself to have added a particular oomph of his or her own along the way to picking out the card that he or she did pick out, the person would reply that he or she had indeed felt himself or herself to do so. But we know that such a feeling in that case is illusory; there, we know, events entirely prior to the moment the person concerned considers (falsely) to be his or her moment of choice are responsible for the particular card being picked. But we cannot in general take the fact that an experience of a certain sort can – in, note, contrived circumstances - be illusory as a reason to suppose that it is generally illusory. So the mere possibility that this impression of our providing an agent-causal oomph to the event-causal chain is illusory should not make us reflect on it as any less evidence that we do generally add agent-causal oomph where it seems to us that we do.

Allow me to imagine then your facing a choice between what strike you as two indifferent options. You ask for a pint of beer at the bar and are offered a choice by the barman of a straight glass or one with a handle. Nothing in your upbringing inclines you one way or the other; you have no cogent character trait or beliefs; in considering the reasons for and against each option, you quickly realise that you have no reason to choose either way. However, not wishing to be like Buridan's ass and not wishing to throw the decision back onto the barman (as you think that would simply delay matters unnecessarily), you say arbitrarily, 'A straight one, please'. In this case, whilst background factors will be relevant to explaining why it was that you faced the choice that you did – had you been brought up by members of the temperance movement, you might have been unlikely to have asked for this type of drink in the first place – nothing other than you explains why you asked for a straight glass rather than a handled one. It will of course still have been the case, during the moment or two that you took between being asked the question and answering it as you did, that mental events were going on in you – your thinking 'I don't want to be like Buridan's ass. I don't want any unnecessary delay in getting my beer.' So, even in a situation where an agent chooses between two options that are entirely indifferent to him or her, the sort of causal oomph provided by agent causation will occur alongside that provided by event causation. But we can, I suggest, disentangle them by introspection and we can see in the mental happening that is your deciding to ask for a straight glass (so as to move beyond a Buridan's-ass impasse and get

your drink as expeditiously as possible), the oomph provided by the agent entering the event-causal world. Well, that's what I say.

The person who maintains the view that I've been sketching maintains then (unless he or she holds that event causation may be reduced to substance causation) that there are two types of cause: agents and events. Agents and agents alone can cause actions. They characteristically do so because they believe themselves to have reasons and their beliefs that they have reasons are events happening in them, but the causal oomph that they provide transcends that of the events that they undergo, including that of the mental events which are their believing themselves to have whatever reasons they believe themselves to have. Agents produce events directly (at the minimum the events of their deciding to do whatever it is they decide to do under the description they decide to will it under), events which are not in fact solely caused (even though in all cases they will be partially caused) by preceding events.

This analysis carries with it then certain metaphysical commitments, commitments which some have considered incoherent. Specifically, through believing in the account, we are committed to believing that there is a certain type of existent – agents – the causal potency of which cannot be reduced to the causal potency of the events that this type of existent undergoes. If we do not think that event causation may be reduced to substance causation (and I have offered no argument that should lead us to think this), we will say then that the *relata* of the causal relationship – the values for *x* and *y* in propositions of the general form, '*x* causes *y*' – may be agents, which are a particular type of substance, as well as events, which are not. Whatever its extravagances when considered alongside a view which posits that only events may cause, this is not, it will be observed, to posit some fundamentally new sort of causation (the *relation* in '*x* causes *y*' stays the same whether the *x* is an event or an agent). All it posits is that agents – not merely events - have causal power. Agents can initiate causal chains, not merely be participants in them through being the things in which events occur; they can, as we have been putting it, add their own causal oomph to the world.⁴ (Of course the view commits one to the falsity of certain philosophical analyses of causation, those which would make it a priori that agents cannot cause.)

So, what sort of thing is this substance the causal efficacy of which cannot be reduced to the causal efficacy of the events that it undergoes? A traditional answer has been that it is a non-physical substance, a soul. But, whilst that may be the right

answer, it is not at all clear that the person who believes in agent-causation need believe in souls, as it is quite possible for someone to maintain that physical substances such as ordinary human beings have causal powers which are independent of the powers of the things out of which they are constituted. In short, souls would do the job, but are not needed to do it.

In order to make some general points in this area that will already be familiar to some here, let us imagine a particular pointillist painting. Obviously, at some level of description, the painting consists of a series of patches of colour – as of course do all paintings, but the pointillist style makes the fact more obvious than some. These patches of colour have various causal powers, e.g. the power to reflect light of certain wavelengths and absorb light of other wavelengths. The picture that emerges if one stands back and looks at the canvass from a distance has other causal powers; for example, it may have the power to make one realize that it depicts people relaxing on a riverbank. We may say of the higher-level entity, the picture, that it has causal powers which the lower-level entities that constitute it, the patches of colour, do not have. Similarly, we may suggest that humans *qua* agents are higher-level entities with causal powers that exceed the causal powers of the component molecules or what have you which constitute humans *qua* biological organisms. It must be conceded that there is a natural inclination to think that higher-level entities cannot have causal powers that are genuinely independent to any extent of, or in any manner exceed the aggregate effects of the causal powers of the lower-level entities that constitute them; they supervene in the sense that they couldn't have been different or indeed have been exercised differently without a change in the causal powers of the lower-level entities or their exercise. That is to say, if the painting has the power to cause the average viewer to realize that it is a depiction of people relaxing on a river-bank, then it only has it in virtue of the patches of colour having the powers to appear the colours that they do to the average viewer and if you had wanted to paint a different painting, there'd have been no way of doing it other than by putting the dots in different places. And this is a natural inclination which the believer in agent causation cannot give-in to in the case of agents (at least if he or she is to continue to locate the agent entirely in the natural world, rather than go down the soul route).

According to the agent causalist who locates agents in the natural world, when the particular sort of higher-level substance that is a human agent causes an effect, he or she does not have the causal efficacy that he or she does solely in virtue of the

casual efficacy of the lower-level entities that constitute him or her at that time; the agent himself or herself can cause something without the efficacy of his or her doing so simply being the efficacy of his or her parts doing whatever it is they are doing. His or her parts doing something may of course be allowed to be necessary for the agent's doing something; it just has to be the case that their doing whatever it is they are doing does not necessitate that the agent does whatever it is he or she does. It just has to be the case that, *qua* agent, he or she adds his or her own bit of causal oomph to that which is being provided by his or her body parts and events. Indeed, when it comes to the higher-level entity that is the human agent, that he or she have brain-cells and that his or her brain-cells be doing certain things is very plausibly causally necessary for him or her to make the choice that he or she does end up making (at least in the ordinary run of things – agents not having their minds 'uploaded' into computers and the like). The believer in agent causation is not committed to denying this (even the believer in souls is not committed to denying this); he or she is just committed to the claim that the causal powers of these lower-level entities do not in themselves exhaust or determine the causal powers of the higher-level entity, the agent, and how they are exercised; thus the space within which the agent *qua* agent makes a difference.

We are now in a position to see then – if we weren't already - that believers in agent causation within the physical world are committed to the actuality and irreducibility of 'top-down causation'; they are committed to the falsity of the claim that there can be no difference in a higher-level property without a difference in lower-level property through being committed to the claim that the higher-level is not in fact determined by the lower. We all believe in bottom-up causation; that's plausibly the story to tell about how we are caused to see a picture of people relaxing on a river-bank by the light reflected by the thousands of tiny dots of colour that constitute a painting. And we can probably be made to believe in top-down causation without much difficulty. That's the story to tell about how, having enjoyed looking at this particular painting, I end up buying a postcard of it from the gift shop on the way out, and thus why it is that the particles that constitute that particular postcard, rather than those that constitute some other, move with me out of the gallery when I leave. But *irreducible* top-down causation and the claim that there can be a difference in a higher-level property without its being determined by a difference in a lower are much more controversial. At this stage we cannot do more than point out that, however controversial they might be, there is nothing incomprehensible about them. Just as agent causation *per se* does not posit a new relation of causation, just that – in

addition to events (unless event causation can be reduced to substance causation) – substances can stand in for xs, the first term in the causal relation, so the person who posits irreducible and non-dependent top-down causation is not positing a new relation of causation either, just that the familiar everyday notion operates from the higher-level down in a way which is not just misleading shorthand for lower-level entities producing epiphenomenal higher-level effects; in fact there can be changes in higher-level properties which are not determined by changes in lower. The assertion that causation can go in this direction is no less comprehensible than the assertion that it cannot.

If we do wish to use souls as the point of origin for our agent-causal oomph, we may of course insert them at this stage non-problematically instead of irreducible top-down causation of the sort I've been discussing and, from then on, the story will be the same. Souls interacting in this way would be providing causal oomph that was quite compatible with the laws of nature (of course the physical universe would have not to be closed). So, we may say that, roughly, if you think it more likely that there would be souls than that there would be irreducible top-down causation of the sort discussed, you should conclude from the fact that we have free will that we probably have souls.

So, I hope this morning to have given you an idea of what I think it is that constitutes our being the authors of our actions and reasons to think the account at least coherent. I've argued that it is quite compatible with our being the authors of our actions that events outside our control causally incline us to do whatever it is we end up doing (or indeed incline us not to do whatever it is we end up doing); incline, but not necessitate. I've argued that it is not necessary that nothing other than oneself as agent would have proved causally sufficient to produce the particular effect that one produced with one's action had one not produced it oneself. (Randomness or brain-chip operators might have taken over had one not chosen to do a particular thing and resulted in one's body nevertheless moving just as it did one's having chosen to do that thing.) The nature of authorship then is not that one's choice is the only cause of what one ends up doing, nor that nothing other than oneself would have caused 'one' to do what it is that one ended up doing had one not caused it oneself, but just that in fact one *did* cause it oneself, one provided – as I have put it – a bit of causal oomph to the chain of events going on in and around one. This causal oomph - though its most immediate effect is an event (the event of one's choosing to perform whatever

action it is one chooses to perform under whatever description one chooses to perform it) – is caused by oneself *qua* agent, *not* caused solely by prior events within oneself. The causal potency of oneself *qua* agent cannot then be reduced to the causal potency of one's parts or of the events that one undergoes or one's parts undergo; thus we avoid the problems of the 'disappearing agent' and of 'luck'. One avoids the problems of the 'disappearing agent' and 'luck', but one does so at a price. Unless event causation may be reduced to substance causation, being an agent causalist commits us to a more pluralistic picture of the *relata* of causation – though importantly not the relation itself – than the person we may dub the 'pure event causalist'. Agents as well as events can be causes. Even so, this is no great reason to doubt that the view is right, especially when we reflect on situations which, it was argued, provide us with direct experience of our providing agent-causal oomph (most vividly situations where we move ourselves out of Buridan's ass impasses by acts of the will). Of course we agent causalists are also committed to positing that we have souls or to positing that even though we are at one level of description entirely physical and may thus be described in the terms of Physics, at another level we *qua* human agents gain causal powers which are not reducible to or determined by the causal powers of the parts that make us up or the events we or any of our parts undergo. This commitment however, as we also saw, is no great reason to doubt the view; again, it does not commit us to a new sort of causation, though it does entail (as would the soul view) the impossibility of 'unifying the Sciences'. It would be an implication of our being the sorts of things I've argued that we *qua* agents essentially are that the sciences dealing with higher-level agent-involving phenomena are not reducible to fundamental Physics. A quick inspection of Sociology, Economics, Politics, Anthropology and the like suggests that the subject matter of their discussion is not reducible to that of Physics. Of course someone who has hopes for total reductionism will be able to point to the fact that these are early days and say that the fact the project of unifying the sciences has not yet been accomplished is not conclusive reason to suppose that it will never be accomplished. And they are of course right when they say this, but whilst not being conclusive reason, it may yet be some reason to be sceptical of the chances of success for total reductionism. That it has not yet been accomplished is certainly no reason blithely to assume that it can be accomplished, especially given that assuming that it can be accomplished is incompatible with an assumption that is already and obviously known to be right, indeed, if what I have been saying is true, is something that we know with greater immediacy than almost anything else, that we are agents.⁵

18th May 2010

¹ In this case though, given that my intention was to tell the truth and that my failure to provide efficacious causal oomph is hardly my fault, so I would remain as praiseworthy as I would have done had I actually told the truth on the assumption that if the effects of someone's praiseworthy choice are supplanted in this way then their prior choice does not thereby lose any of its praiseworthiness.

² Of course if we talk of the outcome in terms of its being that particular agent's action, then his or her agency will be necessary for it; trivially, nothing and no-one other than myself can bring about the outcome of *my* doing such and such. But that is a mere terminological point. Something or someone other than myself could certainly, in principle (in practice the technology may not be available), bring about any outcome that doesn't have built into its description that I bring it about; they could do so by taking control of my body in the right way

³ So, on the view argued for, it is not that something or event (x) that causes me (y) to do something (z) is to be understood as causing an effect, me, who then in turn causes another effect - x causes y and y causes z. Rather, x causes merely an event in me, say the event of my vividly realizing myself to have very good reasons to do z, which inclines (but does not necessitate) me to cause z.

⁴ I view it in this regard as rather similar to Leslie's view that the good can cause; people who object to this view don't in my experience object to it because they regard it as incoherent, just false (albeit false of metaphysical necessity).

⁵ I am grateful for the comments of John Bishop, Peter Kail, and Kevin Timpe on this paper.