



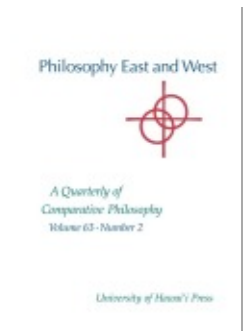
PROJECT MUSE®

Buddhist Paleocompatibilism

Mark Siderits

Philosophy East and West, Volume 63, Number 1, January 2013, pp.
73-87 (Article)

Published by University of Hawai'i Press
DOI: [10.1353/pew.2013.0005](https://doi.org/10.1353/pew.2013.0005)



➔ For additional information about this article

<http://muse.jhu.edu/journals/pew/summary/v063/63.1.siderits01.html>

BUDDHIST PALEOCOMPATIBILISM

Mark Siderits

Department of Philosophy, Seoul National University (emeritus)

/

My topic is a view that I call Buddhist Paleocompatibilism. This is something that I have talked and written about before,¹ and I should apologize in advance for the fact that I shall not have much to say about it that I have not previously said, in one way or another. What I hope to do here is explain the view more clearly than I have in the past. It is a view that I find interesting and perhaps plausible, and I think it serves as a good illustration of the sort of thing that can happen when we bring two distinct philosophical traditions into conversation with one another.

While I call the view Buddhist Paleocompatibilism, I should say right at the beginning that this is not a view that (as far as I know) any Indian Buddhist philosopher actually held. Since we now know of no classical Indian philosopher who claimed that determinism and moral responsibility are incompatible, no Buddhist philosopher would have felt compelled to defend the opposing compatibilist view.² Buddhists were forced to deal with a problem that is in some ways related to the problem of determinism and responsibility—the problem of reconciling karmic justice with the Buddhist doctrine of nonself. And it is their response to this problem that first led me to think that if they were faced with the problem of determinism and responsibility, paleocompatibilism might be their answer.³ So while the view has a Buddhist heritage, it is not strictly speaking a classical Indian Buddhist view. But neither is it a Western view. While the problem it addresses is modern and Western, some of the tools it employs are distinctively Buddhist. And therein, I think, lies much of its interest.

Paleocompatibilists make two key claims:

1. There is no sound argument for the incompatibility of determinism and moral responsibility.
2. In the absence of any compelling incompatibilist argument, we should accept the commonsense view that persons are generally morally responsible for their actions.

While (2) might seem uncontroversial, there are difficulties involved in spelling out what moral responsibility would look like given various possible ways of establishing the antecedent of (2). But I shall skip over those now and focus on the defense of (1). The strategy here uses the doctrine of two truths.⁴ That doctrine has it that there are two ways a statement may be said to be true: ultimately and conventionally. The defense of (1) will be that the form of determinism said by incompatibilists to be

incompatible with moral responsibility could only be true ultimately, while the claim that persons are morally responsible is true conventionally, and there can be no entailment relations across the barrier between ultimate and conventional truth, so nothing about responsibility could follow from the thesis of determinism. Any argument meant to show that the truth of determinism would undermine moral responsibility has to fail.

//

Let me begin to explain all this by saying more about the two truths. Ultimate truth is easy to explain. It is just what most people think of as truth: the property a statement has when it represents reality as being a certain way and reality mind-independently is that way. But the stipulation that correspondence be to how things mind-independently are means that many statements we take to be true are not ultimately true. Take the statement "I am in Sinyang Hall." Since buildings and other such composite entities are not mind-independently real, no statement about them could be ultimately true. That composite entities are not ultimately real is established by the "neither-identical-nor-distinct" argument: the building cannot be identical with the atoms, since it is one and they are many, but neither can it be distinct from them, since it cannot be wholly located where each atom is (it being too large), and the hypothesis that it is only partly located where each atom is leads to an infinite regress (since the building can only be partly located where each of its atom-sized spatial regions is, thus necessitating that it have parts, etc.).⁵ It is important to add that while "I am in Sinyang Hall" is not ultimately true, it is not ultimately false either. For, there being ultimately no Sinyang Hall, the statement that I am not located in it could not be ultimately true.

While the statement lacks ultimate truth-value, the statement is conventionally true. Explaining what this means is rather more difficult. How, one wants to know, could there be any other way for a statement to be true than by corresponding to mind-independent reality? Well, "Hamlet is a Danish prince" is true, we say, yet there is no one in the world to whom the name "Hamlet" refers. The statement cannot be really, literally true, yet we still think there is some sense in which it is true. That Hamlet was a Danish prince is, we say, "true in the story." By this we mean that if the statements that make up the story were fact and not fiction then this statement would be true in the literal sense. Conventional truth is a little like that. There really are no buildings, just atoms arranged in various ways. But given the way that the atoms around me are arranged, if in addition to atoms there were also such things as buildings, then "I am in Sinyang Hall" would be ultimately true. There aren't such things as buildings. We only think there are because it's useful for us to think of atoms arranged a certain way not simply as the many things they actually are, but as composing one big thing such as a building. If this pretense of ours reflected mind-independent reality, the statement would be ultimately true. Our pretense does not reflect reality; indeed, given the neither-identical-nor-distinct argument, it could not. But just as the statement about Hamlet does bear a certain complex relation to marks

on a page, so the statement about Sinyang Hall bears a complex relation to the atoms around us. That is why it turns out to be useful, and why we say it is true when strictly speaking it is not.

Now take the statement that Hamlet lives in Denmark. We know it is not literally true, but could it be “true in the story” or fictively true? Opinions about this vary. The difficulty is that we think Denmark is a real place while Hamlet is not a real person, and it is not clear how a real place could be related to an unreal person in the right way to make the story true. If you think that a real place can figure in the truth-maker for a statement that is only fictively true, then you will think the statement can be fictively true in a perfectly straightforward way. But if not, then you will need to hold that the true-in-the-story “Hamlet lives in Denmark” is not about the real Denmark but about some other sort of place entirely. Something similar happens with conventional truth. We agree that “I am in Sinyang Hall” is conventionally true, and we can also say that a statement like “There are such-and-such atoms arranged in such-and-such a way” is ultimately true. But what about “Sinyang Hall is made of such-and-such atoms”? Since it says something about Sinyang Hall, we know it is not ultimately true (or false either), but could it be conventionally true? Here, too, opinions will vary. The difficulty in this case is that if we can talk of Sinyang Hall and those atoms in the same breath, this will lead to contradictions. Perhaps the easiest way to see why is to think of the sorites puzzles that will emerge when we begin to look into how many randomly chosen atoms we can remove before Sinyang Hall ceases to exist.⁶ And there are other ways to show why permitting such talk leads to dire logical consequences.⁷ Still there are those who would say that there can be a conventionally true statement referring to both Sinyang Hall and the atoms.

Although Buddhist accounts of the two truths do not typically discuss this question, if they were asked then no doubt many Buddhists would say just this. But this is because these Buddhists believe that conventional truth harbors contradictions within itself, and so must be transcended. These Buddhists would not see a reason to try to insulate conventional truth from the contradictions that arise when you let in the referring expressions of ultimate truth. Other Buddhists, such as Dharmakīrti, would disagree. They would say that we cannot account for the usefulness of conventionally true statements if they inevitably give rise to contradictions. So the insulation between the two discourses must be two-way.

To summarize, there are two things we might say about the conventional truth-value of “Sinyang Hall is made of such-and-such atoms.” We might say that it is conventionally true. That policy will lead to there being good reason to accept various contradictions. We might welcome this result as showing the inherent instability of our ordinary ways of talking and thinking. On the other hand, we might take the resulting contradictions to show that conventional truth must be reformed. The proposed revision is that we not allow referring expressions from ultimate discourse to be employed in conventional discourse. Buddhist Reductionists might take either of these two stances. What I want to explore is what happens if we take the second, according to which that sentence has no truth-value, since it is simply meaningless.

Now if that is the view we take of the relation between the two truths, then it follows quite straightforwardly that no argument the conclusion of which belongs to conventional discourse can have a premise that belongs to ultimate discourse. No statement about Sinyang Hall follows from any statement about atoms. That will sound counterintuitive to many, but the Buddhist will not be put off by this. All Buddhists agree that common sense is profoundly mistaken about any number of important facts about the world, so the intuitions shaped by common sense cannot always be trusted. What Buddhists disagree about is whether it is possible to revise common sense in such a way as to rid it of its liability to lead to contradictions. The Buddhist revisionists I am discussing now think it is. That is the point of their two-way semantic insulation. If they can then show that the relevant formulation of determinism could only be ultimately true, while the claim that persons are morally responsible for some of their actions is conventionally true, then they will have a way of establishing (1).

III

This brings me to why I call the view paleocompatibilism. The strategy will be to show that moral responsibility is a property only persons could have, while the relevant form of determinism concerns the parts of persons, such as volitions. Now the first modern Western compatibilists, Locke and Rousseau, both answered the incompatibilists by saying that one should ask not whether the will is free but whether the person is free to do their will. Since by freedom they meant something that is necessary for moral responsibility, they thus held that moral responsibility pertains to the person as a whole and not to one of its parts. And this much seems right. It is persons that we hold accountable for actions; it is persons that are the objects of our praise and blame, gratitude and resentment—not hands, feet, hearts, lungs, brains. Of course the incompatibilists to whom Locke and Rousseau were responding were dualists who thought of the will as a faculty of the mind. And faculties, it can be said, are not parts of the substances in which they inhere; they are powers or abilities of those substances. So if it is a mistake to ask if the will is free, it is not the mistake involved in confusing a whole and one of its parts. While a present-day incompatibilist might not want to affirm that the mind is a substance distinct from the body, they might still claim that the will is not a part of the person but a faculty of the person. But now we must ask what it is about the exercise of this faculty of willing that leads to the thought that the person whose willing this is is not responsible. Answering this question will take us back to wholes and parts.

The thought that if determinism holds universally then the will is not free stems, I submit, from the idea that the event of willing to do A is caused not by the person but by some prior event (which is in turn caused by some prior event, etc., back to a time before the person was on the scene). The kind of causation at work here is event causation: one event causing another event. We also think of persons as causes. But persons are not events; they are substances. When we say that a person caused some happening—say that Kim caused the death of Lee—we think of the effect as an event (the demise of poor Lee) but the cause not as another event (the act of pulling the

trigger, or the act of intending to pull the trigger) but as the thing behind that act, the agent Kim. Agency is a variety of substance causation. And substance causation is not a kind of causation that operates alongside or in competition with event causation. It represents an altogether different model of how the world behaves. Incompatibilist arguments exploit the fact that the two models do not work in tandem. The paleocompatibilist explains this disharmony by claiming that substance causation concerns wholes while event causation involves parts.

The mereological reductionism at the heart of Buddhist Reductionism has no substances in its final ontology. The atoms that are all there really is to what we think of as Sinyang Hall are not indivisible material particles. A particle is a substance, an enduring thing that bears various qualities (such as having a certain mass and a certain charge). And the same argument that shows Sinyang Hall not to be ultimately real likewise shows that no substance, no matter how small and simple in structure, could be ultimately real. The ultimately real atoms are tropes, particular momentary occurrences of qualities such as resistance and heat.⁸ A material particle is a bundle of tropes. And of course it is we who do the bundling, for our own convenience; that is why material particles are not ultimately real. As for the substance known as mind, Buddhist Reductionists analyze that into mental tropes: occurrences of such mental events as feelings of pleasure and pain, desirings, willings, and cognizing. The belief that such events must have an owner arises from the fact that we have conceptually constructed a substance, the mind, as what holds a bundle of such tropes together. Descartes was wrong: from the occurrence of this cognizing event it does not follow that there is an "I" that is cognizing. That this cognizing event occurs is ultimately true. That I exist is conventionally true. There are no entailment relations between the two discourses.

There is more that could be said about the *cogito*, but our present topic is such mental events as desirings and intendings. Incompatibilist arguments typically invite us to think that if determinism is true, then it must hold for events like these. And then a sort of vertigo sets in. We know how to answer questions about the cause of an action, such as the question of who caused Lee's death. Suppose we agree it was Kim. Presumably this means Kim wanted Lee to die and formed the intention to bring about Lee's death. But when we go on to ask about the cause of the intention, the question becomes somewhat puzzling. Are intendings things that I do in the same sense in which I might do the killing of Lee? I know what it is to want someone dead, and to form an intention to bring this about. But how do I cause it to be the case that I intend Lee to die? By intending that I intend for Lee to die? How do I bring it about that this occurs? It begins to look as if when we ask about the cause of the volition or the intention, only another event will serve, and not the substance that is the person. Buddhist Reductionism explains why this should be. Mental events like desires and intentions are among the parts that make up the whole known as a person. This means that such events cannot be said to have persons as their causes. A mere conceptual fiction cannot be the cause of something ultimately real. The causes of such events must instead be impersonal events. So the incompatibilist turns out to be right after all in saying that if the psychological factors involved in performing an action

are causally determined by prior events, then the springs of action can always be traced back to thoroughly impersonal forces. What the incompatibilist neglects to mention is that this is because when we speak of these mental events and their causes, we can only be speaking about ultimate truth, which is by nature necessarily impersonal. Nothing whatever follows about persons.

IV

In defending his agent-causal version of incompatibilism against a Donald Davidson-style objection, Timothy O'Connor (2003) confronts the difficulty that arises when we try to combine our talk of agents with what we say about events internal to the agent. It might prove useful to consider how he handles this, for it may throw some light on how paleocompatibilism both accommodates certain strands of the agent-causation approach and rejects its incompatibilism. Agent-causation theories typically claim that while actions, as events, have causes, the cause of a basic action is not another event but rather an agent, an enduring substance. In O'Connor's account causation is an irreducibly primitive production relation. Causal powers are the properties of particulars, and causation is the manifestation of these powers. There are, then, two kinds of causation: event causation and agent causation. Event causation may be thought of as the manifestation of the particular's having its nature in certain specific circumstances (one event) through the production of the effect (another event). Agent causation is the manifestation of the agent's agent-causal power in conferring on the agent the ability to produce at will some effect or other from a range of options. Only particulars that are in certain respects self-determining have agent-causal powers. Thus a physicalist version of the theory will require that such powers be emergent properties of complex systems with the capacities of representing possible courses of action to themselves and of having beliefs and desires concerning those courses of action (O'Connor 2003, p. 262).

In O'Connor's account, the basic actions produced by agents are initiations of actions through the formation of intentions. This means that I am the agent of the action of killing Lee by virtue of my producing the requisite intention, which is the event through which my agency flows to the successive events in the series. Intentions are indeed things that I do in this view. The Davidson-style objection, tailored to this formulation of the agent-causal theory, concerns the question of what is the relation between agent and action, the action here understood as the primitive action of forming the determinate intention that will in turn initiate the freely willed action (e.g., the arising of the executive intention to kill Lee, from which the grasping and firing of the gun flow). Specifically, is this relation itself an event or not? If so then either this event is an action or it is not. If it is an action, then the supposedly primitive action that is its result is not primitive, and we have the start of an infinite regress. If this event is not an action, then there is a causing by the agent that is not a doing by the agent. If on the other hand this relation is not an event, then since a causing is an event, it is no longer clear what it means to say that the agent caused the action.

O'Connor responds by grasping the second horn of Davidson's dilemma: the relation of the agent to the action is not a distinct event, but is instead identical with the action, understood as the agent's production of the intention. The relation—the agent's causing the action—is a complex state of affairs, consisting of the agent's being the agent-cause of the action. And this is not the sort of thing, O'Connor claims, that could itself be caused (2003, pp. 271–272). The threatened infinite regress is avoided by in effect making the producing relation identical with one of its relata.

The Nyāya school of classical Indian philosophy developed a similar strategy as a way to avoid various Bradley-style regresses. Included in their ontology are such things as universals, substances, and a relation of inherence that connects potness to particular pots. Asked how inherence is in turn related to particular pots, they answer that there is no distinct relation between inherence and a particular, that inherence is a self-linking connector, so that the relation between inherence and a pot is just the inherence itself. (This seems to be what Bertrand Russell had in mind when he sought to stop the original Bradley regress by claiming that relations are not at all like the substances that they relate [Russell 1927, p. 263].) Nyāya likewise holds that there are such things as absences, for instance the absence of a particular pot from this table. But absences, Naiyayikas claim, only exist when there is an existing counter-positive: there can be the absence of the pot from this table only because that pot exists elsewhere (or elsewhen). Asked about the relation between the absence and its counter-positive, their reply is that this relation just is the absence itself, which is a self-linking connector. O'Connor's reply to the question about the causal relation between the agent and the action is similar: by in effect making the action a self-linking connector, he hopes thereby to avoid a Bradley regress.

O'Connor's strategy has a distinguished pedigree, but can it succeed? The notion of a self-linking connector has gotten mixed reviews. The reception is somewhat warmer where we are antecedently inclined to say that the connection in question is across different ontological categories. To suppose, for instance, that some extra tie is needed to connect inherence with substances is to think of inherence as substance-like in being just another particular without intrinsic relations to other particulars. If we instead think of inherence as belonging to a category distinct from that of substance (as does Nyāya), then we may be more inclined to accept the claim that it is a self-linking connector. Those who wish for a more austere ontology containing fewer categories will resist such claims. Buddhist Reductionism represents an extreme form of such austerity, maintaining that we can make do with just the category of momentary trope-occurrences, all else being conceptually constructed on the basis of these elementary events. If so, then the notion of a self-linking connector will turn out to be both incoherent and superfluous. But assessing the plausibility of so Spartan an ontology as that of Buddhist Reductionism is no easy task.

It might, however, be worth looking into the motivation behind this resort to self-linking connectors. For if it should turn out that the problem it is meant to address can be resolved in a simpler way, that will be one consideration against it. We started with the difficulty that arises when we "pop the hood" on my killing Lee in order to ascertain whether I am responsible for this regrettable event. Presumably what we

find when we look under the hood will tell us whether this is genuinely my act or not. The stock answer is that it is my act provided we find that it flowed from an intention wholly my own. How does an intention come to be wholly my own? If we say that it must reflect my character, then we may ask why we should take the relevant elements of my current character as truly my own. If they arose as the result of earlier elements identified as expressive of my nature, we may ask the same question of those earlier elements. Eventually this quest will lead us back to a time before I was born. It is this prospect that leads the agent-causation theorist to claim I can be responsible for the action only if the event of the intention's occurrence is the effect not of a prior event but of an agent, me.

That this line of thought is on the right track might be confirmed by considering the case of the Swampman double who, when the lightning strikes the muck, arises as a fortuitous molecule-for-molecule duplicate of me as I am just before forming the requisite intention.⁹ Should Swampman last long enough to then form the intention and shoot Lee (who happens to be in the swamp that day) before dissolving back into the primordial ooze, there would be no judgment of responsibility. Such judgments require an enduring agent, and Swampman does not endure long enough to count. This, says the agent-causal theorist, is why responsibility drains away when we pop the hood and look for some inner event that might explain the causal chain leading up to the action. As long as our explanations are couched exclusively in terms of event causation, we will never find an enduring agent. This is why we must turn to causation of an altogether different sort.

The paleocompatibilist will agree with this last point, but draw a different conclusion: that agent causation and event causation belong to the distinct discourses whereby we speak of wholes and of their parts respectively. Given the semantic restrictions on these discourses, there can be no problem of explaining how an enduring substance can serve as cause of an event's occurrence at one time rather than another. And it is this dating problem that lies at the heart of the present difficulty. The agent-causal theorist's explanation fails to satisfy because it tries to account for the timing of the formation of the intention in terms of the agent's being in certain circumstances just before the intention was produced. In this case it looks like we should say that it is those circumstances that are the cause, with the result that the agent once again drops out of the picture. As long as agent causation and event causation are thought of as competing ways of explaining the occurrence of an event, event causation will always win out given its ability to solve the dating problem.

The alternative picture offered by the Buddhist Reductionist places event causation at the ultimate level and substance causation at the conventional level. There being no substances at the ultimate level, event causation consists in the relation of universal concomitance and ordered succession between elementary event-types. Substance causation, the sort of causal relation asserted at the conventional level, involves the manifestation of a substance's powers in the production of events (understood as the arising of new properties in this or another substance). When the substance in question is an agent, we say this manifestation can occur "at will." This should not be taken to mean that the agent produces a willing that in turn brings

about the initiation of the action. It means instead that it is up to the agent to bring about the action, that the agent produces the action when they choose to—they “just do it” when they see fit. To seek an explanation of the timing here is to pop the hood, to treat the agent as a system analyzable into components whose properties do the real explanatory work.

To those accustomed to the sorts of causal explanations found in science, it may sound odd to claim that the concept of causation at use in conventional discourse is that of substance causation. But this is not really all that untoward a view. If we follow the lead of Locke (*Essay* I.xxi.4), we should say that the child’s concept of causation begins with its learning to exercise control over its limbs. Seen in this light, it is hardly surprising that the child should attribute agency to inanimate objects. And even when we come to realize that such objects lack the representational and deliberative capacities necessary for agency, we continue to speak of them as the causes of various changes in the world. Just as Kim may be said to be the cause of Lee’s death, so cars, or power plants, or even trees (in Ronald Reagan’s view) may be said to be the cause of the polluted state of today’s air. While current common sense also allows talk of event causation (understood as strictly a relation between events), we may speculate that this has come about under pressure from the ascendancy of scientific explanation. We might even claim that the concept of substance causation is the more primitive concept not just ontogenetically but phylogenetically as well—that it is what John Perry (2010, p. 94) calls a *natural* concept.

V

If determinism is true, then all mental events are causally determined by other events. We have seen why when this thesis is properly understood it turns out to be ultimately true. We can also see why the claim that persons are sometimes morally responsible for their actions could only be conventionally true. Given the semantics of the revisionist form of Buddhist Reductionism I am discussing, no conventionally true statement is entailed by any ultimately true statement. This completes the defense of (1): nothing about moral responsibility can follow from the thesis that all mental events are causally determined by other events. But it might be useful to illustrate (1) in action by looking at what it tells us about Benjamin Libet’s experimental results. In a series of experiments Libet purportedly showed that when subjects are instructed to spontaneously initiate hand movement on the basis of a sudden urge (i.e., an intention formed at the very last minute), the neural events that begin the process of movement occur before the subject becomes aware of intending to move their hand.¹⁰ This is sometimes taken to show that the event of consciously deciding (the formation of a conscious intention) cannot have caused the initiation of movement. Critics contend that the experimental design is flawed, and that the data do not support this conclusion. But let us waive such questions. Suppose we agree that the subjects reliably report the time when they become aware of intending to move their hand, and that in this case the intention must be a conscious event. Since the initiation of movement precedes the conscious intention, the latter cannot have caused the former. (Of

course if we follow Daniel Dennett in thinking of consciousness as a property that mental events only acquire retroactively,¹¹ this conclusion does not follow; but in his view nothing is intrinsically a conscious mental event, so nothing can be a real cause by virtue of being a conscious state.)

The real question, though, is why this should be thought to matter. People often find these results spooky. And the spookiness is sometimes said to consist in this being empirical evidence that there is no free will. What people probably mean by this is that in the experiment the subjects were not responsible for moving their hands, that the movement was not really something the subjects did. The question is why anyone would interpret the results of the experiment in this way. The subjects did, after all, do what they were asked to. They produced a series of hand movements in a relatively spontaneous way: in such a way that an observer would not have been able to reliably predict just when the next movement was likely to occur from anything coming before. Of course the subject is not aware of having decided to initiate movement until after the initiation process has begun. But this does not show that the subject did not initiate the movement. That would follow only if being the agent of the movement requires that one be the agent of the intending. And it might be a mistake to think of intending in this way, as an action performed by an agent.

It is this conception of intending that seems to stand behind the thought that intentions must be formed consciously. The idea seems to be that persons come to have intentions by engaging in a process of deliberation about their situation and their standing desires and then choosing from among the range of available intentions the one that best matches the outcome of the deliberation process. Springs of action not formed in this way are not truly chosen by the agent, and so not genuine intentions of the agent. What else could explain the requirement that the formation of the intention be a conscious process, if not the idea that this must be under the control of the agent, and that control requires consciousness? Surely not the evidence of introspection. We are often hard-pressed to say, even with our most weighty actions, just when we decided to do this rather than that. The descriptions we give of the moment of choice are often no more than retrospective reconstructions. That this way of thinking about intending might be mistaken is also suggested by the fact that it introduces yet another action of the agent into the account, the choosing that yields the intention. As we saw earlier, an infinite regress threatens. We avoid this regress by keeping separate our talk of the agent's responsibility and our talk of the occurrence of the intention. The first concerns a big thing, the second concerns a small thing. Talk of big things and talk of small things do not mix.

Shaun Gallagher (2006) has a somewhat similar view of the controversy about Libet's results. He also thinks we are looking in the wrong place for "free will" when we look at the neurological precursors to the initiation of hand movement. The notion, he says, "applies to intentional actions themselves, described at the highest pragmatic level of description" (p. 117). Neurological accounts of motor control are simply at a level of description ("the subpersonal level") too low to be relevant to the

question of responsibility, just as a detailed account of the workings of the parts making up your car is not relevant to the question whether I should ride in it (p. 115). Instead the question of “free will” or responsibility only arises at the level at which it makes sense to speak of persons and the things they think of themselves as achieving through their intentional acts: doing the laundry, killing Lee, complying with the instructions in the psychology experiment.

Gallagher also holds that responsibility requires that intention be formed through a conscious process. Although this may hold for the sorts of actions that we take as paradigms of “free will,” I am not sure that it holds across the board. But what he says in defense of his claim is important. This is that the time frame that is required to see the sorts of intendings involved in responsible agency is different from the one used in examining processes of motor control. The time between the neural events that initiate hand movement and the occurrence of the movement is measured in hundreds of milliseconds. With the action of killing Lee the intending can count as such only by being placed within a context of awareness of desired outcome (that Lee be dead) and opportunity (there is a gun before me) through initiation of movement (grasping the gun and pulling the trigger) through monitoring the movements (ascertaining that the gun is grasped by the handle and not by the barrel) to awareness of the goal’s having been achieved (seeing the falling of Lee’s body as Lee’s death). The intending involved here is what it is only by being situated within a temporally extended process; it is not the sort of simple neural event that might count as what initiates the process of muscle contraction.

The Buddhist Reductionist will agree and go on to claim that as a temporally extended event the intending involved in the killing of Lee requires an enduring person as its subject.¹² The simple, short-lived mental event that initiates the process of muscle contraction requires no subject. By the mereological reductionism at the heart of Buddhist Reductionism, simple momentary events are themselves ultimately real. Temporally extended processes, on the other hand, can only be construed as properties of conventionally real substances. The intention that Libet studied is the wrong sort of thing to be involved in determinations of responsibility.

VI

So much, then, for the defense of (1). I said earlier that there might be difficulties in connection with the second paleocompatibilist thesis:

2. In the absence of any compelling incompatibilist argument, we should accept the commonsense view that persons are generally morally responsible for their actions.

The difficulties I had in mind stem from the way in which the paleocompatibilist argues for (1) and thus tries to establish that there is no compelling argument for incompatibilism. The belief that persons are generally morally responsible for their

actions is surely a part of common sense, something most people believe unreflectively. The Buddhist Reductionist view that underlies the defense of (1) is, however, profoundly revisionist. The commonsense view has it, for instance, that persons really exist, whereas Buddhist Reductionism denies this. This raises the question of whether the paleocompatibilist is actually affirming the commonsense view that persons are sometimes deserving of moral praise and blame. One way of bringing out the difficulty here is to ask the Buddhist Reductionist why it is that if, strictly speaking, there are no persons it should be useful to hold that there are and attribute moral responsibility to them. The response will be consequentialist in nature: our institutions of moral praise and blame help maximize overall utility, and these institutions require that we think of ourselves as persons. The commonsense view of deserts is not, however, consequentialist. This can be seen, for instance, in the strongly retributivist intuitions most people have about punishment. So the responsibility and desert delivered by paleocompatibilism are importantly different from those of common sense.

This is the familiar complaint about compatibilisms of every sort—that they deliver something other than what was promised. (Hence Kant's complaint that compatibilism is "a wretched subterfuge.") Neocompatibilists are often quite open about this. They claim that the notion of desert behind our commonsense view of responsibility is simply incoherent, and must be replaced by something saner. Here we have a situation that is common in philosophy: we find ourselves with conflicting intuitions and must somehow or other make the best of a bad lot. But perhaps the paleocompatibilist can do slightly better than the neocompatibilist. Many people have the intuition that the person who does an evil deed thereby acquires the very real property of deserving blame and (other forms of) punishment. The neocompatibilist must say that this is simply a mistake. The paleocompatibilist has a different response: what common sense says is true. Of course, in their eyes it is only conventionally true. Still if we are to speak of persons at all then this is what we must say. Moreover, this real desert could not be said to exist if we were to say that the psychological events involved in the production of an action are causally determined by prior events. The incompatibilist is right about this as well. The mistake lies not in our intuitions but in what the incompatibilist makes of them. The intuitions behind our views about responsibility uphold judgments that are true conventionally, while the intuitions behind our views about determinism support claims that are true ultimately, so there can be no conflict here. The responsibility on offer is that of common sense. The conflict lies not in our intuitions but in what philosophers make of them.

So where does this leave us? There is some reason to believe that (1) and (2) are true. If they are, then it can be the case both that persons are sometimes morally responsible for their actions and that determinism is true. We have seen how that could be. I shall end by repeating what I said at the beginning, that I think this view is interesting and might even be true. Much more work is needed to figure out if it is true. And that work requires the active participation of scholars of Asian philosophy.

Notes

- 1 – Most recently in Siderits 2008.
- 2 – The question whether determinism and moral responsibility are compatible has of course been the subject of much recent debate. The question is this: supposing that all events, including those psychological events involved in human deliberation, decision, and action, are caused by earlier events in accordance with strict (non-stochastic) causal laws, would it then follow that no one is ever justifiably subject to moral praise or blame for their actions? This is widely called the issue of “free will,” but because many neocompatibilists find this term problematic, I shall employ the longer and perhaps more accurate label.
- 3 – This was the train of thought behind Siderits 1986, though I did not then call the view “paleocompatibilism.” There is, however, one passage in the *Nikāyas* that I believe can be read as at least hinting at something like this view. At *Aṅguttara* i.173 ff. (*Mahāvagga, Tiṭṭhāyatana Sutta*) the Buddha discusses three views that he claims lead away from the practice leading to liberation: that everything a person experiences is the result of the person’s past actions, is the result of God, is uncaused. These views are all said to lead to a state of soteriological paralysis. In the first two views, the present “I” is a mere conduit of causation, channeling decisions made earlier, over which I have no present control. In the third view, nothing I presently decide can be counted on to have any effect on the world. In all three cases I am left an impotent observer of the events making up my life, unresponsive to thoughts concerning what should and should not be done. Both determinism and the view that nothing is caused are equally corrosive of prudential responsibility. One might then expect the Buddha to seek wiggle room in an indeterminism lying between determinism and utter causelessness, one holding that certain human choosings are not determined by prior causes. But the “middle path” he teaches on this occasion does no such thing. Instead he explicates the formula of dependent origination according to which suffering arises in dependence on a series of psychophysical events beginning with ignorance. In other contexts in which the Buddha invokes dependent origination as a “middle path” in order to dissolve an apparent dilemma, Abhidharma exegetes took him to be utilizing the distinction between the two truths. If we were to read the *Tiṭṭhāyatana Sutta* in that way, the Buddha would be a paleocompatibilist.
- 4 – The distinction between two kinds of truth, conventional and ultimate, was first developed by Abhidharma philosophers. It is central to the view of persons I call Buddhist Reductionism. Not all Buddhists accept Buddhist Reductionism. And not all Buddhist Reductionists accept the formulation of the distinction between the two truths that paleocompatibilism relies on. Many Buddhist philosophers do, however, draw the distinction in the way I am about to describe, for example Vasubandhu, Buddhaghosa, Saṃghabhadra, Dignāga, Dharmakīrti, and perhaps even some Pudgalavādins.

- 5 – The argument is more fully described at Siderits 2007, pp. 104–111.
- 6 – See Unger 1979 for the classic formulation of these difficulties.
- 7 – For evidence that Buddhist Reductionists were aware of the sorts of semantic difficulties that arise when the contentious relatives—the referring expressions of the two discourses—are not kept apart, see Mikogami 1979, pp. 83–86. For discussion of some further considerations behind making the semantic insulation between the two truths two-way see Siderits 2009.
- 8 – Support for the claim that the *dharma*s, the ultimate reals of Buddhist Reductionism, are tropes can be found in Ganeri 2001, pp. 101–102; also see Goodman 2004.
- 9 – For the origins of the Swampman-style thought experiment see Davidson 1987.
- 10 – A useful discussion is Pockett 2009, pp. 15–19.
- 11 – For an overview see Dennett 2005, pp. 131–157.
- 12 – See Metzinger 2003, pp. 422–426, for a useful account of the cognitive architecture at work in the self-ownership of actions—what is required for agents to view the action as their own and thus as something for which they can be held accountable.

References

- Davidson, Donald. 1987. “Knowing One’s Own Mind.” *Proceedings and Addresses of the American Philosophical Association* 60:441–458.
- Dennett, Daniel. 2005. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. Cambridge, MA: MIT Press.
- Gallagher, Shaun. 2006. “Where’s the Action? Epiphenomenalism and the Problem of Free Will.” In *Does Consciousness Cause Behavior?*, edited by Susan Pockett, William Banks, and Shaun Gallagher, pp. 109–124. Cambridge MA: MIT Press.
- Ganeri, Jonardon. 2001. *Philosophy in Classical India*. New York: Routledge.
- Goodman, Charles. 2004. “The *Treasury of Metaphysics* and the Physical World.” *Philosophical Quarterly* 54:389–401.
- Metzinger, Thomas. 2003. *Being No One*. Cambridge, MA: MIT Press.
- Mikogami Esho. 1979. “Some Remarks on the Concept of *Arthakriyā*.” *Journal of Indian Philosophy* 7, no. 1:79–94.
- Nichols, Shaun, and Joshua Knobe. 2007. “Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions.” *Noûs* 41, no. 4:663–685.
- O’Connor, Timothy. 2003. “Agent Causation.” Originally published in *Agents, Causes and Events: Essays on Indeterminism and Free Will*, edited by Timothy O’Connor.

Oxford: Oxford University Press, 1995. Reprinted in *Free Will*, 2nd ed., edited by Gary Watson, pp. 257–284. Oxford: Oxford University Press, 2003.

Perry, John. 2010. "Wretched Subterfuge: A Defense of the Compatibilism of Freedom and Natural Causation." *Proceedings and Addresses of the American Philosophical Association* 84, no. 2: 93–113.

Pockett, Susan. 2009. "The Neuroscience of Movement." In *Does Consciousness Cause Behavior?*, edited by Susan Pockett, William Banks, and Shaun Gallagher, pp. 9–24. Cambridge MA: MIT Press.

Russell, Bertrand. 1927. *An Outline of Philosophy*. London: Allen and Unwin.

Siderits, Mark. 1986. "Beyond Compatibilism: A Buddhist Approach to Freedom and Determinism." *American Philosophical Quarterly* 24: 149–159.

———. 2007. *Buddhism As Philosophy*. Aldershot, UK: Ashgate.

———. 2008. "Paleo-Compatibilism and Buddhist Reductionism." *Sophia* 47: 29–42.

———. 2009. "Is Reductionism Expressible?" In *Pointing at the Moon: Buddhism, Logic, Analytic Philosophy*, edited by Mario D'Amato, Jay L. Garfield, and Tom J. F. Tillemans, pp. 57–69. New York: Oxford University Press.

Unger, Peter. 1979. "There Are No Ordinary Things." *Synthese* 41: 117–154.