

Stephen Law

Honderich and the Curse of Epiphenomenalism

In ‘Radical Externalism’¹ Ted Honderich offers an ingenious and radical new solution to the problem of consciousness — a solution that promises, among other things, to do justice to two important features of consciousness — to both its subjectivity *and* its causal efficacy.

According to Honderich, the main alternatives to his own Radical Externalism are certain forms of dualism, or, as he puts it, ‘spiritualism’, and ‘devout physicalism’. Honderich’s central argument for Radical Externalism is that it succeeds in respecting those features of consciousness to which these two main alternatives fail to do justice. It is, therefore, the superior theory.

But is Radical Externalism superior? Does it have this advantage over its two main rivals?

I don’t believe it does. The central argument of this paper is that Radical Externalism falls foul of much the same kinds of problems concerning causal interaction that plague spiritualism. Indeed, ironically, it turns out that Radical Externalism is vulnerable to a similar objection to that which Honderich himself cleverly levelled against Anomalous Monism almost a quarter century ago.

But before we get to that objection, let’s begin by briefly outlining what Honderich takes to be the two main alternatives to his own theory — spiritualism and devout materialism — and examining their alleged failings.

Devout Physicalism and the Problem of Subjectivity

By devout materialism Honderich means:

[1] Honderich (2006). In this commentary all page references are to this target paper unless stated otherwise.

the belief or perhaps attitude that our consciousness is a fact, property or state of affairs that involves only physical properties ... and in particular properties ... existing and more or less anticipated [by] neuroscience (pp. 10–11).

The devout materialist either identifies mental properties with physical properties, or else eliminates them altogether (as in eliminative materialism).

So what's wrong with devout physicalism? Honderich maintains it fails, among other things, to do justice to the *subjectivity* of consciousness. We're all familiar with the kind of thought experiments involving black and white rooms, fool's pain, homunculi-headed robots and so on to which Honderich is perhaps alluding when he says

[n]o doubt theories are sometimes destroyed by single counter-examples clearly seen, or by crucial experiments well-conducted. It is my inclination to think this of devout physicalism ... (p. 11).

It's certainly a perennial complaint that these various kinds of materialisms somehow fail fully to allow for the subjective quality of conscious experience.

Spiritualism and the Problem of Causal Interaction

Which brings us to what Honderich believes is the other main alternative to his own theory — spiritualism. What characterizes the dualist or spiritualist position, he says, is a commitment to the non-spatiality of consciousness. Spiritualism, says Honderich, is

the theory, rightly associated with Descartes, that your consciousness is somehow non-spatial and hence not physical. It is in fact only misleadingly called dualism, mainly because its distinctive nature and its problems are not owed to its asserting that consciousness is other than physical but rather to its asserting that consciousness is out of space and in fact of a mysterious nature (p. 11).

The problems that plague spiritualism, says Honderich, include the problem of causal interaction. Descartes famously attempts to place the locus of interaction between the mental and the physical realm in the pineal gland. But of course it remains blankly mysterious how this interaction might take place. How can something that is not anywhere causally impinge upon a spatio-temporally extended, physical object?

And so it seems that spiritualism also fails a key test of adequacy, in this case, the requirement that any adequate theory of consciousness

must not make impossible what is actual, which is causal interaction between consciousness and the physical (p. 12).

So we appear to face an intractable dilemma so far as consciousness is concerned. We can either favour some form of devout physicalism, but then we fail to do justice to the subjectivity of consciousness. Or we can embrace some variety of spiritualism, in which case we run into the problem of causal interaction — indeed, we may find ourselves unable to prevent a slide into epiphenomenalism (or occasionalism or pre-established harmony theory, or whatever).

I'm sure many philosophers of mind would acknowledge that we do at least face *something* like the dilemma that Honderich presents us with. Indeed, finding a way out of this sort of dilemma has surely been one of the main preoccupations of philosophers of mind for the last couple of decades.

The question is: does Honderich's new alternative — his third way, as it were — actually allow us to resolve the dilemma? Does it really let us do justice to both of these features of consciousness: to both its subjectivity *and* its causal efficacy?

It's immediately obvious that we might attempt to mount an attack on Radical Externalism from at least one of two directions. First, we might argue that, actually, like devout materialism, Radical Externalism fails to do justice to the subjectivity of consciousness. Or we might try to show that, like spiritualism, Radical Externalism faces problems in allowing for causal interaction. Or we might do both these things.

While I suspect Radical Externalism probably does face problems with respect to subjectivity (see the end of this paper), *it's on problems with causal interaction that I'll focus here*. As Honderich's case for Radical Externalism is that it avoids those problems that respectively plague spiritualism and devout materialism, establishing that Radical Externalism does indeed run straight into much the same old problems regarding causal interaction as does spiritualism would suffice seriously to undercut his case.

Radical Externalism

Let me briefly sketch out both what I take Radical Externalism to be, and how I believe Honderich supposes it allows us to resolve the dilemma outlined above.

At the heart of Radical Externalism lies something Honderich calls *worlds of perceptual consciousness*. One important feature of a world of perceptual consciousness is that it is not located 'in the head' (not *at all*, in fact, unless e.g. you happen to be looking at your own brain). It

encompasses the same tracts of space that those objects of which you are perceptually aware occupy. Hence the ‘Radical Externalism’.

And yet, despite being spatially extended, a world of perceptual consciousness is also supposed to be subjective.

In what sense subjective? Well, a world of perceptual consciousness is something you *have*. Our worlds of perceptual consciousness are numerically distinct. You have yours. I have mine. When you and I simultaneously look at an orange placed on a table in front of us, our worlds of perceptual consciousness may spatially overlap. But still, while we may be conscious of the same orange, there are nevertheless two worlds of perceptual consciousness involved here, not one. Each subject has *their own* world of perceptual consciousness.

A corollary of this is that a world of perceptual consciousness is only immediately accessible to its owner. A world of perceptual consciousness is, in this sense, a *private* world. This kind of privacy is of course commonly supposed to be one of the hallmarks of the subjective (yet note that we are still dealing with a radical form of externalism here — we’re not talking about inner Cartesian theatres, or anything like that).

There’s at least one further way in which a world of perceptual consciousness would seem to qualify as subjective. A world of perceptual consciousness is subjective in that

[w]ith consciousness, what there seems to be is what there is. What there seems to be is all there is (p. 5).

This, I take it, is the familiar claim that while I might be mistaken about there actually being an orange on the table in front of me, I can’t be mistaken about the fact that that is how things subjectively *seem* to me. Within a world of perceptual consciousness, appearance is king. You can’t, as it were, have *fool’s x*, where *x* is something that features within your world of perceptual consciousness.

Resolving the Dilemma

How, then, does the introduction of worlds of perceptual consciousness allow us to resolve the dilemma sketched out earlier? How does it succeed in doing justice both to the subjectivity of consciousness and its causal efficacy?

Let’s start with subjectivity. That Honderich’s worlds of perceptual consciousness do at least do justice to the subjectivity of consciousness might seem obvious. They make perceptual consciousness something you *have*, and they also respect the fact that the illusion/reality

distinction collapses when it comes to conscious experience. They are also private. So perhaps the subjectivity of consciousness is indeed taken care of.

But what of causal interaction between the mental and the physical? How do they allow for that? Well, remember that, according to Honderich, the problem spiritualism faces concerning causal interaction is simply this: how can something that is *non-spatial* causal impinge upon the physical? Honderich's worlds of perceptual experience do indeed appear to sidestep this problem, for the simple reason that they *are* spatial.

So it might seem that Honderich's Radical Externalism does indeed allow us to do justice to both the subjectivity of consciousness and its causal efficacy. A pretty neat trick, if successful.

Unfortunately, I don't believe it is successful, as I'll now try to explain. In particular, some very serious problems concerning causal interaction appear to remain.

Honderich outlines *one* problem about causal efficacy facing spiritualism — the problem of how the non-spatial might causally impinge on the physical. And perhaps Radical Externalism does succeed in sidestepping *that* problem. The problem is that that is not the only problem. In fact, spiritualism faces far more serious difficulties concerning causal interaction, including the kind of difficulty that Honderich earlier raised for anomalous monism.

The problem for Honderich, as I'll try to show below, is that *Radical Externalism faces much the same kind of difficulty concerning causal interaction.*

Honderich's Attack on Anomalous Monism

Let's begin by briefly reminding ourselves of Honderich's own earlier attack on Davidson's anomalous monism (Honderich, 1982). Anomalous monism itself arose in part out of a problem concerning the causal efficacy of the mental. According to Davidson, there can be a causal relationship between events only if they fall under some law. But there are no psychophysical laws. So how then can mental events cause physical events?

Davidson's solution is to point out that laws relate events only under some description or other. A token mental event *a* can cause a physical event *b* if *a* is also a physical event. For then *a* and *b* can both fall under some physical description and so some physical law. But because there are no *psycho*-physical laws — no laws under which mental events *described as mental* cause physical events, so there can

be no possibility of a *type*-identity between mental and physical events.

So Davidson's anomalous monism claims that token mental events are token physical events. But it seems that these token events have two quite distinct sorts of property: mental properties and physical properties. The mental properties of an event are neither identical with nor reducible to its physical properties. So it appears that anomalous monism is still committed to a *form* of dualism — namely, a form of *property* dualism.

Which brings me to Honderich's attack on anomalous monism. Back in 1982, Honderich cleverly pointed out that while Davidson does indeed allow mental events to have causal efficacy, it seems that epiphenomenalism still threatens. The problem is that while anomalous monism may allow mental events to have causal clout, *it's not in virtue of their mental properties that they have such clout*. You could entirely strip away the mental properties of an event, and its physical efficacy would remain undiminished. But, as Honderich puts it, it's surely mental events *as mental* that have causal efficacy. Honderich concludes that Davidson thus fails to do justice to one of our fundamental intuitions about the mental. Surely the mental properties of events *are* causally relevant. Anomalous monism makes them epiphenomenal. So anomalous monism won't do.

Why Non-Spatiality Is *Not* the Only, Or the Most Serious, Problem Concerning Causal Efficacy Facing spiritualisms

Honderich's argument against anomalous monism is neat, clear, and telling. But notice that it *has nothing to do with the non-spatiality of the mental*. (Indeed, notice that it's not even clear that anomalous monism makes mental properties non-spatial. If physical events are spatially located, and physical events have mental properties, then why aren't mental properties also spatially located?) Indeed, surely the Really Big Problem about the causal efficacy of the mental so far as dualism more generally is concerned is *not* that it makes the mental non-spatial (for, as I say, it's not clear property dualism has to make the mental non-spatial, and in any case we might question whether the concept of cause requires both a cause and its effect to be spatially located).² In fact, what more often than not tends to force modern dualists in the direction of epiphenomenalism are not worries about the non-spatiality of the mental role so much as worries about the

[2] After all, there are many who would at least allow action at a distance, and if we allow that, then allowing non-spatial causes is perhaps not such a great further leap.

absence of psycho-physical laws and/or worries about the causal closure of the physical — in particular, the worry that if every physical event has a sufficient physical condition, then the mental ends up being causally locked out of the physical domain. As I say, this sort of worry about causal interaction is quite independent of any worries generated by the alleged non-spatiality of the mental.

But then a problem for Radical Externalism is this. Maybe it does sidestep *one* of the problems for spiritualism concerning causal interaction. By making worlds of perceptual consciousness spatial Honderich does perhaps avoid the problem of explaining how the non-spatial might impinge on the physical (though I have my doubts even about this: even if I acknowledge that a ‘phantom’ pain produced by a severed limb is spatially located [in thin air], that, by itself, still leaves it blankly mysterious how this pain might have any causal impact on the physical). But *that is not the only problem dualism faces so far as causal interaction is concerned*. Indeed, it may yet turn out that Radical Externalism is vulnerable to some of the other classic problems.

In fact it seems to me that Radical Externalism *is* vulnerable. For much the same kind of argument that Honderich wields against anomalous monism can be used against his own Radical Externalism.

Let’s return to worlds of perceptual consciousness. They are supposed to solve the problem of causal efficacy by making consciousness spatial. Non-spatiality is supposedly what leads spiritualisms into trouble with causal efficacy. So that trouble is sidestepped.

Well, yes, *that* trouble is avoided, but the other problems remain. Worlds of perceptual consciousness may be spatially extended. But Honderich makes clear that they are nevertheless numerically distinct worlds. Nor are they identical with the physical world (or, I take it, any part of it). But then *what difference can they make, causally speaking, so far as what goes on in the physical world is concerned?* After all, if we removed these worlds of perceptual consciousness, the physical world would remain, and everything, presumably, would continue on in it exactly as before. So *why aren’t worlds of perceptual consciousness epiphenomenal?*

In particular, if we can explain everything that happens physically entirely by reference to the physical, without our making any reference to worlds of perceptual consciousness or whatever goes on in them *at all*, then aren’t worlds of perceptual consciousness epiphenomenal?

A Reply

To this, the reply may be that, while a world of perceptual consciousness may not be identical with the physical world or any part of it, *it can nevertheless include physical objects*. I am currently conscious of this book. Therefore this very book constitutes a part of my world of perceptual consciousness. As the book is a physical object, it can have physical effects. But then, as my world of perceptual consciousness *includes* the book, *so it too is able to have physical effects*.

This surely won't do. Suppose this book tips over and knocks over a vase. The book tipping over causes the vase to smash. And the book tipping is part of my world of perceptual consciousness. So is this an example of my world of perceptual consciousness having a physical effect? That's an odd-sounding conclusion to draw, to say the least. But even if it did follow, it remains true to say that *featuring in my world of perceptual consciousness has nothing to do with the book's causal efficacy*. Had my world of perceptual consciousness been removed, the causal sequence that was played out in front of me — the book tipping resulting in the vase smashing — would still have been the same. We might put it like this. The event of the book's tipping may have a certain *mental property* — the property of falling within my world of perceptual consciousness. But *this mental property of the event is causally irrelevant to how things play out physically*. The book would have done what it did anyway, whether or not it happened to feature within my world of perceptual consciousness.

Of course, Honderich rightly points out that a world of perceptual consciousness is causally dependent upon what's going on physically. It exists in part as a result of what's going on neurologically. But of course this doesn't make worlds of perceptual consciousness any less epiphenomenal. Honderich allows physical properties can have effects on consciousness. The difficulty is in explaining how consciousness is to have any physical effects. It does seem as if God could have made a physically identical world, but without including any worlds of perceptual consciousness at all. They add nothing so far as how things play out physically is concerned.

But if Radical Externalism falls foul of these same classic difficulties concerning the causal efficacy of the mental, then it seems it has no very substantial advantage over the various spiritualisms to which Honderich thinks it should be preferred.

Does Radical Externalism do Justice to Subjectivity?

I'll finish by briefly turning to the other horn of the dilemma about consciousness — of how we are to do justice to the subjectivity of consciousness. This is something Honderich claims 'devout physicalisms' fail to do. But does his own Radical Externalism fare any better?

Part of my difficulty here is in identifying precisely what worlds of perceptual consciousness are supposed to include. The suggestion seems to be that they can include real physical objects, as opposed to mere subjective surrogates for them. Suppose, for example, that you are conscious of this page. Honderich asks:

What did your consciousness *seem* to consist in? An answer can grow on you fast. It *was for the page to be there*. What your consciousness seemed to consist in was nothing other or more than that. In a better sense of the words than employed by some philosophers, that is what it was like for you to be conscious of the page and that is all that it was like (p. 5).

The idea seems to be that your consciousness of the page is neither more nor less than for the page itself to exist. Well, actually, that's obviously false, as the page can exist without you being conscious of it. It's not entirely clear what Honderich is after, here. But it does at least seem that he wants to make the *physical object itself* feature in your consciousness. He wants to include *it* (and not just some mental surrogate for it) in your world of perceptual consciousness.

But earlier we noted that one of the criteria Honderich thinks any adequate theory of consciousness should meet is to do justice to the thought that when it comes to the realm of consciousness, the illusion/reality distinction collapses. You can be mistaken about what's objectively there, but not about what's subjectively there.

But then an obvious question arises — if real physical objects can crop up in all their objective glory within worlds of perceptual consciousness (as opposed to mere subjective surrogates — seemings, sense- data, ideas, or whatever) why doesn't that make possible exactly the distinction between illusion and reality Honderich wants to avoid?

Perhaps it doesn't, but I don't yet see why it doesn't. After all, it may seem to me like there's a book in my world of perceptual consciousness, but as a matter of fact there isn't. I'm hallucinating. So I'm mistaken about what there is in my world of perceptual consciousness.

In a nutshell, the problem is this. It seems Honderich wants to include real physical objects within worlds of perceptual consciousness. But by including them, worlds of perceptual consciousness no longer appear to be the infallibly given worlds it seems

he requires them to be if his theory is to satisfy his own criterion concerning illusion/reality.

Conclusion

In the preceding section, I raised a worry about how Radical Externalism is supposed to explain what Honderich suggests any adequate theory of consciousness should explain: namely the absence of an illusion/reality distinction within the realm of consciousness. Perhaps Radical Externalism can explain this, though I don't yet see how.

This is merely a worry, however. Perhaps my inability to see how Radical Externalism explains this is down to my not having fully understood it. It seems to me that the more substantial difficulty facing Radical Externalism concerns causal interaction. Perhaps Honderich has solved one of the problems that plagued at least some spiritualisms concerning causal interaction. But in my view the most serious difficulties concerning causal interaction are just as much difficulties for Radical Externalism as they are for spiritualism. So, as it stands, Radical Externalism seems to me to have no very substantial advantage over spiritualism.

Like Honderich, I'm not keen on 'spiritualism' or 'devout physicalism'. I am persuaded that we probably should be looking for a much more radical solution. Honderich's Radical Externalism is bold, imaginative and, I suspect, a very significant step in the right direction. But I am not persuaded that, as it stands, Radical Externalism really does solve the causal interaction problem, as Honderich claims. Not as it stands.

References

- Honderich, Ted (1982), 'The argument for anomalous monism', *Analysis*, **16**, pp. 59–64.
 Honderich, Ted (2006), 'Radical externalism', *Journal of Consciousness Studies*, **13** (7–8), pp. 3–13. [This issue]

REPLY TO LAW BY HONDERICH

Stephen Law in his independent and properly tough-minded piece begins by agreeing, I take it, that devout physicalism — consciousness itself is nothing but physical — can give no adequate account of what has long been called the subjectivity of consciousness, no account of the fact that your consciousness is not a thing, property, fact or state of affairs about you like your location, DNA, weight, synapses or the like.

He depends for his agreement, more than I do, on philosophical thought-experiments about black and white rooms, fool's pain and