

**I might not have been the author of this article:
An argument for Cartesian dualism**

Appealing to Nietzsche's idea of the Eternal Recurrence, in this piece I consider a possible world where I did not exist, everything else remaining the same. The individual who came from the same sperm and egg as the author of this article, might not have been I. This is presented as an argument for interactionist dualism. While Descartes argued that I might have existed even if the world did not, I argue that the world — this same world — might have existed even though I did not. The interpretation that the author might have been a zombie is rejected. Kant's objections to the idea of a Cartesian mental substance in his 'Paralogisms of Transcendental Psychology' are also considered. 'I' might unknowingly 'be' multiple 'I's. Moreover, there are potentially infinite worlds physically identical to the actual world, in each of which the author of this article is a numerically different 'I' from the 'I' (or 'I's) existing in the actual world.

Under what circumstances might I not have been the author of this article?

One case that comes to mind is if the article is plagiarized. It might actually be the case for all you know (if you didn't know me better) that someone else wrote the article and I am presenting it as my own work. Maybe the author and I had an agreement, an understanding that I would claim authorship of the article in return for some consideration, say, help with a loan or an all expenses paid holiday.

Moving into the realm of what might have been in some other possible world, someone not connected to me — say Georgina Kettle — might have written the very article I was going to write, and submitted it to *Philosophy Pathways*. Imagine my shock and surprise. A more bizarre and shocking version of this story would be that a philosopher who wasn't me but who coincidentally has the same name as me wrote the article and submitted it.

So far so easy. We have considered possible worlds where I exist, but I am not the author of the article. However, we also have to consider the possibility that I might not have existed.

Let's imagine a possible world where I don't exist and someone else wrote this article. Georgina Kettle was the author. If you like, you can imagine a back story for Georgina that makes it plausible that she would write the very article I wrote in the actual world. On second thoughts, maybe a feminist linguistic analysis of the article would show that the writing is so typically 'male' that it could hardly have been written by a woman — except, perhaps, as pastiche. (I wouldn't know, that's not my area of expertise.)

We can narrow down the possibilities somewhat. Let's assume that another man wrote the article. We are considering a man similar to me in certain respects, with similar philosophical influences and interests. Perhaps 'we' did our degrees at the same university. He considers himself to be an atheist. He has a squint like Jean-Paul Sartre and consequent double vision, which inclines him (so he supposes) towards a dualist view of reality.

But why stop there? In a possible world where I was not born, let's say that my parents had a first son, whom they named George. Maybe George and I share birthdays, went to the same schools, led similar lives right up to the point where this article was written. Or why not just go the whole hog and consider the possibility that my parents had a first son whom they named Geoffrey, who by coincidence had the identical genetic makeup as me, who went on to live the same life, experience the same experiences, think the same thoughts — and write the same article.

The other 'me' is not *me*. That was our assumption. The accepted view — at least since Saul Kripke's landmark essay *Naming and Necessity* (1971, 1980) — is that regardless of genetic makeup or any other physical properties, if the other me came from a different sperm and egg pair, then we are not the same. If he came from the same sperm and egg pair, then he and I are one and the same — in which case I would, after all, be the author of this article.

Is that the best we can do? I don't think so. As we have been venturing into the realms of the metaphysically possible, it would be remiss not to consider an hypothesis first advanced by the Stoics and revived by Nietzsche in the nineteenth century: the Eternal Recurrence.

In his book, *The Gay Science*, and then again in *Beyond Good and Evil* and *Thus Spake Zarathustra*, Nietzsche invites the reader to suppose that the universe endlessly repeats, and, each time the universe repeats, I return; my very self, not merely someone exactly like me. Conventional wisdom has it that view this is incorrect. Barring circular time, the minimal criteria for personal identity rule out the possibility that my very self 'comes back' in the next universe along, and so on ad infinitum. I may indeed have a special relationship to all 'my' future selves, I have every reason to be concerned about all the mistakes they will inevitably make, the misadventures that will befall them. But none of them is I.

The Eternal Recurrence is just one version of a range of universe-duplication scenarios. If these require that the laws of physics be bent in various ways — ever so slightly, or maybe a lot — then let them be bent. The details aren't important.

It follows from what has been said that it is possible that there might, in some other possible world, have existed a universe identical in every respect to the actual universe, the only difference being that *I* am not in that universe. In terms of our picture of the Eternal Recurrence, this universe might have been some other universe in the series, one in which I do not exist. In that possible world, I am not the author of this article but 'my' counterpart is. 'My' universe need never come to be. In its place, we have an identical universe where *I* never existed.

Regardless of one's view of possible worlds — or how talk of 'possible worlds' is to be analysed — my claim is incompatible with naturalism and empiricism, however broadly construed. Empiricism holds that the assertion of non-identity must be based on some discernible or describable *difference*, or it is meaningless. But there is not, and cannot be an empirical explanation why I am here, rather than not here, in the

sense that I have explained. In short: *I might not have existed, everything else in the physical world remaining the same.*

Using the hypothesis of an evil demon, Descartes argued in his *Meditations* that there is a possible world where I exist but the physical universe does not. I have argued, in effect, that there is a possible world where this physical universe exists but I do not. In either case, the same conclusion follows: mind-body dualism. One advantage some will see that my argument has over Descartes is that my argument doesn't depend on the possibility that idealism is true. That is not an issue that worries me, as I am perfectly prepared to consider that idealism might be true — although I do not see (and my argument does not provide) any strong reason for believing the idealist theory.

The main concern will be that my argument, as with Descartes, rests upon an *intuition*, or what Descartes would call a 'clear and distinct idea'. I have a clear and distinct idea of *the possibility of my non-existence*. I can see no credible *reason* why I am here, rather than not here, at least no reason that physics or biology (or etc.) can provide. That I am here is a brute fact, but it is not — it cannot be if my clear and distinct idea is true — a brute physical fact.

The mind-body dualism in question must be *interactionist*. We are not considering the possibility that the individual writing this might have been a zombie, everything else remaining the same, which has been used as an argument for dualism of the epiphenomenal variety. (According to Robert Kirk writing in the Stanford Encyclopedia, David Chalmers first proposed the zombie argument against physicalism in 1996.) The problem with the zombie argument is a glaring loophole: if a *zombie* would say exactly the same thing I am saying, then surely whatever explanation is available to account for the zombie's saying that — or, rather, making those marks, or emitting those sounds — is also available in my case. (That's not a knock-out refutation of the zombie hypothesis, but it comes close.) Whereas, I am saying that if the author of this article had not been I, then the author would necessarily have been, not a zombie but another person exactly like me: a *different I*.

A similar objection applies to the parallelist theory. According to parallelism, there is a physical ‘author’ of the physical article, plus a non-physical author of the non-physical ‘article’. (Note the scare quotes.) The same explanation as in the zombie case of what caused the physical ‘author’ to write what ‘he’ wrote leaves no room for an additional, fundamentally different explanation of *my* authorial actions.

How many different ‘I’s might ‘I’ have been? An infinite number, it seems. The otherwise identical author of this article in the possible world we have been considering, asks himself the very same question as I am asking, imagines ‘himself’ not being ‘I’ but being a ‘different I’. It follows that there are potentially an infinite number of possible worlds identical to the actual world, except that in each world the author of this article is a numerically different ‘I’.

This startling conclusion has echoes of Kant’s arguments, in the ‘Paralogisms of Transcendental Psychology’ in the second Part of the *Critique of Pure Reason*, against a Cartesian mental substance. If I am a Cartesian mental substance, then it is possible that there are a thousand ‘I’s or mental substances thinking the thoughts I am thinking, as ‘I’ (or, rather, they) write these words. It is possible that my ‘I’ is annihilated every five seconds and is replaced by a qualitatively identical ‘I’, all ‘my’ memories remaining intact. (Kant’s argument is used in an article by P.F. Strawson, which has been studied by generations of students: ‘Self, Mind and Body’.)

These are difficulties for the version of mind-body interactionism I have argued for, but not conclusive objections. We have thrown naturalism and empiricism overboard. Presumably, there must be *some* explanation of why I am here, but it is a non-natural one. Let’s say the explanation is *unknowable in principle* for beings like ourselves, reliant on sense perception for our knowledge. In Kantian terms, the non-physical self is a *noumenon*. There might indeed be, in reality — however unlikely it may seem — a thousand noumenal ‘I’s composing these words on this one computer screen. Or perhaps my noumenal ‘I’ is replaced every five seconds. We can’t ever know, so it’s pointless to speculate.

Pointless but not meaningless. Any objection to the meaningfulness of the hypothesis of multiple 'I's requires the additional assumption of verificationism. The standard reply to the verificationist is, just because you can never, or never in principle *know* the answer to a question — in this case, 'how many 'I's' — doesn't mean that the theory which prompts the question doesn't make sense. The onus is on the verificationist to prove the validity of the verification principle, in the face (as many critics have pointed out) of its own *unverifiability*.

At the other extreme, is the notion that all subjects of consciousness are one. There is ultimately only one point of view — one *Atman* — the theory held by Vedanta Hinduism, and also proposed, interestingly enough, by the American philosopher Thomas Nagel in his book *The View From Nowhere* as a potential solution to his riddle about the meaning of 'I am TN'. (He calls it the 'objective I': in this sense 'I am TN, and 'I am also GK, and also, etc.)

One great virtue of the Atman theory is the heroic way it deploys Occam's Razor. However, on my picture the same thing applies to the Atman as applies to the Cartesian soul. In another possible world, there might have existed a numerically different but otherwise identical Atman from the Atman that exists in the actual world. That's the point of the argument: that the 'I' is ultimately *contingent*, but not contingent on any *physical* state of affairs.

There will no doubt be other objections to the case presented here for mind-body dualism, which do not necessarily rest upon the assumption of verificationism. However, any objection that depends on naturalism or empiricism must be discounted, because it begs the question.

So, for example, if you argue that *if* physicalism were true, then a self-conscious being such as myself *would be tempted* by the thoughts I am expressing here, then the reply is that you have no right to make that assumption. The case is quite different from the analogous objection to Darwin's Theory of Evolution, where the Creationist argument that a wing or an eye could not have evolved is met by a detailed explanation, using computer modelling, of exactly how, according to evolution, a

wing or an eye might evolve. A.I. research has not yet reached the point — and never will, if my argument is valid — where the possibility that a computer running a program could have a ‘point of view’ like you or me has been demonstrated.

The so-called ‘locus of interaction’ has been a bugbear for Cartesian dualists. Descartes thought the pineal gland was the place where the soul diverts the movement of the ‘animal spirits’. The problem with Descartes’ proposal is that this ‘diverting’ mechanism is permitted in Cartesian mechanics, but not in Newtonian mechanics. Descartes’ theory thus conflicts flat out with Newton’s principle of the Conservation of Energy. However, if interactionist dualism is true, then we know that there must be wriggle room somewhere. — I am not going to speculate where that might be.

Perhaps the most worrying objection will be that if you throw too much overboard, there is nothing upon which to test one’s intuitions, ‘clear and distinct’ or otherwise. If we are no longer to be good empiricists, then what is to become of us? Does this signal a return to the bad old days when metaphysics held sway over science and logicians debated how many angels can dance on the head of a pin?

I don’t know where this might lead or what *post-empirical* philosophy will look like — if it ever comes to be, which is looking increasingly unlikely in our technological age. My case is that, if we discard our natural prejudice, as inheritors of the legacy of logical positivism and the linguistic turn, the intuition in question is a very strong intuition indeed, requiring something equally clear and distinct to refute it.

Bibliography

Descartes, R. *Meditations on First Philosophy*, 1641

Kant, I. *Critique of Pure Reason*, 2nd Edition 1787

Kirk, R. ‘Zombies’, *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Zalta, E.N. (ed.)

<https://plato.stanford.edu/archives/sum2015/entries/zombies/>

Kripke, S. 'Naming and Necessity', *Semantics of Natural Language*, Harman, G. and Davidson, D. Eds. Synthese Library 1971

Kripke, S. *Naming and Necessity* Basil Blackwell 1980

Nagel, T. *The View From Nowhere* Oxford University Press 1986

Nietzsche, F. *The Gay Science* 1882

Nietzsche, F. *Beyond Good and Evil* 1883

Nietzsche, F. *Thus Spake Zarathustra* 1886

Strawson, P.F. 'Self, Mind and Body', *Freedom and Resentment and Other Essays* Methuen 1974

© Geoffrey Klempner 2017

Email: klempner@fastmail.net