

Is Consciousness Transcendent?

Comments on James Tartaglia's *Philosophy in a Meaningless Life: A System of Nihilism, Consciousness and Reality*

Philip Goff*

Abstract

In this paper, I will discuss James Tartaglia's view on consciousness, as laid out in chapters 4 and 5 of *Philosophy in a Meaningless life*. Chapter 4 is an excellent critique of physicalist accounts of consciousness. In chapter 5, Tartaglia develops an original and intriguing alternative: the 'transcendent hypothesis', the view that both consciousness and the physical world it puts us in touch with are elements of a reality whose nature is entirely unknown. I will raise small concerns about the critique of physicalism. More broadly, I worry that there is a tension between chapter 4 and chapter 5: it seems to me that if the arguments of chapter 5 succeed in demonstrating that consciousness is unknowable, then this undermines the anti-physicalist arguments of chapter 4. Finally, I will respond to Tartaglia's rejection of more standard alternatives to physicalism.

In general, analytic philosophy has less to say about the meaning of life than other philosophical traditions. Many analytic philosophers doubt that questions concerning 'life's meaning' are themselves meaningful questions. Certainly the vast majority of analytic philosophers (and I put myself in this category) will feel that they can get on with their own little branch of philosophy – ethics, epistemology, philosophy of mind, or whatever – without relating it to questions pertaining to the meaning of life. In so far as the problem of life's meaning is taken seriously, it is seen as an isolable philosophical issue that a philosopher may or may not be interested in.

In this context, it is wonderful to find, in James Tartaglia's *Philosophy in a Meaningless Life*, a return to true systematic philosophy. Three central topics of philosophy – consciousness, time, and universals – are dealt with through a single approach, an approach focusing on the meaning of life. Each chapter is rich and thought provoking, but in this essay I will focus on the two chapters of the book (4 and 5) that deal with consciousness.

Chapter 4 is an excellent critique of physicalist accounts of consciousness. In chapter 5, Tartaglia develops an original and intriguing alternative: the

* Associate Professor, Central European University, Hungary. Email: goffp[a]ceu.edu.

‘transcendent hypothesis’, the view that both consciousness and the physical world it puts us in touch with are elements of a reality whose nature is entirely unknown. I am broadly in agreement with the rejection of physicalism, but I will raise some responses to the argument of chapter 5 in support of the transcendent hypothesis. More broadly, I worry that there is a tension between chapter 4 and chapter 5: it seems to me that if the arguments of chapter 5 succeed in demonstrating that consciousness is unknowable, then this undermines the anti-physicalist arguments of chapter 4. Finally, I will suggest that Tartaglia presents a less than conclusive case against more conventional alternatives to physicalism, such as dualism and panpsychism, and that these might be better options for the anti-physicalist.

Can physical science account for consciousness?

Physicalism is the view that fundamental reality is entirely physical. There are broadly speaking two approaches the physicalist can take to consciousness: reductionism and eliminativism. The reductionist tries to account for consciousness in terms of physical processes in the brain. The eliminativist denies that consciousness exists, and in this way dispenses with the need to account for it.

The latter option is not very popular. For most phenomena, you’ll find some philosopher willing to deny its existence: free will, moral value, the material world. But the reality of consciousness seems so evident, that few philosophers are prepared to embrace out and out eliminativism about it. What could be more evident than your present experience of colours, sounds, emotions, etc.? Perhaps because it has few adherents, the case against eliminativism is not well explored. Most are happy to take the reality of consciousness as a non-negotiable starting point.

Tartaglia offers an extremely interesting argument against eliminativism, going beyond just pointing out its basic implausibility. In the case of eliminativism about other phenomena – free will, moral value, God, or whatever – the eliminativist is able to make sense of the *rational basis* for belief in the entity in question. The eliminativist about, say, free will, ultimately thinks that belief in free will is false, but she will have something to say about *why* people believe in free will, something that makes sense of how rational women and men could come to believe in such a thing. Many eliminativists would say that people

believe in free will because it sure *feels* like we make free decisions.

In the case of eliminativism about consciousness, however, it's hard to see what could be said. One cannot say (as one would probably want to say in the case of free will) that it *feels* as though we're conscious, because of course to accept the reality of feelings just is to accept the reality of consciousness. It seems that the consciousness eliminativist will have to say that we just have a basic, ungrounded, disposition to claim that we have experiences, contrary to the reality. As Tartaglia puts it:

... the enduring inclination is to judge that I am having a certain kind of experience. But if I am not, and there is nothing in the objective world that I could mistake for an experience, then this inclination can have no rational basis, and must rather be an automatic and senseless reaction. This, however, is not how it strikes us at all: it is not as though I inexplicably find myself wanting to spout the word 'I'm having an experience' without knowing why; as if believing we have experiences were like suffering from Tourette's syndrome.' (p. 94)

This seems to me a powerful and ingenious way of strengthening the case against eliminativism.

What about reductionism, the more popular form of physicalism? Tartaglia argues, quite powerfully in my view, that reductionism collapses into eliminativism, and so ends up being just as implausible. The problem is that our concepts of consciousness richly characterise it, and that that characterization is inconsistent with the characterization physicalists give of it. Physicalist U. T. Place argued that our concept of a green experience amounts to nothing more than 'the sort of thing we have when we see something green.' Such a minimal characterization leaves the metaphysical nature of the green experience completely open, and hence leaves it open that the green experience could turn out to be a brain state (which is exactly Place's view).

The problem is, as Tartaglia puts it, 'our conception of conscious experience is not remotely this anodyne' (p. 96). He argues that, contra Place, our ordinary mental concepts characterise green experience as *having a green quality* ('although not 'green' in the same sense we use to describe a patch of light'). And more broadly, we have a rich understanding of what an experience is: it is a certain kind of self-aware state. These kinds of essentially subjective properties

have no place in the objective picture of the world we get from physical science. The reductionist has managed to account for ‘consciousness’ by re-defining it. It’s a bit like claiming to have proved the existence of God, by redefining ‘God’ as the physical world, and then claiming that the existence of the physical world entails the existence of God.

The currently dominant form of physicalism – the so-called ‘phenomenal concept strategy’ – tries to avoid this difficulty by claiming that we refer to our conscious states *directly*, rather than in virtue of any of their properties. We think about ‘water’ in terms of its superficial characteristics, such as its being colourless and odourless, and its being the stuff that falls from the sky and fills oceans and lakes. But, according to the phenomenal concept strategy, we don’t think about our conscious experiences in terms of any of their characteristics; we simply have a capacity to, as it were, blindly point at them through introspection: pain is ‘that thing’ *points at introspectively*. If this is our relationship to pain – a kind of blind pointing – then there seems no way of ruling out that the thing we blindly point at turns out to be a physical brain state.

Tartaglia’s reply is that to point blindly isn’t really to have a concept at all:

[the phenomenal concept strategy] forgets what it means to have a *concept* of something. My concept of a tree is what I believe the tree to be: it is how I think of it. It is true that I can have a false conception of a tree and yet still manage to refer to it; just as I can refer to a man at a party as the one drinking a martini even if he is drinking water. However, even a false conception presents my conception of what the things is; the Phenomenal Concept Strategy cannot exempt introspective concepts of this basic requirement ... (p. 98)

I’m not so sure. In *Naming and Necessity*, Kripke claims that proper names refer in virtue of a causal connection between the name and the referent, rather than in virtue of an associated description. A school child can refer to ‘Galileo’ without knowing anything about him (perhaps they mistakenly think he was a famous explorer), which shows that they can’t be picking him out in terms of any of his characteristics. The child manages to pick out Galileo because they use a term, i.e. ‘Galileo’, which is causally connected in the right kind of way with Galileo himself. This account of proper names seems to me fairly plausible, and it suggests that the concept expressed by a proper name is a kind of blind pointer.

I think it's better to say, not that there are *no* blind-pointer-type concepts, but simply that it's pretty implausible that *experiential concepts* are blind-pointer-type concepts. When I think about pain in terms of how it feels, I know something about its essential nature. That's what prompts my concern when others feel pain; I know what pain is, and hence I know it's a terrible thing to endure. We might bolster this claim with reference to Tartaglia's claim that we characterise experiences as self-aware states: to characterise a state as self-aware is to have a positive conception of it, not just to point blindly at it.

However, these are finer points of dialectical strategy. I agree with Tartaglia that the physicalist can accommodate consciousness only by *revising* our ordinary concept of experience, pretending that those concepts are more minimal than they in fact are. And this means that, in terms of consciousness as we ordinarily conceive of it, the physicalist is an eliminativist. And as Galen Strawson remarks (reported by Tartaglia), eliminativism about consciousness is 'the silliest view ever put forward.'

The Transcendent Hypothesis

Tartaglia rejects not only physicalism, but also the standard alternatives to physicalism, such as dualism, panpsychism and idealism. I will get to his concerns about these views presently. But first I want to explore the line of reasoning that gets Tartaglia to his favoured position: the transcendent hypothesis. According to the transcendent hypothesis, ultimate reality transcends our understanding of it. Ultimate reality, for Tartaglia, contains consciousness, but not in a form that corresponds to human understanding of it. The transcendent form of consciousness is entirely unknown.

The move towards the transcendent hypothesis begins with some reasons for thinking that our ordinary conception of consciousness is hopelessly flawed. As Tartaglia notes, it is a traditional view in modern Western philosophy that the mind is better known than the body, indeed that we have a perfect grip on the nature of our experiential properties. He argues, however, that upon reflection, our experiential concepts turn out to be mere 'shadows' of our concepts of properties in the material world. We seem to find in our experience shape-like properties: my experience of the tree in front of me seems to be tree-shaped. And yet 'our only notion of spatial arrangement and shape belongs to the objective world. We have no other notion, and besides, this is evidently the

notion we have in mind; the shape we discern in the experience is the shape the tree has, just as a photograph of a tree has the shape the tree has' (p. 110).

What about the greenish quality I find in my experience of the tree's leaves? Colour, as Tartaglia notes, has often been taken by philosophers to reside primarily in the mind rather than the physical world. But he then rejects this traditional view, with an argument reminiscent of Berkeley's attacks on Locke: 'if experiences do not have a size and shape, they can hardly have a colour which fills that size and shape ... the greenness I had in my mind filled the contours of my experiences; but since experiences can have no contours to be filled, this cannot be a property that experiences possess' (p. 111). What philosophers think of as the properties of experience – 'phenomenal shape' and 'phenomenal colour' – are so described using concepts borrowed from our concepts of features of the external world. We dress experience up in the clothes of external reality in order to make sense of our idea of it as an internal mirroring of that external reality.

I remain unpersuaded. Phenomenal colour and phenomenal shape *represent* their external analogues: when I see a tomato I have an experience which represents a red, round thing at a certain distance from me. But I see no reason to doubt that there is 'mental paint' doing the representing, mental paint with an intrinsic character known through introspection. We call a certain intrinsic property of experience 'phenomenal colour' because it represents phenomenal colour, and another intrinsic property of experience 'phenomenal shape' because it represents phenomenal shape. But this fact is not inconsistent with their having an intrinsic character of their own known through introspection. I do not take myself to have given an argument for this view; my only claim is that I can't see what reason Tartaglia has given us to doubt it.

Tartaglia presses the Berkeleyan argument:

... as Berkeley has pointed out, resemblance between ideas and physical objects makes no sense; and we might express this point by saying that if experience does not belong to the objective world, then an experience cannot resemble something objective in virtue of shape. We might try to get around this by talking of an abstract isomorphism rather than a resemblance, such that something about the nature of experience systematically correlates with the shape of the tree. But to say this is to admit that the notion of shape is inapplicable to experience, despite the

fact that when I attend to my experience, all I find is something shaped like a tree; the experience of the leaves is *above* the experience of the trunk and so on. (p. 110)

Certainly there is something we find in experience that we feel inclined to describe in such spatial terms. But I still don't see why I cannot say that that thing is an intrinsic phenomenal character – known in terms of its essential nature – and that it is external shape which is unknown – or rather known only in terms of its structural commonalities with phenomenal shape. On this view, it is our conception of external shape that is a mere 'shadow' of our conception of phenomenal shape, rather than vice versa. Again, I have not given an argument for the epistemic priority of phenomenal properties over external properties, but I cannot see that Tartaglia has given an argument for his converse prioritising of external properties over phenomenal properties.

But what is this 'phenomenal character' that we know through introspection? Tartaglia complains that we can't say much about it:

If I want to tell you what the experience of green is, for instance, I have but three very inadequate options: I can compare it to another colour experience; tell you how to get it so you can find out for yourself; or say some very general, philosophical things, such as that it is a subjective state that alerts people to the presence of green light. None of this remotely compares to the detailed knowledge than can be imparted about things in the objective world (p. 107).

It is a familiar point that experiential qualities are in some sense ineffable. But this is plausibly due to the fact that the concepts we use to pick them out are primitive. Compare to other plausibly primitive concepts: existence, metaphysical possibility, causation, the notion of a reason. It is arguable that none of these notions can be explained in more basic terms. If someone asks you what it is to have a reason to perform a certain action ϕ , you might say 'It's for it to be the case that something counts in favour of ϕ ing.' But this is really just to use different words to express the same concept. There is nothing mysterious about this, we have just reached the basic epistemic building blocks of our picture of the world. If, as seems plausible to me, experiential concepts are also epistemically basic, then we should likewise take it that we can't say much very

informative about the nature of a given experience. The nature is known, but it is known only through actually having the experience itself. If you have to ask, you'll never know.

Tartaglia goes on to argue that our conception of experience is not only inadequate, but flawed. This further argument goes as follows (p. 112):

1. In order to make sense of experiences representing the external world, it is necessary to conceive of experiences as being causally determined by the environment.
2. However, given that experiences are not part of the objective world, they cannot be causally impacted by features of the objective world.
3. Therefore, the notion that experiences represent the external world is incoherent.

The second premise seems to me to depend on equivocation concerning the word 'objective.' Experiences are not 'objective' in the sense that they are subjective properties, i.e. properties which characterise the subjective experience of an individual. But they are perfectly 'objective' in the sense that the facts about experience are perfectly objective facts about reality. If I am having a pain with a certain phenomenal character, then it is a fully real fact about the world that I am having a pain with that particular phenomenal character. If someone thinks I am not having that pain, then they are wrong. And if experiential properties are in this sense perfectly objective features of reality, then why should they not causally impact on the physical world? In the last section we saw reason to doubt that experience properties are *physical*; but just because a property is not physical it does not follow that it cannot causally impact on properties that are physical. The belief that a non-physical God impacts through miracles in the physical world, whether or not it is true, does not seem to be incoherent.

I am therefore not persuaded that I have reason to abandon the traditional view, which seems to me quite plausible, that the nature of experiential properties is (more or less) perfectly known through introspection. And for this reason, although I find the transcendent hypothesis intriguing and worth further consideration, I do not as yet take myself to have reason to accept it.

A tension between the two chapters

In chapter 4 Tartaglia argues against physicalism on the grounds that we have a rich conception of consciousness and one at odds with the characterizations of consciousness offered by physicalists. In chapter 5, Tartaglia argues that we know absolutely nothing about the true nature of consciousness. This is most explicit at the point when his discussion touches on considerations pertaining to the famous ‘knowledge argument’ against physicalism. In the much discussed version formulated by Frank Jackson, the genius brain scientist Mary has spent all of her life in a black and white room, from where she has learnt everything there is to know about the physical processes involved in colour vision. One day she escapes her room and, for the first time, sees something red. For proponents of the knowledge argument, Mary at this point learns something new: what it’s like to see red. This is supposed to show that there is more to red experience than can be known from physical science.

Tartaglia certainly wants to agree with the conclusion of the knowledge argument: red experiences are part of unknowable transcendent reality, and hence their nature cannot be known through physical science. But if red experiences are transcendent, then Mary’s experience of red cannot teach us anything about their nature either:

... it is easy to be misled by the thought that in having experiences we learn something new, namely *what it is like* to have them. There is something important to this though, since there is no experience in the objective world ... [h]owever ... it is a mistake to infer ... that knowledge of ‘what it’s like’ is knowledge of a different part of reality than the objective world. These are mistakes because in having an experience, and making sense of it in the only way we can – namely with objective thought – we are forming a misconception. (p. 115-6)

Why do I need to make sense of experience in terms of objective thought? Why can’t I form a perfectly adequate conception just by conceiving of my experience in terms of what it’s like to have it? It seems that I can entertain the possibility of solipsism – the hypothesis that all that exists is myself and my mental properties – and in doing so I think about my experience without bringing in the idea of anything from the external world. Why is this not a

perfectly adequate conception of my own experiences? I suspect there may again be an equivocation in talk of ‘objective’ thought and reality. The conception I form of myself and my experience when I entertain the epistemic possibility of solipsism is ‘subjective’ in the sense that it characterizes reality in terms of properties of subjective experience, but it is also perfectly ‘objective’ just in the sense that it is an objective fact about reality that I really am instantiating those subjective experiential properties.

Moreover, if I have no adequate conception of my experiences, and hence know nothing about their nature, how can I know that their nature is not entirely physical? The anti-physicalist arguments of chapter 4 depended on my having rich knowledge of the nature of experience, which enabled me to rule out that my experiences are physical brain processes. But if I know absolutely nothing of the nature of my experiential properties, then these arguments seem to be undermined.

Indeed, Tartaglia adopts a response to the knowledge argument popular among the phenomenal concept strategists, the very physicalists he argued against in chapter 5:

... it is said that someone acquainted with objective thought’s final story about the nature of red, would upon seeing red for the first time learn what it is like to see red – which is not something they could have known before ... [this is] best accounted for in terms of their acquisition of demonstrative concepts. Thus when they see red, they acquire a new demonstrative concept of it as ‘that’ property, the one they are indirectly aware of when they have a certain type of experience; as well as a shadow concept of ‘that’ as an experience of red. However although this will allow them to recognize red when they see it, and the experience of red when they have it, it will not teach them anything new about reality under either interpretation. (p. 116-7)

According to many phenomenal concept strategists, the reason we can’t know a priori the conscious states are physical states is that our experiential concepts are demonstratives, and hence do not reveal to us the physical nature of their referents. It is an empirical fact that our experiential demonstratives pick out physical brain properties. Given that Tartaglia also thinks of experiential concepts as demonstratives that leave us in the dark about the nature of

experience, how can he be so confident that experience doesn't have a purely physical nature? The arguments of the previous chapter depended upon our having a rich conception of experience, but the demonstrative account we find here seems to undermine that.

What's wrong with dualism and panpsychism?

Tartaglia spends much more time rejecting physicalism than he does rejecting the traditional opponent of physicalism. There has been much discussion recently of Russellian monism, a view which has two components:

1. Physics fails to reveal to us the intrinsic nature of matter, telling us only about its structural or dispositional properties
2. The intrinsic nature of matter is made up of conscious, or proto-conscious, properties.

Tartaglia's quick rejection of this view is premised on understanding Russellian monism as in competition with physical science, a competition which the former looks certain to lose:

... philosophers are in no position to pontificate on the inadequacy of our conception of matter; this is a concern which long ago passed into the hands of empirical science. And besides the inadvisability of philosophers stepping into core scientific territory as anything more than interested bystanders ... the fact remains that in a contest between our ordinary conception of consciousness and objective thought, objective thought wins hands down. Objective thought provides the foundation for all of our understanding of the world, and cannot be put in doubt by a philosophical problem. (p. 99-100)

However, Russellian monists do not see themselves as in competition with natural scientists, or trying to interfere with their work. The point is that, from Galileo onwards, natural scientists have been involved in a rather limited kind of project: mapping the causal structure of the universe. They have simply not been in the business of speculating about the intrinsic nature of matter. Philosophers should certainly leave it to physicists to tell us what the causal structure of

matter is. But they shouldn't leave it to physicists to tell us its intrinsic nature, because this is simply not the job of a physicist.

Russellian monists do not propose a competition between theorists of the objective physical world and theorists of consciousness; rather they propose a way of bringing both together in contributing towards a unified picture of reality. On this unified picture, consciousness is the intrinsic nature of the stuff physics describes extrinsically.

Tartaglia is even quicker in his rejection of dualism. In order to make sense of the interaction between mind and the physical world '... the dualists must attribute to the brain a unique capacity to interact with the non-physical world' (p. 102). But Tartaglia has already rejected any such 'metaphysical specialness' of the brain: 'we know from objective thought the brain is not radically unlike everything else in the world; it is the most complex organ in the human body and the one science currently knows least about, but it is nevertheless still a physical object, metaphysically on a par with every other' (p. 101-2).

This is reminiscent of the standard anti-dualist argument from the alleged 'causal closure' of the physical world. If the physical world is causally closed, then there is no space for the mind to do any causal work by making changes in the brain. However, although often stated, the causal closure of the physical is not often defended with empirical argument. It would be nice to hear a little bit more from Tartaglia of the case for causal closure.