

WHAT IS AT STAKE IN ILLUSIONISM?

James Tartaglia

Keele University

Abstract

I endorse the central message of Keith Frankish's 'Illusionism as a Theory of Consciousness': if physicalism is true, phenomenal consciousness must be an illusion. Attempts to find an intermediate position between physicalist illusionism and the rejection of physicalism are untenable. Unlike Frankish, however, I reject physicalism; while still endorsing illusionism. My misgivings about physicalist illusionism are that it removes any rational basis from our judgement inclinations concerning consciousness, undermines the epistemic basis required to explain the genesis of our physical conception of the world, and leads to a widespread scepticism about the basis of philosophical reflection. I endorse the core of physicalist intuition, but not its metaphysic, and sketch my alternative illusionism, which resists physicalism's merging of philosophy with science without thereby impinging on science. I conclude that physicalism is fostered by inattention to metaphilosophy and threatens philosophy's distinctive voice; but that illusionism itself is an important insight.

Keywords

Phenomenal Consciousness; Physicalism; Illusionism; Eliminative Materialism; Metaphilosophy; Transcendence; Identity Theory; Phenomenal Concept Strategy

Word Count (including notes and references)

8082

1. Towards Another Kind of Illusionism

As someone with strong illusionist sympathies, I found myself agreeing with so much of what Keith Frankish says in this refreshingly clear and insightful essay, that I formed the hope that it will become a landmark in the philosophy of consciousness from which future discussions take their lead. For with the issues set up this well, the debate might be able to move away from the features which have characterised it in the last twenty-plus years: lack of vision combined with tiny dialectical moves, needless proliferation of terminologies, and the tedious reciting and consequent dissection of a few simple thought experiments. With this paper as a new starting point, the debate might take a fresh turn and become more interesting. I think it deserves to acquire this status, because the central case it makes is of crucial philosophical importance. It is that if physicalism is true, then consciousness must be a kind of illusion. All the murky, intermediate positions trying to hold on to both consciousness (conceived in the ordinary, phenomenal way) and physicalism (conceived in the ordinary, non-revisionary way) are untenable. The philosophers at the opposite poles, namely anti-physicalists and illusionist physicalists, have been saying this for years. They were right. Attractive as it might seem, there is no stable middle ground on this issue. You cannot have your cake and eat it: consciousness (as ordinarily conceived) or physicalism (ditto) has to go.

Frankish makes this case extremely well, but it is worth reminding ourselves of the *prima facie* incompatibility from which it derives, lest we become tempted by the thought that a subtle philosophical distinction might save the day. Phenomenal properties are conceived in direct opposition to physical properties, namely as subjective properties typically caused by objective, physical properties. Thus phenomenal-red, for instance, is subjective in that unlike physical properties, it exists only from the perspective of the person experiencing it, and to

get a grasp on its notion requires us to contrast the redness a flower may possess, with the redness the flower causes within us. Phenomenal-red can only be introspected; red can only be perceived. The former is only there for me; the latter for everyone. In short, to understand the notion of phenomenal-red, you have to contrast it with red; that is how we all get a grip on the notion when we start studying philosophy. So the starting point for would-be intermediates is rather like that of trying to show that our conceptions of a square and of a circle are really just two conceptions of the same thing.

If we are conceiving the same thing in directly opposed ways, then one of these conceptions must be wrong. But in that case, since the physical conception is the one conventional physicalists trust to tell us how the world is, the intentional object conceived by the phenomenal concept cannot really be there. There cannot be anything in the world that fits the phenomenal conception, if it is directly opposed to another that provides an accurate conception of the same thing. So if the phenomenal concept is a concept of a brain state, it must be a radical misconception of it; we must be misconceiving the brain state beyond all recognition, in fact. We are thinking of a brain state as a subjective experiential array, but that is not what it is at all. Consequently, the array must be an illusion, even if thinking about it somehow allows us to think about real brain states.

Like Frankish, then, it seems perfectly clear to me that we cannot pretend our phenomenal conception of consciousness is OK if we think conventional physicalism is true. Our phenomenal concepts cannot be getting it right if our physical ones are, so if we are to preserve the *prima facie* situation, we must either accept that they are presenting us with illusions, abandon physicalism, or move into the dubious territory of making predictions about the future of physics. The only option for philosophers who want to both have and eat

cake, then, is to try to cast doubt on the *prima facie* conceptual incompatibility. But since it concerns how we are *conceiving* of the world, this project is doomed from the outset. For this is a factual matter; we cannot make our ordinary conception of consciousness something it is not. The insurmountable obstacle to the would-be intermediate, then, is that we *do* think about consciousness in a way which is problematic for physicalism; which is something they concede by entering the debate. They could resolve the problem by arguing that this conception is wrong, as the illusionist does, but there is no room to argue that we have the conception we do *and* that it is unproblematic for physicalism.

This is the standard approach nonetheless. Thus the mainstream of physicalism since the 1990s, inspired by Loar (1990), has tried to find a middle way by claiming that phenomenal concepts are empty and non-committal. The basic idea is that the concept of phenomenal-red, for instance, does not really tell us anything about its referent, but rather just directly picks it out; ‘directly’ in the sense that it does not connote any features of its referent, in the way the concept of Phosphorus connotes a morning appearance. This move is supposed to vindicate our ordinary conception of consciousness on the grounds that the conception of phenomenal-red is not telling us anything incompatible with physicalism – because it is not telling us anything about its referent at all, just referring to it. Thus the *prima facie* incompatibility is supposed to be removed in a manner which leaves phenomenal concepts intact.

The root problem with this proposal remained hidden for many years, due mainly to the wide variety of subtly different forms in which it was defended (it took about two decades to even acquire a standard name: ‘The Phenomenal Concept Strategy’). But now the dust has settled, I think it is plain that its central claim about the nature of phenomenal concepts is both unsupported and patently false. Phenomenal concepts are rich and theoretically committal. If

I look at a flower and think about what I am aware of as phenomenal-red, then I think my attention is focused upon an inner experiential array of phenomenal-colour, quite unlike anything to be found either on the surface of the petals or inside my brain. So as a matter of fact, phenomenal concepts are not as the Strategy describes them. I have previously laboured this point (Tartaglia, 2013), which Frankish gets across nicely in two paragraphs (pp. 13-14).

The Strategists did not argue for their ‘empty’ reading of phenomenal concepts; rather they defended it as a neglected theoretical alternative capable of solving the problem of consciousness (Tartaglia, 2013). If our concepts were as they say they are, however, the problem would never have arisen. So since these philosophers must originally have been working with the ordinary conception, their ‘empty’ reading must be revisionary. The Strategists forgot that a person’s *conception* of something is how that something seems to be to the person. They must have forgotten, since they argue, in effect, that although it seems to us that phenomenal states are private presentations with no place in the physical world, they also seem to be nebulous occurrences whose nature completely eludes us. This is obviously incoherent unless a temporal distinction is introduced, thereby revealing the proposal to be the overtly revisionary one of abandoning our original conception in favour of a new one. But then they cannot claim to have found a way of maintaining the *prima facie* situation.

Construed as a proposal to amend our phenomenal thinking, the Strategy asks the impossible of us. I cannot think of my experiences as presenting themselves as I-know-not-whats, because they present themselves as very familiar subjective states, the likes of which I have never encountered in the objective world (and cannot imagine how I could). Like Frankish, I take these presentations to be illusory, but I cannot eradicate my conception of them in favour of something more neutral; only build on it with the extra thought that they are illusory. To

remove the *prima facie* conceptual incompatibility, however, this is what the Phenomenal Concept Strategy asked of us. It took its lead from the Identity Theorists of the 1950s, with their original proposal that our experiential concepts are ‘topic-neutral’, such that we think of experiences as simply ‘something going on’ in specifiable circumstances. Place (1956) and Smart (1959) showed considerably more insight, however, because they recognised that this would require phenomenal thinking to be renounced.

That we do not conceive of our experiences in this anaemic manner did not present the same problem for the Identity Theory as it does for the Strategy. For this broadly functionalist conception is at least part of what we intend, and Place and Smart were quite clear that they thought the other, phenomenal part was untenable; the Strategists, by contrast, claimed to have preserved it. What the Identity Theorists were not clear about, quite excusably given that they were pioneers, was the implication that our ordinary conception of consciousness is of an illusion. They were proposing a new way of thinking about conscious states that was compatible with physicalism, and do not seem to have thought much about the old one. But it did not go away, and the intuitions it continued to engender kept physicalism mired in controversy. Things took a turn for the worse when physicalists who had been persuaded by Nagel (1974) and Jackson (1982) that the phenomenal conception was legitimate, tried to combine it with physicalism in order to find that elusive middle way. With this desideratum set, then once Kripke’s (1972) causal theory of reference entered the mix, we found ourselves landed with the Explanatory Gap and the Phenomenal Concept Strategy. Thus the original illusionist thrust of physicalism did not last long, being preserved only by the likes of Rorty (1982) and Dennett (1991). It now seems to be regaining focus, however, and this I welcome; hence my hope for the fortunes of Frankish’s paper.

Now given what I have said so far, and the fact that not so long ago I published a paper defending a physicalist form of illusionism along very similar lines to Frankish (Tartaglia, 2013), you will probably not be expecting what comes next (unless you read the abstract). For I now reject physicalism; though not illusionism. As I have said, I think the only stable options in this debate are physicalist illusionism or a rejection of physicalism. I opt for the latter; and yet I am still an illusionist.

You might well find this combination puzzling, if you think of illusionism as simply a consequence of physicalism. I do not. Evidently, then, I do not reject physicalism because I think that illusionism is obviously crazy, as many anti-physicalists do. And I am not about to defend dualism, panpsychism, or any of the other metaphysical positions familiar in this debate; my views on such positions are similar to those of most conventional physicalists. Neither am I going to reject physicalism on the grounds that it cannot be rigorously formulated; such considerations do not move me. Rather I shall be taking things up to the metaphilosophical plain, which is something I learnt to do from Rorty, one of the great illusionists.¹

2. The Problems of Physicalist Illusionism

Any credible metaphysic must be compatible with the manifest situation which we all, as conscious beings, find ourselves in. So it must be at least possible that what the metaphysic says about consciousness explains why, for instance, I seem to be here in my office on a certain Monday morning, with a point of view that takes in my computer screen (as viewed from a certain location and with certain eyes), my awareness of my thoughts in writing, and

¹ Arguably, he was its originator. See Tartaglia, 2016a, for my take on Rorty's illusionism; and for an unusual exposition of his position (written as if by him), see Tartaglia, 2016b.

sensations such as the coffee-taste lingering in my mouth. If the metaphysic would rule out this being my manifest situation, were it true, then it is not true. The manifest situation will not go away no matter how we metaphysically portray it, as Frankish realises (p. 9); if our metaphysical understanding of what is going on is incompatible with it, then we are simply confused. That is why ‘illusionism’ is a better name than ‘eliminativism’; for we cannot eliminate the manifest situation.

This, I think, is the core wisdom to Chalmers’s (1996) injunction to ‘take consciousness seriously’. Stripped of any further connotations, this might just seem obvious. And indeed it is; but plenty of physicalists have overlooked it by focusing on the details of the metaphysic (physical vs. functional, for instance), without concerning themselves with how *that* could be true given our manifest situation. Without an attempt at reconciliation, however, consciousness is not being taken seriously. Illusionists like Frankish and Dennett take it seriously. Dennett demonstrates this by sandwiching his account of consciousness between two renditions of a vivid example of the kind of manifest situation that needs to be explained; by the time of the second, he takes himself to have done the job (1991, pp. 26-7 & p. 407).² I do not think any philosophers really doubt the manifest situation, although they are sometimes disingenuously accused of doing so; as when Searle invites us to pinch ourselves as a preface to his critique of Dennett (Searle, 1997, p. 97ff). Such talk is best construed as saying, elliptically and with frustration, that the metaphysic is not compatible with the manifest situation; not that the theorist is unaware of that situation (or somehow does not share it). The problem of philosophers not taking consciousness seriously is not that of denying the manifest situation, but rather of not providing anything meant to explain it.

² For a critique of Dennett based around this example, see Tartaglia, 2016c, pp. 92-5.

Frankish seeks to explain it with intentionality, as I did when I was a physicalist. The idea is that the phenomenal arrays we are apparently presented with are objects of intentional acts of mind. As Brentano emphasised, such objects need not be real: we can think about real people, but also fictional ones like Sherlock Holmes. The manifest situation, then, is supposed to be explained by the fact that human beings have intentional states about fictions, i.e.

phenomenal states. This, Frankish suggests, explains their apparent causal potency; they ‘move us in the same way that ideas, stories, theories, and memes do, by figuring as the objects of our intentional states’ (p. 15).

The initial difficulty this proposal faces is that thinking about fictional objects like Holmes *feels* like something; this is even more obvious in the case of imagining them.³ We do not typically find the Holmes case puzzling because we do not bring the concrete phenomenological presence into dispute when we claim that it is not indicative of a real man. Illusionists must, however, so the parallel with thinking about fiction will not take them far before they must reconstrue it. They must claim that any apparent presence is itself a merely intentional object, not a real one. But then, how is it that it *feels* like something? As we understood this before, conscious awareness of fictional objects was facilitated by phenomenological presences. So if the presences are themselves to be reduced to intentional objects, we need a new account of the appearance of a concrete presence which occurs when we think about illusory objects; it can no longer be a real presence.

The account both Frankish (pp. 17-21) and Dennett (1991, p. 364ff.) provide is that the appearance of a concrete presence is a matter of our judging or representing that there is one. And we certainly do judge this, even if we then correct ourselves. So the issue is now recast

³ The illusionist should agree; if they are trying to account for our manifest situation, they should not concede ordinary terms like ‘feels’ to phenomenal realists.

in terms of the basis of this judgement inclination; whether we judge this way, and hence it seems this way, because of a real presence or not. The point the physicalist illusionist overlooks at this juncture, however, is that if we opt for ‘not’, as they suggest, then the judgement becomes evidentially baseless. In essence, the view is that when I seem to see a vivid patch of red on closing my eyes – that is, when I am inclined to judge there is such a presence, even if my illusionist commitments make me subsequently refrain – then there is no rational basis for my inclination whatsoever. My judgement inclination is rationally inexplicable from my own personal perspective, even if a causal explanation can be given for how my central nervous system is affecting my speech centre, say. This must be the case, because as physicalists of all persuasions continually emphasise, there is no evidence for the existence of phenomenal consciousness in the physical world. Nothing like that vivid red patch is to be found in the brain or anywhere else; which is why astute physicalists think it must be an illusion. But then, whatever physicalist illusionists think my judgement inclination is actually sensitive to, it cannot be a reasonable interpretation of the data.

This is the consequence of the physicalist illusionist’s substitution of bare judgement for phenomenology as an account of seeming. It is unavoidable, because they need this account of seeming for the intentional object account of phenomenology which makes their account explanatory; and they need the claim that there is no evidence for the existence of phenomenal presences for their physicalism. As I have said elsewhere (Tartaglia, 2016c, p. 94), the result is an account according to which our irrepressible urge to make judgements about the phenomenal character of our experiences is somewhat akin to suffering from Tourette’s syndrome – in that something in our wiring must be making it practically impossible for us to resist sincerely uttering sentences like ‘there is a vivid red patch’, even

though we have no evidence for the presence of such a patch, and even though we may be sold-up physicalist illusionists.

Two further consequences of making this move are as follows; the second will begin to take us into the metaphilosophical heart of the matter. The first is that if these judgement inclinations are rooted in an illusion, then our confidence in physicalism must immediately be shaken. This is because we drew up our physical conception of the world on the basis of (phenomenal) conscious experience.⁴ You could deny this outright by claiming that we would have come up with the same conception even without conscious states; that consciousness is epistemically epiphenomenal. If a claim this implausible were required by illusionism, however, we could dismiss it out of hand. But if we concede, as we must, that consciousness had a role in shaping our conception of the world, the illusionist claim that conscious states inspire us to believe in non-entities immediately places that conception in doubt. If a book central to a thousand years of Egyptology was discovered to be completely unreliable, our understanding of Ancient Egypt would be thrown into disarray. You do not need the foundationalist conviction that consciousness is our ultimate sourcebook to see that if judgements about conscious experiences are completely unreliable, then given how thoroughly intertwined they are with the genesis of our physical conception of the world, our confidence in that conception must be profoundly shaken. And our confidence in illusionism must be too. For how, on the basis of an illusion, could we discover true reality, and hence be confident that our evidence was indeed illusionary?

⁴ Frankish only says that the phenomenal aspect of experience is illusory, not experience itself, which can be understood functionally. I made the same distinction in Tartaglia, 2013, following the lead of the original identity theorists. However we are only aware of our experiences, such that they can enter into our reasoning about what caused them, what their function is, etc., because of what phenomenal realists understand as their phenomenal aspects. Since the distinction will not affect my case, I will hereafter omit it.

Since representations of phenomenal presences are misrepresentations, according to the illusionist, what they accurately represent must be brain states or distal objects; Frankish considers both options (p. 8), saying that consciousness might be an introspective illusion (brain states) or transparent (distal objects). The idea that they represent brain states holds particular appeal for physicalists, because on their account, that is what the representations ultimately are; hence they are representing themselves, which promises an account of self-consciousness. But then the misrepresentation could hardly be more radical; for brain states are nothing like vivid red patches. This is clearly quite unlike paradigm cases of misrepresentation where a representation which has its content fixed by sheep, say, is inadvertently triggered by the presence of goats; or Frankish's example of misrepresenting a flat hologram as a three-dimensional object (p. 24).

The situation is little changed if we think of conscious presences as misrepresentations of distal objects. On the face of it, this would allow us to invoke the primary / secondary quality distinction to find an element of accurate representation. The illusory appearance of an apple does at least seem to track its shape and size, and that starting point might allow us to determine that its illusory phenomenal-colour tracks certain types of physical surface. But this is to forget that the shape we discern in the experience is itself illusory (Tartaglia 2016c, p. 110). It is the shape of an illusory phenomenal presence; or, if we insist upon transparency⁵, it is the shape outlined by an illusory presence projected onto the object. Since my conception of this shape is a misrepresentation, then, I am in no position to assert any kind of isomorphism between the experience and apple. I have no reason to think my mistake can tell me anything about the shape of the apple; and hence no basis to draw up the

⁵ I would not; see *ibid.*, p. 113-4.

functional correlations between conscious states and distal conditions which physicalism relies upon.

Frankish looks hard for veridical elements to the illusion; perhaps phenomenal illusions and brain activity might be thought to share ‘intensity’ (p. 6), he suggests. But if I am wrong about the intensity of the experience, since phenomenal intensity is an illusion, then I have no basis to infer that it is indicative of a real feature. Since I can only cherry-pick the illusion for veridical elements once physicalism tells me what those veridical elements might be, this tactic cannot render the *genesis* of its conception of reality plausible; instead it is a tactic which simply dilutes the metaphysical insight of illusionism for the sole purpose of trying to preserve the physicalism it epistemically undermines.

Perhaps the problem is alleviated if we do not think of representation as causal covariation, but rather in terms of the functional-role semantics Frankish prefers (p. 23). I think this only makes matters worse. For then our representations have their content fixed by abstract brain functions, but misrepresent those functions as concrete presences; the misrepresentation must now cross the metaphysical divide between abstract and concrete, in addition to everything else it gets wrong. Nevertheless, supposing the magic trick is performed, we could then represent the function of the illusory phenomenal presences through how they relate to each other; that is, represent the functions of misrepresentations of brain functions (e.g. the functions of illusory red patches). But then to glean anything about the external world from that, we must correlate these functions with external conditions (e.g. red surfaces); and the only way to do that is to delve into the illusory content of the misrepresentations again (their phenomenal redness), thereby taking us back to the original problem.

By rendering consciousness illusory for the sake of our physical conception of the world, the physicalist forgets where that conception came from. With the physical conception in place, physicalist illusionism makes sense. But human beings built up that conception by working out what their conscious experiences were telling them about the world. If fictions were integral to our reasoning, we should never have believed the physical conception that tells us they are fictions. Frankish suggests that consciousness is an illusion hard-wired into our psychology by evolution (p. 24). But this view is inherently unstable when presented against the backdrop of physicalism. For why should we trust an evolutionary illusion to lead us to the true nature of reality? As Dennett says, ‘If some creature’s life depended on lumping together the moon, blue cheese, and bicycles, you can be pretty sure that Mother Nature would find a way for it to “see” these as “intuitively just the same kind of thing”’ (1991, p. 381). Evolution cares about survival, not getting things right. Perhaps getting it right explains how we have survived, but this is hard to maintain when pressing the case that our starting point was an illusion; for conscious experiences guided us to evolutionary theory.

This leads onto the second consequence, which is that physicalist illusionism renders our manifest situation illusory in a wholesale manner that goes far beyond debates about consciousness. For consciousness provides the basis of the intuitive way we have of thinking about our lives and world. Many aspects of this have been central to philosophical debates for millennia – from freedom to personal identity to conscience – but a particularly vivid one in this context is time. My experience tells me that this particular point in time is where I am in my life at the moment, and hence has a special significance (soon to be lost). But physicalist illusionism tells me this is no basis for a serious judgement. I should only believe the results of collective, scientific rationality; for physicalism holds that science must determine our metaphysics. I then find that contemporary physics strongly suggests a ‘block universe’

conception of reality, according to which my personal ‘now’ has no ontological significance. So I am not currently aware of even illusory phenomenal presences in any substantive sense. Rather, I am reporting an instantaneous state of four-dimensional reality; a projection or three-dimensional ‘time-slice’ on a par with any other projection of the four-dimensional reality that constitutes my life. In a very real sense, then, I am not even *here*; or better, I am everywhere that, while alive, I ever sincerely want to call ‘here’. Of course, some philosophers dispute the block conception; but since their starting point is how consciousness presents our manifest situation, physicalist illusionism denies them any good grounds for doing so.⁶

Does this mean that physicalist illusionism is incompatible with our manifest situation, and consequently false? I do not think we can answer this without determining whether physicalism is true. Frankish accepts that we have ‘introspective subjectivity’, which need not amount to ‘intrinsic subjectivity’ (p. 19), going on to say that for a (physicalist) representational theory of mind, the difference between a non-veridical introspective representation of a feeling and a veridical one is subjectively ‘no difference at all’ (p. 21). But without presupposing physicalism, I see no way of determining whether I could sincerely make judgements about my manifest situation in the absence of real phenomenal presence. Even if we made a machine that enacted the perfect illusionist programme, I could not step inside its introspective subjectivity to see if its judgements were sincere. I could not *become* the machine, only implement functionally identical representations in my own body; that of a conscious being.⁷ If physicalism is true, then I agree with Frankish that the machine’s judgements would have to be sincere; but if it is not, they might not be. The real crux of the

⁶ See Tartaglia, 2016c: chapter 6.

⁷ If on interfacing with the machine I were to lose consciousness, scientists would simply adjust the programme. When they got it ‘right’, this might simply indicate their ability to alter an extant consciousness, rather than manufacture it in a machine.

issue, then, is whether physicalism is true. Since illusionism undermines physicalism's evidential basis, I think we already have good reason to suspect that it is not. But with an alternative on the table capable of satisfying the core of physicalist intuition, I think we can do better.

3. The Nature of the Illusion

The core of physicalist intuition, it seems to me, is that natural science provides our best vocabulary for describing what actually exists. So when philosophers put forward consciousness as an obstacle to scientific explanation, this immediately seems dubious; given that philosophy has only *a priori* reflection, and science has equipment like the Large Hadron Collider! It sounds like either a throwback to the bad old days when philosophers tried to practise natural science *a priori*, as in idealist works of *Naturphilosophie*, or else just a badly informed defeatism about the limits of science, destined to be disproved by the next scientific breakthrough; if indeed one is even needed in this case. Speculation about a non-physical reality which interacts with or supplements the physical one, seems only slightly more outdated than speculation about a psychical 'inside' to reality and predictions of a currently unimaginable revolution in physics. For if science is to be supplemented, this will be done by scientists for scientific reasons; philosophers are kidding themselves if they think their reflections will lead the way. I agree with all of that. And once you have that core of intuition, the only place to go is physicalism, right? Wrong.

Physicalism is a philosophical position, according to which our metaphysical conception of reality is to be dictated by physical science. If, like Quine (1975), you do not recognise a distinction between philosophy and science, then this will not strike you as a significant

statement, and the core of physicalist intuition will immediately lead you to physicalism. If you do, however, then you will not think that the possibility of science providing a complete objective description of reality automatically settles the nature of the metaphysics we should adopt; because such a description might not capture the independent nature of reality, which is what metaphysics is interested in. To pursue this possibility without abandoning the core of physicalist intuition, you must ensure that your account of this independent nature does not interfere with science's objective picture by requiring anything of it. It must remain a purely philosophical matter, to which science can remain indifferent. This is a possibility which may hold appeal if physicalist illusionism strikes you as having problematic and unattractive consequences. And here is a sketch of how it might pan out, which will in turn shed light on what might be meant by a 'purely philosophical matter'.

Suppose you are currently dreaming. Granted, real dreams about reading philosophy papers may never happen (I hope not), and if they did, they would not have the phenomenological detail and lucidity of your current experience. But Descartes was surely right that you *could* be dreaming right now, even if we have to support the 'could' with science-fiction scenarios about brain manipulation. Now suppose that during this (super-) dream, you start to think about the problem of consciousness. You would no doubt find this just as puzzling as in waking life. If you are a physicalist, you would think about both your conscious experiences and the brain states in your head, and wonder how to make sense of the claim that they are the same thing; the impossibility of reconciling these directly opposed conceptions might lead you to illusionism. But note that the supposition that you are dreaming alters the significance of this thought process. For when thinking about the brain states in your head, you would naturally think of the head within the dream; the one with eyes focused upon these words. But of course, if it is a dream, then there is no prospect of identifying your experiences with

those brain states; for they are not real ones. They are no more real than anything else in the dream. Rather if your experiences are to be identified with any brain states, then it must be with the ones inside your real, sleeping head; the one resting on the pillow of your bed. For whatever reality there is to your experiences is not to be found within the dream.

From the experiential perspective of the dream, the world in which you are asleep in bed is a transcendent context; it transcends the space and time of the dream. Now you might think, in line with a venerable philosophical tradition according to which the mind is temporal but not spatial, that the real world transcends only the space, not the time, of the dream. You might think this on the grounds that although the space of the dream is illusory, it nevertheless takes place in real time; that the events of the dream could be synchronized with the clock above your sleeping head, for instance. However, this thought is confused, as I think must be recognised if we take illusionism seriously – illusionism without the physicalism, that is.

Whether or not this is a dream, the phenomenal presence of the screen I am looking at has a top and bottom. This spatiality is something that could only be found in the objective world, where the phenomenal presence does not belong; so the spatiality is illusory. But if this is a dream, then in order to think of the presence as at least causally connected with brain states, as I must in order to make sense of the possibility of altering it by causally affecting regions of objective space, I must misrepresent the presence as an occupant of the objective world. But then, it is equally true that to conceive the presence as in causal communion with my objective brain, I must think of it as within objective time; as in sync with the clock above my sleeping head. So given that this misrepresentational projection into the objective world gives us no reason to think experiences really possess spatiality, and that dream experiences take place in a unified space and time, I see no more reason to trust dream temporality than

spatiality when it leaves the context of the dream. The dream-presences simply do not belong to the objective world; the latter transcends both the space and time of the dream.

If this is a dream, then, any reality there is to my experiences belongs to a context that transcends the dream. I must project them into the objective world in order to explanatorily interweave them with that world, given what I know about it, but in doing so I misrepresent them as really having the spatiality and temporality which they seem to possess in the dream. Now if we take dream experience as indicative of how experience in general works, namely by enclosing us in an illusory experiential world from the perspective of which the independent reality of those experiences could only be found within a transcendent context, then we are ready to form what I call the ‘Transcendent Hypothesis’ (Tartaglia, 2016c, pp. 101-121). For then we will entertain the thought that consciousness in waking life works the same way. If it does, then the independent reality of our experiences is not to be found in the objective world described by science; flawlessly, at least in principle, and in no need of supplementation on the basis of philosophical speculation. Rather, it belongs to a context of existence which transcends the objective world.⁸

In that case, there is nothing we can cogently say about this independent existence, except that (1) it is transcendent (it is not accurately characterised as belonging to the world of objective space and time, though its existence is responsible for our finding this world⁹), and that (2) it exists. We can say nothing substantive about independent existence; but by misrepresenting it, and thereby thinking of experiences as the kind of things that might be systematically correlated with objective states of affairs, we have arrived at our objective

⁸ Thus the objective world transcends dreams, and independent reality transcends the objective world; independent reality transcends dreams too, but at one step removed. This extra step is required to account for the existence of dream experience, since dreams are immediately transcended by the objective world, where, as all illusionists realise, no kind of (phenomenal) experience can reside.

⁹ I work through the details of this in *ibid.*, pp. 101-167.

picture of the world in all of its astonishing detail.¹⁰ Freed of physicalism, it is no longer puzzling that we should form objective conceptions of individual experiences to facilitate causal explanations involving the impact of the world on these experiences and *vice versa*, and that once the picture was complete it had no place for our starting point. For a completely objective picture is centreless. This is only puzzling if we try to eliminate the starting point by making it a profoundly misleading revelation of objective reality's independent nature; for then we must engage in desperate *ex post facto* cherry-picking for the elements that might have been revelatory. But the starting point is not eliminated if it is transcendent; the illusion was that it had objective features. Our objective conception of the world is the end product of working out the implications of these supposed features. All of them, for we learnt about light and surfaces from so-called 'phenomenal-colour', and science will ultimately correlate all features of experience with conditions in the outside world, our brains, or both; despite the fact that experiences are absent from the world they are correlated with.

The fact that our objective picture of reality cannot describe independent existence, given the nature of consciousness, is a purely philosophical concern which connects with the core subject-matter of the discipline ever since Plato's invocation of transcendence (ibid., pp. 61-81; Tartaglia, 2016d). It is of no concern to the science of consciousness, which can continue to map the correlations between experiences (misrepresented as objective and reported accordingly) and physical conditions, without bothering itself with the philosophical implications of using the language of causation, identity, or misrepresentation. It can use whatever language is most convenient (surely causation), and ignore the notion that it is 'missing out' something; a notion pressed so successfully by some philosophers, that they

¹⁰ This includes our notion of representation; see ibid., pp. 147-67. This is not idealism, transcendental or otherwise. Neither is it phenomenal realism, since I am illusionist enough to think that our understanding of independent reality as consciousness is a misrepresentation. However *what* we are misrepresenting (as conscious states in causal communion with an objective world) is certainly real; which is why this account cannot be recast in fictionalist terms acceptable to a physicalist (thanks to Frankish for this suggestion).

have spread their own confusion about the nature of their discipline into science, such that many scientists are convinced they have a scientific challenge to overcome, or else, as is often the case, that they must denigrate philosophy for suggesting it has not already been overcome. For the only thing being missed out is the metaphysics, and there is a well-established discipline to deal with that. Science aims to perfect our objective representation of the world. But reflecting on the status of this representation within the context of transcendence is a philosophical concern.

Physicalism gained its ascendancy in the twentieth century, when the achievements of science became so hard to ignore that it came to seem that any serious investigation of the nature of reality must be scientific. Philosophy wanted to associate itself with science rather than religion, with which science was increasingly at odds; and traditional metaphysics seemed to be, and often was, on the side of religion. But religions make specific claims about the nature of the transcendent context; most typically that it is a meaningful one capable of vindicating human strivings. Such claims are necessarily baseless according to the transcendent hypothesis, which suggests only that the evolution of consciousness, as we conceive it objectively, enclosed us within an illusory phenomenal world which, when thus conceived in objective terms, allowed us to make maximal, objective sense of reality; as well as grasp philosophically that this sense could not capture reality's independent nature.

I think that what led philosophers to embrace physicalism, and the gut-instinct which sustains sympathy with it to this day, is a desire to align philosophy with science against religion and the supernatural. Much-touted motivations such as causal closure, by contrast, only seem powerful against the backdrop of Cartesianism; for closure rules out practically no traditional metaphysical positions except interactionist dualism. Once physicalism had been embraced,

due to a lack of historically-informed metaphilosophical self-consciousness which thereafter steadily increased, the remit of philosophy was radically reduced; for the core of metaphysics had been handed over to science. Philosophers began to think that claims not made on the basis of science lacked ‘conservatism’, as Frankish puts it (p. 12); thereby forgetting to conserve the insights of their own tradition. Philosophy continued to thrive nonetheless, but only because outside of philosophy of mind, little heed was paid to physicalism and its illusionist consequences, such that philosophers felt able to continue using the manifest situation as their starting point; while inside philosophy of mind, opposition to physicalism provided a continual feed of tasks for physicalists to respond to.

If physicalism were to win the day, philosophy would shrink to a point, and those who have been calling for its replacement by science ever since Comte would finally have their way. This is immediately clear in the case of philosophy of mind. For if it became generally accepted that science tells us what the mind *is*, there would only be so much interest to expend in the question of how best to formulate that insight; remove all the bats, black-and-white scientists, zombies and Chinese rooms from this area of philosophy, plus their predecessors in the form of misgivings about behaviourism and topic-neutral analysis, and the topic as we know it is decimated. Further remove all the reflections rooted in our manifest situation, and little or nothing would remain. After all, the idea of philosophy of mind as a distinct area within the wider discipline emerged only with physicalism (Rorty, 1982).

Why should this matter? Surely our only interest is in having an account of the nature of mind which is true. Two reasons it matters are as follows. The first is that physicalism’s consequences for consciousness are thoroughly problematic. Physicalism could be true if bare judgement can account for our manifest situation. But only physicalism suggests that it can,

and physicalism is badly motivated; mainly by inattention to metaphilosophy. Moreover, if judgements about consciousness were indeed evidentially baseless, this in itself would cast serious doubt on physicalism.

The second reason is that this debate is occurring at a time when the science of consciousness and artificial intelligence research is racing along; hardly a coincidence. Before long, we may have machines that, from the outside, appear to have minds just like ours. If they pull sad faces or make sad noises when we go to switch them off, we will have qualms. Now all philosophers want to account for the manifest situation. But the way physicalist illusionism does this provides an ideal justification for allocations of machine consciousness. For we can programme machines to represent the same kinds of physical things as us, and also to report phenomenal illusions; so if we believe that representation is all that is required to create the illusion of a manifest situation, our natural inclination to think they are like us so long as they act like us receives an intellectual stamp of approval. Without physicalism, we cannot know they share our manifest situation, because we cannot know that bare judgement is enough; *they* are making the judgements, so we cannot know how things seem to them, only how they tell us they seem. With physicalism, however, they must be conscious; because physicalism leads inexorably to illusionism.

This matters because if these machines are invented, they will have a profound influence on human life. Maybe it will be good, maybe bad; but how we should treat them, and whether indeed we should press on in inventing them, is a debate we need to have. Physicalism removes philosophy's voice from that debate. It removes the rational basis from appeals to our manifest situation; the intuitive way we have of thinking about our lives and world. With physicalism as an unquestioned background, there is no room to question whether they will

be genuinely like us when their behaviour and programming is objectively indistinguishable. With the conceptual space of transcendence closed down, this distinction cannot be made without disputing, rather than interpreting, the scientific world-view; and the latter is all philosophy can credibly do. So unless philosophy rapidly gets its act together, it will have no significant part to play in this evidently philosophical debate. Essentially, there will be no debate. Rather, anti-physicalist philosophers will be related to the side-lines, along with religious leaders, as the loony 'anti-science' faction, while the scientists developing the technology will do their best to make it safe ... and then we will all just have to wait to see what happens when the machines join us. The idea that there might be a distinctively philosophical perspective to be had on these issues, besides cheering or smearing science, will be passed over. This is what is at stake in illusionism: philosophy's voice in a debate where it might be persuasive and important.

Illusionism is a sound and important metaphysical insight, which undermines the pseudo-challenge of phenomenal realism to science. But it can be taken in different directions, and my suggestion is to pursue it outside of the self-imposed shackles of physicalism. Frankish says that 'evolutionary theorizing about consciousness can flourish, once freed from the metaphysical preoccupations of [phenomenal] realism' (p. 17). I do not think serious scientific theorizing about consciousness should give two hoots about metaphysics; if it does, then this is the result of either scientists dipping into debates that are not their concern, philosophers trying to exert undue influence on science, or – what I think is actually the case – both. Philosophers should stay aware of how scientific debates go, so as to inform their work on the status of our objective representation of reality. But they should not try to take part in them; not in their day-jobs, at least. I am all for joined-up thinking across disciplines, but not if this results in philosophy being incapacitated and science being distracted by

philosophical problems it cannot solve. So I prefer Frankish's sentence in an amended form: philosophical *and* scientific theorizing about consciousness can flourish, once freed from the metaphysical preoccupation of physicalism.

References

- Chalmers, D. J. (1996) *The Conscious Mind*, Oxford: Oxford University Press.
- Dennett, D. C. (1991) *Consciousness Explained*, Boston, MA: Little, Brown and Company.
- Frankish, K. (2016) Illusionism as a theory of consciousness, *Journal of Consciousness Studies*.
- Jackson, F. (1982) Epiphenomenal qualia, *Philosophical Quarterly*, **32** (127), pp. 127–136.
- Kripke, S. (1972) Naming and necessity, in Harman, G. & Davidson, D. (eds.) *Semantics of Natural Language* (pp. 253-355), Dordrecht: Reidel.
- Loar, B. (1990) Phenomenal states, in Tomberlin, J. (ed.) *Philosophical Perspectives*, *4: Action Theory and Philosophy of Mind* (pp. 81-108), Atascadero, CA: Ridgeview.
- Nagel, T. (1974) What is it like to be a bat? *Philosophical Review*, **83** (4), pp. 435-450.
- Place, U. T. (1956) Is consciousness a brain process? *British Journal of Psychology*, **47** (1), pp. 44-50.

- Quine, W. V. O. (1975) A letter to Mr. Ostermann, in Bontempo, C & Odell, S. (eds.) *The Owl of Minerva: Philosophers on Philosophy* (pp. 227–30) , New York, NY: McGraw-Hill.
- Rorty, R. (1982) Contemporary philosophy of mind, *Synthese*, **53** (2), pp. 323-48.
- Searle, J. R. (1997) *The Mystery of Consciousness*, London: Granta Publications.
- Smart, J. J. C. (1959) Sensations and brain processes, *Philosophical Review*, **68** (2), pp. 141-156.
- Tartaglia, J. (2013) Conceptualizing physical consciousness, *Philosophical Psychology*, **26** (6), pp. 817-838.
- Tartaglia, J. (2016a) Rorty's philosophy of consciousness, in Malachowski, A. (ed.) *A Companion to Rorty* (forthcoming), Oxford: Wiley-Blackwell.
- Tartaglia, J. (2016b) Rorty and the problem of consciousness, in Leach, S. & Tartaglia, J. (eds.) *Consciousness and the Great Philosophers* (forthcoming), London: Routledge.
- Tartaglia, J. (2016c) *Philosophy in a Meaningless Life*, London: Bloomsbury.
- Tartaglia, J. (2016d) Is philosophy all about the meaning of life? *Metaphilosophy*, **47** (2), pp. 282-303.