'TRUST ME, I'M AN ILLUSIONIST.' A CRITICAL RESPONSE TO KEITH FRANKISH'S ILLUSIONISM AND ITS PLACE IN CONTEMPORARY PHILOSOPHY OF MIND

PAUL STENNER

Introduction: the spell of the trickster

In the following I offer a critical account of Keith Frankish's piece on illusionism. Frankish is inspired by a paradoxical but prominent tradition within philosophy of mind that, crudely put, seeks to eliminate its subject matter ('mind', especially in its conscious aspects). The tradition is exemplified in the work of Daniel Dennett, who wrote an article in 2019 praising Frankish as follows:

An illusionist is an expert in sleight-of-hand and the other devious methods of stage magic. We philosophical illusionists are also illusionists in the everyday sense — or should be. That is, our *burden* is to figure out and explain *how the 'magic' is done*.

Dennett correctly observes that illusionism is the business of illusionists. Let me risk a generalisation: Illusionists are tricksters who are preoccupied with their own tricks and with the tricks of others. And they know that whatever a trick is, it is the effect of a carefully designed illusion, and hence not something *real*. The trickster tricks others but does not want to be tricked. An error typical of the trickster is that they would rather sacrifice a wider grasp of truth than risk making an error. The trickster casts a pessimistic spell.

A little history: machines and magic

To get to my point a little historical background is needed, so let me start with a story that begins less than 400 years ago. Seventeenth Century Europe, dominated by Christianity, was a world still reeling from the wars of religion that had decimated its political and cultural landscape following the reformation. From out of this liminal experience, with the growth of



311

natural philosophy, particularly in the field of physics, a new idea was gaining ground. The universe, our solar system, and all that it contains, is a gigantic machine of a purely physical nature, determined like all physical things by natural laws. When this idea was articulated by Galileo and refined by Newton, both geniuses presupposed that this machine could only have been created by a divine mechanic, so precise are its workings. Both took their contributions to physics as testimony to God's existence and nature. These beginnings of modern science were theological in a new sense, and thoroughly political. Questions remained as to whether God continued to intervene and otherwise interfere with the physical machine, or whether, having completed his task, he left it to its own devices. Questions also remained as to whether the human soul is a little piece of God's sovereign agency distributed amongst his favoured creation the better to consciously appreciate his divine work. For example, another seventeenth century physicist-philosopher, Descartes, argued that whilst plants and animals are nothing but machines, the machine bodies of men (or at least white European men) are also inhabited by souls (un-extended 'thought substance'). For Descartes, this exceptional feature gave mankind natural sovereignty over all other life and nature, which is purely machinic.

Time, as it does, moved on. In the eighteenth and especially nineteenth century this idea of mankind as a soul-bearing part of a world-machine made by God started losing ground. The clever people now began entertaining the opposite proposition: that God is the creation of the functionings of the human machine. To many people, physicalist materialism had come to prove so powerful that there was no longer any need to entertain old-fashioned nonphysicalist notions like 'soul' and 'spirit' or even 'thought' and 'mind'. All of that can be explained as part of the operations of the purely physical machine. Even the quaint notion that we are each a 'person' – a coherent being capable of feeling and thinking things through, coming to decisions and acting on them, etc. - comes then to be seen as a curious illusion: a mere effect generated for functional reasons, by the brain of a clever machine. This raises some deep paradoxes. The one who 'believes' these things must also be a purely physical machine and so cannot in fact believe them. The materialist's beliefs too must be creations of a purely material machine enjoying the illusion that it is a person with beliefs. Keith Frankish thus cannot really believe his own thesis since this would be incompatible with his illusionist philosophy. So even here he is a trickster creating the false impression that he entertains an illusionist theory. In making this argument, and in calling him an illusionist, a trickster, I am not critiquing him, but acknowledging that this is the only position he can authentically adopt given his premises. Like him, I do not believe Frankish's premises, and nor do I agree with them. But can we break the spell that he casts over us because it has been cast over him?

The Pledge, The Turn and The Prestige

Let me put it differently. Is it wise to trust someone who tells you straightforwardly that they are a trickster type? A magician no less? Perhaps it is sensible to trust them to show you a trick, and some magicians may even tell you how the trick works, whilst some keep their secret carefully hidden. Christopher Nolan develops this theme beautifully in his 2006 film *The Prestige*. Two rival stage magicians in London compete to perform the 'transported man' trick. A man *seems* to be instantly transported between two cupboards at opposite ends of

the stage. The character played by Michael Cain explains that a magic act has three phases. First is The Pledge when something ordinary is presented, like a bird in a cage. Second is The Turn, when the ordinary becomes extraordinary. To continue the example, the bird is made to disappear. Third is The Prestige, when the magnificent magician re-establishes normality: the bird reappears. The skill of the magician, as everyone knows, is to manipulate impressions. To the audience it looks like the bird has disappeared and reappeared, but actually the cage has been designed so that in a swift movement it can crush the bird pancake thin. The cage also allows entry to a second bird carefully selected to exactly resemble the first bird, now dead. A main message of Nolan's film concerns desire: the audience wants to be tricked. That, after all, is why they have paid good money to see the show. They are tickled by the fantastical nature of the trick. The more unbelievable and counter to common sense, the better. But of course they will quickly spot it if the trick is not expertly enacted. This is why ever more technical and complex tricks are required, some of which may be dangerous to the assistants and stooges on stage. Certainly the magician knows it is all a trick, which is why they guard the techniques so secretly. The whole idea is to entertain by deception.

It is quite instructive to treat Frankish's interview text as if it were a magic act. First comes The Pledge. He opens his act with something ordinary: 'I don't think there is anything special about it [the problem of consciousness], I don't believe it presents a hard problem... radically different from the problems we face in explaining other psychological and biological processes'. Nothing unusual here, just a helpful philosopher observing jobbing psychologists doing their business as usual.

Second follows The Turn, when the reader is invited to entertain something extraordinary. It is common sense both to ordinary people and to most philosophers and scientists that when we examine our experiences we think of them as private and subjective: an ache in our back, savouring the scent of a tulip, or the luminous colour of its petals. But, 'ta da!', actually this 'phenomenality' is nothing but an illusion created by the hidden functions of a number of brain systems. We possess a 'psychology which disposes us to believe in mental qualities, even though they are not real'. Illusionism in philosophy is needed to clarify how our brains create the illusion of having private subjectivity. The full Turn takes a number of pages and involves building some quite elaborate metaphysical castles in the air.

Finally, to end the trick normal service is resumed with The Prestige. After all, philosophers must allow psychologists to 'get on with the task of trying to explain the reactions'. Indeed, his own air castle building complete, Frankish discourages others from building metaphysical castles in the air lest they distract scientists from this important task of explaining the biological basis of consciousness, about which we should now have lost 'our intuition that there is more to explain, or at least cease to trust it'.

The motive: bolstering the crumbling edifice of materialism

But why would Frankish want to undermine our trust in the idea that we have private subjective experiences? Why would he want to make what seems obvious disappear? As with any good trickster, he does not let his motives loom large in the foreground but keeps them

safely off-stage. But the stakes are clear to a careful observer of tricksters. He wants to reassert a doctrine of physicalist materialism against what he perceives to be a growing number of dissenters and 'radicals'. The doctrine he defends has become ever more implausible since the time of Galileo, and resorting to 'illusionism' is one of the few ways left to defend it. Frankish's 'physicalism' is the assumption that all of psychology can be reduced to brain processes and, as he puts it, 'explained in physical terms' and 'fully characterized in thirdperson terms' (p.9). The physicalism he assumes is that of Galileo's machine, but unlike Galileo, Newton and Descartes he does not see the need for God, the soul or any other 'non physical' stuff. His physicalism is uncompromising: physical matter is all there is. The rest is illusory 'magic'. Non-physicalist descriptions and explanations are thus to be rejected a priori and all scientific efforts encouraged to adopt an exclusively physicalist approach. There is nothing else. Psychology is his main target, but this physicalist position extends to the social sciences and humanities (because it extends to everthing). As he puts it, he holds the view that 'everything is constituted by the fundamental entities posited by physics'. It is accepted that psychology, let alone history or sociology, has not yet been able to provide physicalist explanations of its subject matter. Some would even say that this path has been a barren dead-end and that only research that has deviated from it has proved fruitful. But regardless of this, Frankish wants it to be established that it must carry on along this same physicalist path. A trickster must please the crowd, and so this dogmatism must be presented as its opposite. As he puts it 'I'm not dogmatically committed to it. I'd be willing to treat consciousness as fundamental if there were a strong enough theoretical case for doing so'. His trick, however, is to define in advance that only his preferred type of physicalist account can count as a 'strong enough theoretical case'. All other theories are to be made to disappear, replaced by physicalist substitutes. After all, he has already decided that only the physical is.

Here a detour is needed for a couple of paragraphs into a quite complicated matter in the history of science and psychology. Either bear with me or, if it gets too abstract, skip ahead. Frankish's 'physicalism' is nothing original, and nor is his trick of using theory to impose limits on what can count as scientific theory. During the 1920s and 1930s Carnap and others famously asserted a 'logical positivism' stipulating that for any statement to be scientific it must be expressed in physicalist language permitting the application of quantitative values attached to definite coordinates in the space-time system. Logical positivism thus proposed a comparable restriction of scientific propositions to those dealing with directly observable existents and couched in a model of theory-building based on a highly theoretical understanding of reality as a purely physical unity displaying demonstrable regularities. One reason it failed is that the 1920s and 1930s also happened to be the time during which physicists conceptions of the nature of basic physical entities and of the nature of the space-time system went through an unprecedented transformation. Building on James Clerk Maxwell's unified theory of electric and magnetic fields, major thinkers like Albert Einstein and Niels Bohr found it necessary to re-figure basic physical entities as events within a space-time system re-thought as a field. At the same time, this transformation in basic physical concepts was accompanied by a transformation of cosmology in the philosophy of physics, at least amongst those philosophers competent enough to understand the actual science (see McHenry, 2017). Whether one considers A.N. Whitehead's The principle of relativity, Bertrand Russell's *The analysis of matter* or Ernst Cassirer's *Zur Einsteinschen Relativitätstheorie* the position is remarkably similar: far from being composed of irreducible physical stuff located in the coordinates of a fixed set of laws, fundamental reality is composed of a processually unfolding *field* of *events*.

The problem is not just that logical positivism was already out of date when it was first proposed (and, indeed, somewhat imposed via the Chicago School). The problem was that its advocates were not aware that the old physicalist ontology they had taken for granted as simply 'real' was actually not brute reality but a high-level theoretical abstraction. In using this newly challenged theory to define what counts as scientific theory (and to exclude any theorising that does not follow its pattern), they were actually limiting scientific possibilities by imposing the errors of their own thinking. Retrospectively this explains why logical positivism was such a disaster for psychology, yielding labyrinthine and ultimately sterile theoretical programmes like that of Clark Hull, and legitimating some of the cruder simplifications of behaviourism. The richest and most fecund psychological theory of the time came from the Gestalt psychologists who, being directly acquainted with relativity and quantum physics, were confident enough to ignore the positivistic philosophers and articulate their own field-theoretical accounts of psychological events. The main point I want to make here is three-fold. First, that when Frankish claims not to be 'dogmatically committed' to physicalism and to be open to any proposition with a 'strong enough theoretical case', he speaks knowing full well that his trick is to have set the rules in advance so that only physicalist arguments can count as theoretically 'strong enough'. Second, he borrowed this trick from the logical positivists. Third, that the real lesson is to understand that what might at first glance appear to be established empirical fact can actually be a distortion created by high-level theoretical abstractions. We must be wary of what we take to be 'ultimate truths', not least when those truths concern complex concepts like physicality, and especially when we are tempted to impose these on others.

'Not like that, like this!': substituting 'qualia' for subjectivity

Detour over, let me return to Frankish's trick. The plausibility of his argument relies upon an illusionist 'sleight of hand'. The one I will dwell on here is a *substitution*. The trademark joke of the British comedy magician Tommy Cooper was to clumsily expose a trick to the audience saying: 'not like that, like this!' Frankish's substitution is analogous to the *disappearing bird trick* involving the swift and indiscernible replacement of the dead bird with a live one. What is it, in fact, that Frankish is telling us is an illusion? Sure, he says clearly a number of times that the central question is 'whether conscious experiences have phenomenal properties'. But what does 'phenomenal properties' actually mean?

When he first introduces it, Frankish uses quite broad terminology: we tend to think that our experiences 'have a private subjective aspect', but this is an illusion. This illusion brings with it the idea that, being private and subjective, consciousness 'can't be explained in third-person scientific terms'. The audience gasps in astonishment at this broad claim: what! No subjective experience! Impossible! The assertion that we cannot doubt the experience of our own subjectivity has been accepted as fundamental to modern philosophy ever since Descartes used it as the one certain thing in an otherwise completely uncertain world! The

trickster tormenting Descartes could make him doubt the very existence of the world outside him, but it could not make him doubt the fact that he doubted. The subjective operation of doubting is real and proves the existence of the doubter. Even if the doubter ceases to doubt, it is still the case that they *did* doubt and that now they *no longer doubt*. And if they start to doubt anew, this too becomes a new fact that in no way disqualifies what came before or what will come.

And while the audience is absorbed in its own astonishment at this breathtaking Turn, Frankish makes a quick substitution (achieved using brackets – which is quite hard to do in an actual verbal interview, if it ever happened...). Not like that, like this! Broad subjective phenomenal properties are switched to something much narrower and more defined: qualia or private mental qualities '(also called "qualia"; I'll use the terms interchangeably)'. The big question of subjectivity is replaced by something much more thinly abstract: now it is 'mental qualities' that are not real and that we must cease to believe in because they are illusions. But the new claim is much smaller and very different. What is a 'mental quality' and who believes in them anyway? What would a belief in 'mental qualities' even look like? When Frankish tells us things start to look very suspicious indeed. When gazing 'at a cloudless sky', those who believe in mental qualities believe 'that the blue quality we are aware of is actually a private mental quality'. Hmm. If a companion turned my attention to the sky and said 'Paul, that blue is a private mental quality', apart from worrying that I was in the company of a pseudo-philosopher I would probably say, assuming they were not hallucinating the blue, 'yes, but it looks blue to me too'. I might wonder: is she getting me back for pointing out during last night's sunset that the sun was not in fact going down? Whilst I know plenty who would protest against a denial of their subjective experience, I know of nobody who is committed to insisting that 'qualities' are purely 'mental'. On the other hand, every school child is taught that the colours they experience are the effect of a complex interaction between their visual systems and reflected light from the world beyond. Take away either the light or the eyes and the blue becomes black. Only those who have no time to think about it would conclude on this basis that the blue is just a private mental qualia existing 'only in our minds'. We assume that other people also see the blue sky and think of them as colour blind if they don't (or looking in the wrong direction). And yet Frankish wants us to think that this is 'a natural conclusion to draw'. It is not a natural conclusion and, because of its patent epistemic fragility, those who draw it can easily be disabused of it.

So the trick is to subtly switch something astonishing because it challenges everyone's basic assumptions (subjective experience is an illusion!) for something mundane that nobody is actually committed to believing anyway ('qualia in our minds' are illusions!). The value of this substitution for the illusionist comes from a double transfer of meaning. First, for those who don't notice what he has done, the *certainty of experience* from the initial claim is carried over to the second, which as a result gives the impression of being surprising. And at the same time, the *uncertainty of experience* from the second is carried over to infect the certainty of the first, so we come to imagine that subjective experience is as doubtable a proposition as a theory of mental qualities. How does Frankish exploit this double transfer of meaning? Well, *that* we experience the blue colour is actually something quite concrete and

certain. It is very hard to make us doubt it. If the sun then sets and the sky is now orange, this does not alter the fact that it *was* blue but *now* is orange. Soon, when the light fades, it will be black. All the experiences of colours and changes are very concrete and hard to doubt. But, if we are fooled, we do come to doubt it because without noticing the substitution, we are called upon to doubt something quite different masquerading as concrete experience. The theory of 'qualia' is not a simple statement of experienced fact but an abstract theory. It is not about doubting or affirming whether or not we experience blue, but rather about calling into question the assumption that the blue *belongs exclusively to the sky*. As the growing child learns something about science they learn that the blue they see does not belong *exclusively* to the sky but is an experience that arises in themselves qua observer as a function of an interaction between their visual system and light from their environment. This *theoretical knowledge* is inherently more doubtable, but considerably more doubtable is the proposition that these so called qualia 'exist only in our minds'.

Frankish gives us this very doubtful and indeed 'unnatural' proposition (things called 'qualia' exist 'only in our minds') and tells us that this is 'a natural conclusion to draw'. He can then 'surprise' us with his cleverness by pointing out that in fact it is not the only possible conclusion to draw. Indeed, it is at this point that he pulls his illusionist theory out of the hat as the true alternative to this 'natural conclusion'. But only somebody suckered by his trick would take the proposition 'qualia exist only in our minds' as anything other than highly questionable if not plainly false.

Galileo's real real error

It is at this point that Frankish refers his reader to another of his papers entitled 'Galileo's real error'. If you follow up the reference you find him arguing against another philosopher (Philip Goff) who argued – according to Frankish at least – that Galileo made an error in 'denying that sensory qualities are present in the physical world'. Frankish urges instead that Galileo's real error was not simply to get the location of sensory qualities wrong (by insisting that they exist only in the mind), but to assume, apparently like most people, 'that sensory qualities are instantiated somewhere'. The real error was to assume that qualia exist anywhere at all. And hence we return to the core illusionist thesis that 'qualia' are illusions because the only reality is physical. This is paradoxical if we take into consideration what Galileo actually argued. He actually argued, as part of his discussion on heat in *Il Saggiatore*, for a distinction between those qualities that would no longer exist in the world if sense organs were removed, and those qualities that would continue to exist. It is worth paying close attention to what Galileo actually wrote about the tastes, smells, colours and so forth back in 1623 (Chapter XLVIII):

Per lo che vo io pensando che questi sapori, odori, colori, etc., per la parte del suggetto nel quale ci par che riseggano, non sieno altro che puri nomi, ma tengano solamente lor residenza nel corpo sensitivo, sì che rimosso l'animale, sieno levate ed annichilate tutte queste qualità; tuttavolta però che noi, sì come gli abbiamo imposti nomi particolari e differenti da quelli de gli altri primi e reali accidenti, volessimo credere ch'esse ancora fussero veramente e realmente da quelli diverse. (p. 199)

Ma che ne' corpi esterni, per eccitare in noi i sapori, gli odori e i suoni, si richiegga altro che grandezze, figure, moltitudini e movimenti tardi o veloci, io non lo credo; e stimo che, tolti via gli orecchi le lingue e i nasi, restino bene le figure i numeri e i moti, ma non già gli odori né i sapori né i suoni, li quali fuor dell'animal vivente non credo che sieno altro che nomi... (p. 201)

These tastes, smells, colors, etc., with regard to the object in which they appear to reside, are nothing more than mere names, and exist only in the sensitive body; insomuch that when the living creature is removed all these qualities are carried off and annihilated; although we have imposed particular names upon them ... and would happily persuade ourselves that they truly and in fact exist. But I do not believe that there exists anything in external bodies for exciting tastes, smells, and sounds, but size, shape, quantity, and motion, swift or slow; and if ears, tongues, and noses were removed, I am of the opinion that shape, quantity, and motion would remain, but there would be an end of smells, tastes and sounds, which abstractedly from the living creature, I take to be mere words.

Galileo, obviously, was neither a psychologist nor a biologist. If he had been he would have thought twice about the idea that he could get at the truth of things like sounds, tastes and smells by amputating ears, tongues and noses. Equally obviously, sounds, tastes and smells DO exist when those organs ARE intact and in contact with the appropriate features of the environment. Galileo accepts this when he proposes the thought experiment of 'removing' the 'living creature', for he cannot remove what doesn't exist. There is an ambiguity here, however, because he is simultaneously asserting two things, the first of which looks quite 'illusionist'. On the one hand he asserts that sounds tastes and smells are 'mere words' and on the other hand that they 'exist only in the sensitive body'. Galileo has clearly not thought this through. It is clear that annihilating only words like 'smell', 'stink', 'fragrance' etc will not get rid of the actual smells, although it might surely influence the capacity of the sniffer to make fine grained distinctions between complex bouquets. Removing the sensing parts of the 'sensitive body', by contrast, will be a much more effective operation, but the mutilations would be of real existing things. So we can conclude that Galileo got a bit over excited about his new perspective when he allowed himself to write that these sense experiences are 'nothing more than mere names'. In fact he always checks himself and asserts the relevance of the living, sensitive creature: only when taken 'abstractedly from the living creature' are they 'mere words'.

The organism tastes and smells food with its tongue and nose. In the same way food – a tasty strawberry for example – is not food without an organism that wants to eat it. Food comes into real existence along with its eater, just as its taste comes into real existence with the tongue of its taster. Animals with tongues and noses DO exist in the world, even if a physicist like Galileo can make his calculations and devise his theories of physical motion without ever taking them into consideration. The other qualities he discusses, by contrast, cannot be discarded and are vital to physics. Shape, quantity and motion are qualities that can be successfully described in terms of mathematical quantities. They, from Galileo's exciting new perspective, can provide the basis for a purely physical science, whereas the secondary qualities just discussed lack this feature. The secondary qualities can, however, be explained

by means of the primary qualities. The secondary qualities can thus be discarded from his new modern *physical* science where previously it had been assumed that they were real properties of things.

In fact, what Galileo called into question was an entire metaphysics of substance that had been assumed since at least the times of Plato and Aristotle, built as it is into most languages. This metaphysics assumed that reality is ultimately composed of substances qualified by attributes: a red rose, a grey stone. The tulip was understood as a substance qualified by a number of attributes such as its distinctive scent. The notion of 'qualities' that Frankish makes so much of derives from this metaphysics because 'qualities' are the attributes that qualify a given substance. Galileo's thought challenges this metaphysics, but only partially and inadequately. He was a great mathematical physicist but not such a good philosopher. He shows that qualities like colour, smell and taste, unlike qualities like shape, size and motion, do not actually belong in any straightforward way to substances. He thus challenges the 'attribute' side of the 'substance/attribute' metaphysics but he retains the substance side. Also since Plato and Aristotle, substance had been defined as something that exists in an entirely self-contained manner, dependent upon nothing else. The colour of the stone may change with accidents of light, they reasoned, but the stone substance itself is timeless and sovereign. Unfortunately, a downside of Galileo's otherwise excellent influence is that this substance metaphysics gets further compacted and buried deeper. This is because Galileo develops a highly limited theory of what physicality actually is. Note, this is a theoretical abstraction. It has nothing to do with antics like banging one's fist on the table and saying 'that is what is real!' As the discussion above clearly shows, his theory is limited by the fact that it deliberately excludes 'sensitive bodies' from consideration. And yet Galileo himself was just such a sensitive body, making sensitive judgements about the world around him. Of course he knew that he existed, and he knew that other living creatures were real, but he excluded these from his new theoretical concept of nature. This left him with the entirely dead vision of nature as a machine that was taken forward so effectively by Newton, Descartes, etc. He could avoid thinking this problem through because he could simply assume that this machine had been built by God. Anything 'sensitive', in other words, could be consigned to an entirely different ontological realm.

Descartes did the honours of formalising this as a self-conscious metaphysics. And neither did Descartes get rid of the basic notion of substance. On the contrary, he duplicated it into his distinction between extended substance and thought substance. But the entire metaphysics is clearly deeply problematic. Understanding this took several more developments in science (not least the development of biology) and philosophy. Articulating a plausible metaphysical alternative took still longer (see the discussion of fields and events above). Galileo's real error was certainly not that he failed to purge his concept of physical matter of 'qualities'. It is that, brilliant though his discoveries were, he accepted a metaphysical distinction between primary and secondary qualities that allowed him to retain what would prove to be an untenable notion of physical materiality as self-contained substance. It is that notion that, for reasons known only to himself, Frankish is intent on preserving. But he can do so only by trickery, because – like Harlequin's cloak – its plausibility as valid science and philosophy lies in tatters.

References

Frankish, K. Illusionism and its place in the contemporary philosophy of mind. Interview with Keith Frankish by Katarína Sklutová. *Human Affairs*, 32(3), 300–310.

Frankish, K. (2021) Galileo's real error. Journal of Consciousness Studies, Autumn.

Galileo, G. (1623) *Il Saggiatore* [The Assayer]. FreeEditorial. https://freeditorial.com/en/books/il-saggiatore/related-books

McHenry, L. (2017) Whitehead and Russell on the analysis of matter. *The Review of Metaphysics*, 71, 321–342.

School of Psychology and Counselling Faculty of Arts and Social Sciences The Open University Gardiner 2 Building Walton Hall Milton Keynes MK7 6AA United Kingdom

Email: paul.stenner@open.ac.uk ORCID: 0000-0003-4505-6422