

Rationally Speaking #242: Keith Frankish on “Why consciousness is an illusion”

Julia Galef: Welcome to Rationally Speaking, the podcast where we explore the borderlands between reason and nonsense. I'm your host Julia Galef, and my guest today is Keith Frankish.

Keith is a philosopher based in Greece. He is in the Department of Philosophy at the University of Sheffield. His specialty is the philosophy of mind, and he's also the author and editor of several books on philosophy of mind, specifically on consciousness, including the book *Illusionism*. This is probably what Keith is most well-known for, being a proponent of a theory of consciousness that he's dubbed illusionism.

We're going to be talking about what illusionism is, and the argument for it -- and also I'm hoping that Keith can help me kind of get a handle on the landscape of theories of consciousness. Why people disagree with each other, what could possibly settle the disagreement. I have for a long time felt very lost in trying to untangle the debate over consciousness.

So that is what we're going to be talking about today. Keith, welcome to Rationally Speaking.

Keith Frankish: Hi, Julia, thanks for inviting me.

Julia Galef: You know, every six months or so I say to myself, "I should really do an episode on consciousness, theories of consciousness," and then I'm like, "Well, let me first do enough reading so that I really understand it well enough that I can ask really good, clear questions" ... and that has just never come true.

So rather than putting it off indefinitely until we've settled the debate by uploading our minds onto computers, I figured I would just kind of pop one level out, one level meta in the debate... and in addition to asking you about illusionism, have you help me get a handle on how the debate has been shaping up. But we'll get to that.

Let's start by talking about illusionism. When I first heard of illusionism I thought that you were claiming that my experience of consciousness is illusory. Like, "You, Julia, think you're experiencing things, but really you're not," which seemed very implausible to me on the face of it. But after reading a little more it started to seem like I had misunderstood what you meant by illusionism, is that right?

Keith Frankish: It all depends.

Julia Galef: Or does this all hinge on the words --

Keith Frankish: It all depends, it all hinges on what you mean by consciousness.

Julia Galef: Of course it does.

Keith Frankish: Perhaps the simplest way to put it is that I'm offering a different model of what consciousness is, a model which rejects a very central model, a model that dominates most people's thinking about consciousness. And it says that, that model is wrong, and that consciousness in that sense is illusory. It doesn't exist *in that sense*.

It doesn't mean that we stop talking about consciousness, it means that we've introduced a different way of thinking about it. But this model is so central -- the one I want to dismiss is so central -- to some people's thinking about consciousness that it seems that I am saying that consciousness doesn't exist, that it's illusory.

... The idea is that while I may be wrong about the external world -- I may be wrong that there's a cup in front of me. I may be wrong the cup is blue, perhaps my eyes are playing tricks on me, perhaps the lighting's strange. Perhaps I'm hallucinating. Perhaps I'm even in The Matrix, and there isn't any external world at all.

But I couldn't be wrong about the character of this internal world. I couldn't be wrong that I'm having an experience with this kind of blue feel to it.

Julia Galef: It seems to me that you could be a *little* wrong. In the sense that there are all these visual illusions that are built into our brains, and our eyes I guess -- where I think that I'm perceiving color up to the edge of the periphery of my vision, but actually I'm not. And so I'm wrong about that.

So it seems to me that we could be wrong at that level, or to that degree, but it would be really hard to understand how we could be wrong about having an experience at all that's kind of "blue cup"-related.

Keith Frankish: I think that's an excellent point. And I think you've already pointed out here that our common sense view of what our inner experience is like is not as solid and reliable as we think. Because as you say we tend to assume, [our] representation of the visual field is complete in every detail, right out to the periphery. That's how we kind of judge it. But that turns out to be wrong, that is an illusion.

But that's an *introspective* illusion. Even in The Matrix scenario we would still be having that illusion of having this complete visual field, even though there's no world out there. The *inner* world isn't as it seems.

Now once you allow that then I think you're opening a wedge [to allow] the idea that introspection might itself be a kind of construction, a construction of the brain.

Julia Galef: Well, I was trying to say that even though you are opening a wedge, it does not seem to me that that wedge could be big enough to dislodge the basic fact that I'm having an introspective experience.

Keith Frankish: It certainly doesn't seem that way. But I think you've opened a wedge, and once we see the reasons for sort of pushing at this, at this wedge, and once we see what alternative stories we can have, then I think the big picture begins to look a bit more plausible.

But let me say a bit more about what I call the "realist" picture first. Because I want to stress how odd this picture is.

Some characteristics of this internal picture -- Daniel Dennett calls it a "Cartesian theater." This idea that there's this kind of inner display of experience with conscious awareness. So the outer world sort of affects us, it creates this private cinema screen, as it were.

Now first of all, who are "we" here that is witnessing this? Who is this inner "I" that is witnessing?

Julia Galef: I'm the little character sitting in the theater!

Keith Frankish: In the theater, that's how it seems.

Julia Galef: Yeah.

Keith Frankish: I think that this sort of view, that there is this introspective display, does presuppose some kind of introspector at all. So that's one thing that needs to be cashed out.

And the relation between this introspector, and what they're introspecting -- between the spectator and the presentation on the screen -- is a very, very intimate one. Because while I might be wrong about the cup out there being blue, or there being a cup at all... I *couldn't* be wrong, it seems, when I'm actually attending directly to it.

This is a bit of the intuition that you thought stood up. If I'm directly focusing on what the central area of my experience is, right now, and it seems it's kind of blue -- vivid ultramarine blue, say -- I couldn't be wrong about that. Maybe, as you say, the edge of my visual field is another aspect [about] which I might be wrong.

But this, what I'm actually closely attending to right now, I couldn't possibly be wrong about it. It seems to be presented to me directly. Transparently. Immediately. In a way that doesn't allow for any possibility of error.

Now I'm not saying that this realist picture couldn't be true. I'm saying I don't see how it could be true in any kind of mechanistic picture of what the brain is. You'd have to have some kind of metaphysical speculation here.

I mean, obviously for Descartes, it was quite easy because the thing that was witnessing all these was an immaterial soul. And the idea that an immaterial soul has some kind of special privileged access to its own state -- well, if it's immaterial, all bets are off, I suppose.

Now most people, most philosophers, including realist ones, think that, "Well, it is just the brain, we're not two things, we're just one thing. We're just an embodied brain." But somehow the brain has these special properties that are kind of presented to the person constituted by the brain in a specially immediate way.

I find that hard to understand. I am not saying you couldn't produce some metaphysical theory that would underwrite it, but I find that... can we do it without that first?

Here's another reason for being kind of cautious about these properties. And cautious is all I want to be, really. It's that they're private, and private in a really deep sense. Obviously in a sense everything that happens in my brain is private -- it's locked away in my skull, and you'd need special instruments and so on, to detect what's going on in there.

But these things are private in a much more radical way. Neuroscientists could map every single neuron firing in my brain, but they would never detect the *quality* of the experience that is produced by it. It's always remained conceivable that my experience of the world -- my color experience, say -- is different from yours.

We both agree that that's blue, and that's yellow, and those are the labels we put on things of that color. But what is that color like to

me on the inside, subjectively? What kind of qualia are produced by that color? What does coffee taste like?

Julia Galef: Yeah, but why is that suspicious, though? I can see how that's *inconvenient*. It makes it very hard to understand and theorize about consciousness. But it sounds like you think it's *suspicious*.

Keith Frankish: It makes it unlike any other aspect of the world, I think. Certainly unlike any aspect of the world that can be studied by science. Because the whole point about science is that you try to... you have instruments, you try to triangulate things, you try to compensate for individual bias, and get a picture of what things are like in themselves. This requires having multiple perspectives on the same thing.

The trouble with these things is that there's only one perspective on them, the perspective of this mysterious thing in my head. And now that means that these are unlike the rest of the natural world, and you can't do science on them.

Julia Galef: I see, so that ... Yeah. So the reason it's suspicious then is just, it seems unlikely that there would be this thing in the world that is unlike everything else, and can't be studied?

Keith Frankish: That it would just happen to be something, an aspect of us that we kind of focus on when we start to think about our own mental processes in a kind of reflective way. That I think, too, is suspicious.

I mean... We are pretty complicated creatures by anybody's account. We have some sort of self-awareness of our own mental processes. And it wouldn't be surprising I think, if that picture was not wholly accurate.

Nature hasn't equipped us to be super neuroscientists or to have a super understanding of our own mental [states] -- we don't need that. We maybe need some sort of rough and ready understanding of what are brains are doing, or what kind of states we are in. But we don't need to know it in detail. So maybe it's provided with something much more sketchy, and caricatured, and distorting.

Here is a way of putting it that I used in one paper. What everyone should agree, I think, is that these properties -- these qualia, phenomenal properties -- they're different words and they're used with different connotations, so I'm kind of smearing over those here. Some people might want to say that they matter. But I don't think they do, to the big picture.

These properties are anomalous. They're odd, they're unusual. I think everyone could agree with that. They're not like other properties of the human body -- properties like digestion, respiration, reproduction -- which we've got a pretty good grip on. They're not like other mental properties like emotions, insofar as they don't involve consciousness. We have good cognitive accounts of what's going off there.

These seem really anomalous. There seems nothing else like this. Now there's kind of three approaches we could take to that:

One is to say, "Yes, they seem anomalous, and they are anomalous, and we've got to do some really radical theorizing in order to account for them." We have to do some heavy duty metaphysics. That these things are not part of the physical world, but are kind of some non-physical aspect of the world, perhaps.

And this is a suggestion that is gaining popularity, that they are something like a fundamental feature of physical reality. The intrinsic nature of all matter is conscious in this way -- all matter has this intrinsic phenomenal aspect to it.

Julia Galef: Are you saying that approach is gaining popularity among philosophers or in popular culture?

Keith Frankish: Yes, yes, [among philosophers].

Julia Galef: Oh, philosophers, okay. Interesting. What's the name for it?

Keith Frankish: Panpsychism, because it's "psyche" "everywhere."

It's actually something like pan-phenomenalism. Because what is supposed to be everywhere is not thought and intelligence and emotion, and all of these things, but a feeling --

Julia Galef: Like, the raw building blocks.

Keith Frankish: Exactly, and so the idea here is that physics tells us what the electron does, but it doesn't tell us what an electron *is*. If you were to get down there, and sort of look at an electron, what would you see? Would it be colored? What it's like in itself?

... The very conception we have of the intrinsic nature of things is the intrinsic nature of ourselves, what it's like to be us. This phenomenal aspect. And so maybe it's the same for an electron. An electron has its own sort of private, kind of subjective nature, which is a bit like --

Julia Galef: That seems incredibly implausible. How would an electron have an internal... Anyway, I'll let you continue. I'm just registering that that sounds very implausible.

Keith Frankish: Well, it's a very simple form of it. It's not subtle like tasting a particular wine or something, it's not [rich], but it has a very primitive form of phenomenality.

Then the idea is that our phenomenality is somehow a construction from the phenomenality of all the molecules that compose our brains. In the way that cognitive functioning is a consequence of the dynamics and structure of all those physical... So you put things together in the right structure and the right way, and you get the dynamics of the brain. You put the phenomenality of the bits in place, and you get the consciousness of the whole brain.

Julia Galef: I feel like this is – like, in “hot and cold,” I feel like we're getting colder. We're making the world more confusing, with this.

Keith Frankish: I think there's one way in which these people are right, which is that you can't get realist about consciousness -- in this phenomenal sense, in the sense that I think is the bad sense -- you can't be realist about that without paying some quite heavy metaphysical price. If you really want to be realist about it you've got to fit it into the natural world somewhere, and it doesn't fit in easily.

So now maybe that's one way to get it into the natural world. You've got to find a place for it somehow.

Maybe it just pops into existence when you get complex enough brains, that would be another. That's sort of a form of emergentism. Nature starts building brains, and you get small brains, and they don't have this internal aspect to them, they are just information processing mechanisms. They get bigger and bigger, and bigger, and at some point between the first organism and us, the lights came on inside. And with all that complexity, all that information processing, which was doing all the work, the lights came on and there was an interior aspect.

Now then the question is, "Where did that happen?" There's absolutely no way of knowing because we cannot tell which other creatures have this other than [humans] -- strictly speaking we can't be sure anyone else has it apart from you.

So there you have a kind of arbitrariness. At some point the lights come on, in the natural world. Well, at least panpsychism has a consistency.

Julia Galef: Is it really more arbitrary than defining life? There is a bit of fuzziness at the boundary of what we consider life – like, viruses are not really life. The world is kind of messy.

Is consciousness worse on that score?

Keith Frankish: I think it is worse, in two ways.

One is that there doesn't seem to be a sort of in-between condition of, "Just a little bit of an interior world." Either there is something it's like to be, either it has this first person perspective, or it doesn't.

It may be a very impoverished perspective, and there's only sort of something very boring that it is like to be an electron or an amoeba, or whatever it is, but there's still either something it's like -- it either has this first person perspective -- or it doesn't. It's hard to see how I could have sort of "half a perspective." It's either like something or it's not. The inner lights are either on or off.

Second, there's just no way of -- I mean, with life, the reason it's difficult is just because we have difficulty specifying precisely what we *count* as life and what we don't. There's no hidden facts here. It's just "do you call that life"? Which features are necessary for life?

It's a kind of a termonological [question]. There's no hidden facts.

But no matter how clearly we define this thing, we just can't tell where it is and where it isn't. And if someone said, "Look, my cup has it," there is no test you can do to prove that their cup doesn't have it.

Julia Galef: Yeah, although as a subjective Bayesian, you could take humans -- who you're very confident have it -- and look at, I don't know, apes and dogs, and other creatures that you're somewhat less confident have it... for reasons involving both observing their behavior, and also reasoning via evolution, about how similar our brains should be... Then you could look at which features of lesser and lesser species are shared with the higher species. And form credences about how likely it is that the lower species each have this property of subjective experience.

I'm just saying -- even though you can't know for sure, I'm saying you can guess, with better than random chance, I would say. Right?

Keith Frankish: I'm not sure. I think that's a very natural way of approaching it. But it assumes that consciousness has some kind of function, that it is

doing something. And that we can detect it, we can correlate it with certain behaviors, and certain reactions, and so on.

Now with this kind of “bad” picture of consciousness you can't do that.

Julia Galef: It might be easier to let go of this model if we have an alternative.

Keith Frankish: Absolutely. Well, let me just walk really quickly through what the three, or four positions you can take on this are.

One, is just sort of accept this conception of this very anomalous thing that we have, this very anomalous aspect of our nature, and say, "It's real. It's there. It can't be accounted for by science as we know it. We're going to have to expand science, or go into metaphysics."

We can maybe go with panpsychism. We maybe say that there are these radically emergent properties that are not explicable in terms of lower level properties, just the way you get certain kinds of structures. This extra property just pops into existence inexplicably.

Or you could start appealing maybe to some strange effects like quantum effects, that... somehow, they produce this stuff. But it means going well beyond existing cognitive science, and into completely new realms. Some people think they have to do that.

Another response is a more conservative response. It says, "No, we can kind of explain all these using the standard resources of cognitive science, by talking about representations in the brain, and maybe some sort of self-awareness." This is where the self-awareness comes back in. Maybe when we start to represent our own experiences to ourselves, that's when this apparent subjectivity comes in.

So you have this whole bunch, and their standard take on this is that, "Yes, experience does have this strange subjective aspect, but it's just because of the way that we look at it introspectively." When we focus on the qualia of an experience -- the blueness of the experience itself, as opposed to the blueness of the cup, as it were -- what we're actually focusing on is just some kind of pattern of neural activity. That is what we're aware of, but we're just looking at it from an unusual angle, as it were.

So this is an identity theory. The idea is that pain is just a firing of certain neurons. They're the same thing. The idea is we have some sort of introspective mechanisms that allow us to sort of scan our

brain activity, and when they scan brain activity, this is what it's like. This is what it is.

Julia Galef: I think that's the view that, if you'd asked me to explain consciousness, I would have fumbled way towards something like that.

Keith Frankish: That's been the standard view.

Now, with illusionism... the fundamental ontological commitments are the same as the physical view, but it just goes a bit further. It says, "Well, yes, there is some sort of introspective mechanism involved here, but what they're doing is misrepresenting their targets."

It's not that the targets, these brain states really *are* like that, really do have these simple, essentially private qualitative properties. Brain states couldn't have those. They're just way too different. It misrepresents brain activity as being like that. That's why it's an illusion.

Julia Galef: Okay, yeah.

Keith Frankish: It takes seriously the idea that, "Yes, there is some sort of self-monitoring going on here" -- so, a representation of our own experiences, but it's a misrepresentation. And a misrepresentation that we can't unpick, precisely because we don't have any further mechanisms scanning that. We don't have any way of checking that, or challenging that. That's kind of that, as far as our monitoring goes. We're just left with that -- it's like that, that's how it is for you, but that in itself is just a representation.

My problem with the physicalists, with that sort of picture --

Julia Galef: Physicalists, they're called physicalists?

Keith Frankish: Well, it's a physicalist approach.

So, say people have representational theories of consciousness. They identify phenomenal consciousness with the existence of certain kinds of representational states. Or certain kinds of second-order monitoring of representational, or perceptual states. Now, they do it all in terms of mental representations, that kind of stuff, the bread and butter of cognitive science.

My problem with them is not that they're not giving good explanations of what's actually happening, but that they're passing

them off as explanations of phenomenal consciousness, as explanations of consciousness as conceived in this folk way. Now I think mechanisms like that may very well be operative. But they cannot explain consciousness in this strong, phenomenal sense. Because that is impervious to explanation.

Here is an analogy. In the Middle Ages, people thought that sometimes people were possessed by demons. And I guess some people sincerely thought they themselves were possessed by demons. This maybe was very convincing to them, the first person perspective.

Now let's say modern clinical psychiatry comes along and gives quite a different explanation of what's happening. And now, do you say "That's what demons really are"? That they really are schizophrenia, or whatever characterization modern psychiatry gives of it?

Or do you say, "No, stop thinking about it in terms of demons altogether. This isn't an explanation of what demons really are. Thinking in terms of demons is just the wrong way. Don't even start there. Start with this other conception"?

And that's kind of what I'm asking to do. Because so long as you try to pass off some sort of cognitive explanation as an explanation of phenomenal consciousness, the realists won't be happy. They will say:

"But you're dodging the hard question, the hard problem. You're not really explaining this thing that we have in mind when we talk about consciousness. This thing that is essentially private, that is immediately presented to us, that zombies could lack. You're not explaining that. Because I can imagine a zombie having all those mechanisms of introspection you've just described but not being phenomenally conscious."

So you're never going to satisfy them.

Julia Galef:

Does it seem at all worrying to you that... I sort of feel like if you're going to argue that something is an illusion, you need to show people how the illusion is done. Like a magician's illusion. And the way you know that you succeeded is if people have this, "Ah, *that's* how it's done feeling," and they no longer feel like there is a mystery, or magic, to be explained anymore.

Do you think it's a bad sign about your theory that it is not producing that "Ah, okay, now I'm no longer confused," reaction in people?

Or is the fault in them, for failing to grasp the explanation?

Keith Frankish: Look, if it is an illusion -- and of course nothing too much hinges on that, the word "illusion," and the analogies, and the connotations that "illusion" has. It's a mistaken belief, a systematically generated mistaken belief I suppose.

If it is an illusion, it's a very sort of deep-rooted one. Now, how far it's the product of our introspective hardware, and how far it's a product of reflection on the deliverances of that hardware -- we kind of introspect into our minds, and then we sort of theorize it in a certain way, we apply a certain folk theory to that -- I don't really know.

Some people think that it's kind of hardwired into our brains. Nicholas Humphrey, the psychologist, has quite a lovely theory about how this illusion has been an adaptive feature. That this feature was selected for because it made the world, it made our lives, and our experiences, seem richer and more important to us. We started to have an extra level of interest in our own existence. It made us seem special.

Julia Galef: Was that adaptive somehow?

Keith Frankish: He thinks so. He thinks that we began to relish experience for its own sake.

Julia Galef: So it increased our will to survive?

Keith Frankish: He thinks so. He also thinks it played an important role in socialization, I think. That we began to see each other as centers of specialness. So we weren't just like the natural world, and we weren't even just like other animals.

We were these creatures with souls, it seems... We began to conceive of ourselves as creatures that are special, that have these inner worlds, which could perhaps survive death. And others around us, as having these special inner worlds too, and so mattering, in a way that other creatures didn't. Their lives mattering, in a way that other creatures didn't.

He sees this as kind of the root of all that makes human life special.

Julia Galef: You mentioned earlier that panpsychism was growing in popularity among philosophers. Are there any other trends you can point to in the last 20, 30 years of discussing consciousness, and the hard problem of consciousness, that... like, some theories have been rejected, or some theories people have converged on, at all?

And to what extent have any of those trends been the result of empirical evidence, as opposed to just people making better arguments?

Keith Frankish: Right, I think the modern debate -- when I say modern, I mean the last 20 or 30 years -- was really shaken up by David Chalmers' 1996 book, *The Conscious Mind*. I mean, he was building on a lot of work that had been done in the past, but he did a very forceful presentation of an anti-physicalist view of consciousness, drawing up things like the zombie argument. He really kind of helped to revive interest in the problem, and to revive interest in the viability of anti-physicalist approaches to it.

Now the standard physicalist response to that was to say, "Well, yes, there is something here we need to explain, but we can explain it in terms of a sort of conceptual dualism, rather than any metaphysical dualism."

There's two different ways of conceptualizing these brain states. One is the perspective of the neuroscientist, and the other is the introspective perspective of the person who's undergoing the states. And they conceptualize these states very differently, but they are in fact the same state.

This was often known as the conceptual phenomenal concept strategy. That we have special introspective concepts for thinking about the experience. ... That was the view you were attracted to originally. So that tried to reduce this metaphysical division to a mere conceptual division, to different ways of thinking about the same thing.

And I think what's happened, probably in the last 10 years or so, is that faith in that strategy has kind of waned. Increasingly, people are seeing that that strategy won't really do it.

Julia Galef: So that's why panpsychism rushed in to fill the void.

Keith Frankish: And I see illusionism as the other response to the breakdown of that kind of compromised strategy.

Julia Galef: Got it.

Keith Frankish: That strategy, the phenomenal consciousness strategy, kind of took seriously all these intuitions that the realists have, but say that they're compatible with physicalism. Now that's what the phenomenal concept strategy was supposed to do. But increasingly, people see that if you take those intuitions seriously, you've then got to do something like, I don't know, panpsychism or emergentism, or something. You've got to actually pay the metaphysical price -- or reject those intuitions and go illusionist.

Julia Galef: So all of this evolution of your collective understanding of consciousness is down to argument, not to empirical evidence, right? Or I guess, I don't know, stuff like Mr., was it Mr. Crabgrass who-

Keith Frankish: Clapgrass.

Julia Galef: Clapgrass, yeah, or empirical evidence about split-brain patients or something, I can imagine that kind of thing influencing the consciousness debates. Was there anything like that, that ended up influencing it?

Keith Frankish: I think all that stuff is slowly pushing towards the sort of picture -- I probably would say this, wouldn't I -- the sort of picture that I have. Because it's showing how that simple introspective picture breaks down, and how it can gradually dismantle this user interface. And when the interface doesn't function properly we begin to get glimpses of how it's actually being constructed.

I think all that stuff does some work.

But the reason a lot of this debate is conducted at a distance from the empirical literature is simple: Because for a lot of people it's a metaphysical debate. It's a debate about the place of these things in the physical world.

And of course if they're not actually aspects of the physical world, if they're essentially subjective, then science is going to tell you zero about them.

Julia Galef: Although, maybe they would feel less drawn to the metaphysical explanations if they had a more compelling physical explanation. I could imagine empirical evidence indirectly helping resolve these debates that *seemed* metaphysical...

Keith Frankish: I think that's what should happen, and I think it probably will happen. But the thing is, some people start with a very, very strong intuition that qualia are presented to us in a way that is immediate,

and transparent. Transparent in the sense that they are revealed to us, there's nothing hidden about them.

Julia Galef: That we're not making an inference of any kind.

Keith Frankish: You don't need to. Just by having the experience, and just attending to the experience, we can know everything about the character of that property that we're experiencing.

Now, it's pretty obvious, I think, if that's your conception of what needs explaining, then science isn't going to give you much help on that... How could we know any feature of the physical world in that immediate, transparent, revelatory way? Certainly I can't think of any other model for that in any other area of natural science. We're out of that realm.

And therefore, to suggest that science should shed light on this is, for some people, to kind of miss the point of the whole debate.

Julia Galef: Can you conceive of any scientific evidence -- that maybe we can't obtain with our current scientific technology, but that we could in theory obtain -- that you think would settle the debate?

Keith Frankish: I could think of some that I think *should* go a long way, but I don't know whether it would work.

It was something like this: Suppose we could start to map the introspective mechanisms responsible for creating this user illusion. So we can explain why we have this introspective picture of ourselves, why it seems to us that we are confronted with this immediate access to this private world of qualia. We can explain all our judgments about it, all our intuitions about it, why it seems so mysterious and inexplicable. We can explain all of that in cognitive terms.

So we can say, "Look, I know you think this stuff is real," but even if it weren't real you would still believe all of that.

Julia Galef: This feels very parallel to attacking moral realism by explaining evolution, and how evolution would have produced these moral intuitions in us, even if there were no such thing as moral truths.

Keith Frankish: Exactly, it's a debunking argument.

So now, David Chalmers has recently written a fascinating thing called *The Meta-problem of Consciousness*. The meta-problem is precisely that, of explaining all our intuitions about it. He thinks it's

quite possible that we can do that in standard cognitive science terms, in physical terms. We can explain all our intuitions about it in this way, thereby providing the premises for a debunking argument against realism.

He's quite open about this, and he's quite clear that that will provide a good argument for the illusionist.

Julia Galef: And he is not an illusionist. So that's kind of a nice thing for him to do, to lay out what would prove him wrong.

Keith Frankish: And he's very clear that if he were a physicalist, he would be an illusionist. He thinks it's the coherent way for a physicalist to go, and I'm absolutely on board with him about that. I think he's right.

But he says despite the existence of this debunking argument, we would still know that consciousness was real. Because we're directly acquainted with it.

Now in response to that, I wrote a little paper, a comment on that. I try to run a parallel argument, which I think is analogous:

[Suppose] we take a bunch of UFO reports, and we manage to show that all of the observer's beliefs are explicable in mundane terms. There was some atmospheric disturbance; there was some aircraft; they were hallucinating; whatever it was. And we show this to everyone's satisfaction.

But the observer says, "Yes, okay, I accept that. I accept that there was an atmospheric disturbance there. But there was still definitely a UFO there as well. I know that my belief wasn't caused by the UFO, I know it was caused by this disturbance in the air, but there was a UFO behind it all the same."

Or, "I know that I was hallucinating, and I didn't actually see the UFO, but there definitely was one there all the same."

And we say, "How do you know that?" And he says, "Well, I've got a kind of sixth sense for UFOs. I've got this extra sense for UFOs... when they're present, they just immediately reveal themselves to me, and I know they're there with absolute conviction."

Julia Galef: It's not a very flattering analogy.

Keith Frankish: I am afraid it isn't. But it does seem that that's kind of the dialectic. That the argument is "Even if my beliefs about phenomenal

consciousness are not produced by phenomenal consciousness, they're still true.”

Actually David articulated this position in his book, back in 1996, *The Conscious Mind*. He called it the Paradox of Phenomenal Judgment -- the zombies who are just like us, but don't have phenomenal consciousness, would believe they were phenomenally conscious. And his zombie twin would be as convinced of the reality, and non-physicality, of consciousness as he is...

I didn't really answer the question you asked earlier, about how, even if you have the illusion explained to you, it doesn't disappear.

Julia Galef: Yeah, and whether that fact is a sign that you haven't successfully explained things, that your theory is missing something still.

Keith Frankish: What I talked about there was about Nick's theory, about it being sort of a hardwired feature of us.

Now if it is, that would explain why we can't shake it. I mean, if we're hardwired to conceptualize ourselves in that way, then sure, we're not going to shake it. It's going to be like the Muller-Lyer Illusion – however much you don't trust it on reflection, it's still going to *seem* to be there. And that might be right.

Another approach, which I think Dennett is more sympathetic to, is the idea that what we're doing here is projecting a sort of philosophical theory or folk philosophical theory onto introspection. And that we can reconceptualize it -- that is, we're not hardwired to think that way. People who are into Buddhist philosophy tell me that this is what Buddhist thinkers have been doing for a long time.

So I think it's an open question, to what extent we could shake off this fake picture.

Julia Galef: Okay, well -- Keith, I think that's probably as unconfused as I am going to get in one session. So before we wrap up I wanted to ask you if there is a book or a thinker who you substantially disagree with, but who you nevertheless have gotten a lot of value out of reading. Anyone like that come to mind?

Keith Frankish: Well, in terms of consciousness I would say David Chalmers, his 1996 book *The Conscious Mind*. It's a magnificent book. He's a brilliant philosopher. He argues with exceptional rigor and clarity, and he's presented a case for a non-physicalist view of consciousness in the most powerful, and persuasive way.

He structured my thinking about the whole debate. And in some ways what I'm doing is kind of having an argument with him. But it's an argument that is in many ways shaped by his way of forming the debate.

I mean, I think he's got a lot of the structure of the debate absolutely right. I think he made some wrong turnings at certain points, but I think that the way that he sets up the debate is right.

I think he's a brilliant writer. He certainly had a big influence on me, in terms of the quality of his writing as a philosopher.

I would encourage everyone to read the strongest presentations that you can find of your opponent's views. They present a challenge, they stimulate you, they make you really want to sharpen up your own ... Don't read people who agree with you, that's far too easy, and that won't sharpen you up. You want to get a work out find somebody who can really challenge, and really push you, really make you go the extra mile, and Chalmers does it.

Julia Galef: My knee-jerk reaction was to give an Amen, but I want you to keep following me, so I'm now vainly searching for some reason why people *shouldn't* read views that disagree with them...

Anyway, Keith, thank you so much for coming on the show. I am still confused, but I have at least a much clearer, and richer picture of the sort of landscape of what people are arguing over, and what the sticking points are, which is really helpful. So thank you so much.

Keith Frankish: Well, thank you, I think it was ... Again, as Chalmers has said there's no view of consciousness that *isn't* crazy, so thank you, Julia, I really enjoyed it.

Julia Galef: You just have to pick which flavor of crazy you want.

Julia Galef: We'll link to your website, and your blog, and your Twitter handle, which I really enjoy following. And to Illusionism, and if you want to give me a couple links to Chalmers' best papers or books for people to read, that present the other side in the strongest possible way, let me know.

Keith Frankish: Okay, will do. Thanks so much.

Julia Galef: This concludes another episode of Rationally Speaking. Join us next time for more explorations on the borderlands between reason and nonsense.