

00:04

Who am I? Who is anyone, really?

00:09

When I wake up in the morning and open my eyes, a world appears. These days, since I've hardly been anywhere, it's a very familiar world: there's the wardrobe beyond the end of the bed, the shuttered windows and the shrieking of seagulls, which drives Brighton residents like me absolutely crazy. But even more familiar is the experience of being a self, of being me, that glides into existence at almost the same time.

00:35

Now this experience of selfhood is so mundane that its appearance, usually, just happens without us noticing at all. We take our selves for granted, but we shouldn't. How things seem is not how they are. For most of us, most of the time, it seems as though the self, your self, is an enduring and unified entity -- in essence, a unique identity. Perhaps it seems as though the self is the recipient of wave upon wave of perceptions, as if the world just pours itself into the mind through the transparent windows of the senses. Perhaps it seems as though the self is the decision-maker in chief, deciding what to do next and then doing it, or, as the case may be, doing something else. We sense, we think and we act. This is how things seem.

01:31

How things are is very different, and the story of how and why this is so is what I want to give you a flavor of today. In this story, the self is not the thing that does the perceiving. The self is a perception too, or rather, it's the collection of related perceptions. Experiences of the self and of the world turn out to be kinds of controlled hallucinations, brain-based best guesses that remain tied to the world and the body in ways determined not by their accuracy, but by their utility, by their usefulness for the organism in the business of staying alive.

02:13

Now the basic idea is quite simple, and it goes back a very long way in both science and philosophy -- all the way back, in fact, to Plato and to the shadows cast by firelight on the walls of a cave, shadows which the prisoners within took to be the real world.

02:30

Raw sensory signals, the electromagnetic waves that impinge upon our retinas, the pressure waves that assault our eardrums, and so on, well, they're always ambiguous and uncertain. Although they reflect really existing things in the world, they do so only indirectly. The eyes are not transparent windows from a self out onto a world, nor are the ears, nor are any of our senses. The perceptual world that arises for us in each conscious moment -- a world full of objects and people, with properties like shape, color and position -- is always and everywhere created by the brain, through a process of what we can call "inference," of under-the-hood, neurally implemented brain-based best guessing.

03:17

Now ... Here's a red coffee cup. When I see this red coffee cup, when I consciously see it, that's because "red coffee cup" is my brain's best guess of the hidden and ultimately unknowable sensory signals that reach my eyes. And just think about the redness itself, for a moment. Does the color red exist in the world? No, it doesn't. And we don't need neuroscience to tell us this. Newton discovered long ago that all the colors we experience, the rainbow of the visible spectrum, are based on just a few wavelengths of electromagnetic radiation, which itself is, of course, entirely colorless. For us humans, a whole universe of color is generated from just three of these wavelengths, corresponding to the three types of cells in our retinas. Color-wise, this thin slice of reality, this is where we live. Our experience of color -- indeed, our experience of anything -- is both less than and more than whatever the real world really is.

04:25

Now what's happening when we experience color is that the brain is tracking an invariance, a regularity in how objects and surfaces reflect light. It's making a best guess, a top-down, inside-out prediction, about the causes of the relevant sensory signals, and the content of that prediction -- that's what we experience as red. Does this mean that red is in the brain, rather than the world? Well, no. The experience of redness requires both the world and a brain, unless you're dreaming, but let's not worry about that for now. Nothing in the brain is actually red. Cézanne, the great impressionist painter, once said that color is where the brain and the universe meet.

05:11

Now the upshot of all this is that perceptual experience is what I've come to call, drawing on the words of others, a "controlled hallucination." Now this is a tricky term, prone to misunderstandings, so let me be clear. What I mean is that the brain is continuously generating predictions about the causes of sensory signals, whether these come from the world or from the body, and the sensory signals themselves serve as prediction errors, reporting a difference between what the brain expects and what it gets, so that the predictions can be continuously

updated. Perception isn't a process of reading out sensory signals in a bottom-up or outside-in direction. It's always an active construction, an inside-out, top-down neuronal fantasy that is yoked to reality in a never-ending dance of prediction and prediction error. Now I call this process controlled hallucination to emphasize just this point. All of our experiences are active constructions arising from within, and there's a continuity here, between normal perception and what we typically call hallucination, where, for example, people might see or hear things that others don't. But in normal perception, the control is just as important as the hallucination. Our perceptual experiences are not arbitrary. The mind doesn't make up reality. While experienced colors need a mind to exist, physical things, like the coffee cup itself, exist in the world whether we're perceiving them or not -- it's the way in which these things appear in our conscious experience that is always a construction, always a creative act of brain-based best guessing.

07:00

And because we all have different brains, we will each inhabit our own distinctive, personalized inner universe.

07:09

Now I've digressed quite far from where we began, so let me end by returning to the self, to the experience of being you, or being me. The key idea here is that the experience of being a self, being any self, is also a controlled hallucination, but of a very special kind. Instead of being about the external world, experiences of selfhood are fundamentally about regulating and controlling the body. And what's important here is that the experiences of being a self are composed of many different parts that normally hang together in a unified way, but which can come apart in, for instance, psychological or neurological disorders. There are experiences of being a continuous person over time, with a name and a set of memories shaped by our social and cultural environments. There are experiences of free will, of intending to do something, or of being the cause of things that happen. There are experiences of perceiving the world from a particular perspective, a first-person point of view. And then, there are deeply embodied experiences, for instance of identifying with an object in the world that is my body. These hands, they're my hands. And then, of emotion and mood. And at the deepest-lying, most basal levels, experiences of simply being a living body, of being alive. Now my contention is that all these aspects of being a self are all perceptual predictions of various kinds. And the most basic aspect of being any self is that part of perception which serves to regulate the interior of the body to keep you alive.

08:57

And when you pull on this thread, many things follow. Everything that arises in consciousness is a perceptual prediction, and all of our conscious experiences, whether of the self or of the world,

are all deeply rooted in our nature, as living machines. We experience the world around us and ourselves within it, with, through and because of our living bodies.

09:25

So who are you, really? Think of yourself as being like the color red. You exist, but you might not be what you think you are.

09:38

Thank you.

09:40

David Biello: A stand-in for the audience. Anil Seth: David is clapping.

09:43

(Laughter)

09:45

AS: That makes me feel better. DB: It was great. Thank you for that.

09:49

I have to say that the thought of my brain floating around in a bony prison is a disturbing one. But how do all those billions and trillions of neurons give rise to this experience of consciousness, in your view?

10:07

AS: First, I mean, consciousness is experience, so I'd use the two terms synonymously there. It's the same thing. And by the way, the idea of your brain wobbling around in its bony vault of a skull is presumably less disturbing than it doing something else and doing something outside of the skull. (Laughter) That would be the more worrying situation. But the question, of course, this is the big question. You start off with a simple question, "How does it all happen?" And this is why there is a long way to go here.

10:40

And there are, I think, two ways to approach this mystery. So the fundamental question here is ... What is it about a physical mechanism, in this case, a neurobiological mechanism, 86 billion neurons and trillions of connections, that can generate any conscious experience? Put that way, it seems extremely hard, because conscious experiences seem to be the kinds of things that cannot be explained in terms of mechanisms, however complicated those mechanisms might be. This is the intuition that David Chalmers famously called "the hard problem." But my approach, as hinted at in this talk, is that we can characterize different properties of consciousness -- what a perceptual experience is like, what an experience of self is like, what the difference between sleep and wakefulness is like. And in each of those cases, we can tell a story about how neural mechanisms explain those properties.

11:36

In the part of the story we've touched on today, it's all about predictive processing, so the idea is that the brain really does encode within it a sort of predictive generative model of the causes of signals from the world, and it's the content of those predictions that constitutes our perceptual experience. And as we sort of develop and test explanations like this, the intuition is that this hard problem of how and why neurons, or whatever it is, in the brain, can generate a conscious experience, won't be solved directly -- it will be dissolved. It will gradually fade away and eventually vanish in a puff of metaphysical smoke.

12:16

DB: Katarina wants to talk about anesthesia, that experience of having your consciousness kind of turned off. What do we know about this ability to switch a person off, in a matter of seconds? What is actually happening there, do you think?

12:33

AS: Firstly, I think it's one of the best inventions of humanity, ever. The ability to turn people into objects and then back again into people -- I wouldn't want to live at a time in history without it. Whenever we have this, like, "Wouldn't it be nice to live in Greek antiquity or something, when people swum around, philosophizing, drinking wine?" Yes, but what about anesthesia? (Laughs) That's my response. It does work, this is a fantastic thing. How? Here's an enormous opportunity for consciousness science, because we know what anesthetics do at a very local level. We know how they act on different molecules and receptors in the brain. And of course, we know what ultimately happens, which is that people get knocked out. And by the way, it's not like going to sleep. Under general anesthesia, you're really not there. It's an oblivion comparable with the oblivion before birth or after death. So the real question is, "What is happening?" How is the local action of anesthetics affect global brain dynamics so as to explain this disappearance

of consciousness? And to cut a long story very short, what seems to be happening is that the different parts of the brain become functionally disconnected from each other, and by that I mean, they speak to each other less. The brain is still active, but communication between brain areas becomes disrupted in specific ways. and there's still a lot we need to learn about the precise ways in which this disconnection happens -- what are the signatures of the loss of consciousness? There are many different kinds of anesthetic, but whichever variety of anesthetic you take, when it works, this is what you see.

14:21

DB: I think some folks such as Jasmine and more anonymous folks are troubled by this idea that what I call red might be a different color for you and for everyone else. Is there a way of knowing if we're all hallucinating reality in a similar way or not?

14:42

AS: Again, this is a lovely topic, and it really gets to the heart of how I've been thinking about perception, because one of the aspects of perception that I think is easy to overlook is that the contents of perception seem real, right? The redness of this coffee cup, it seems to be a mind-independent, really existing property of the external world. Now, certain aspects of this coffee cup are mind-independent. Its solidity is mind-independent. If I throw it at you, David, across the Atlantic, and you don't see it coming, it will hit you in the head, it will hurt. That doesn't depend on you seeing it, but the redness does depend on a mind. And to the extent that things depend on a mind, they're going to be different for each of us. Now, they may not be that different. In philosophy, there's this argument of the inverted spectrum, so if I see red, is that the same as you seeing green or blue, let's say? And we might never know. I don't have that much truck with that particular thought experiment. Like many thought experiments, it pushes things a little bit too far. I think the reality is that we see things like colors, maybe we see them similar, but not exactly the same, and we probably overestimate the degree of similarity between our perceptual worlds, because they're all filtered through language. I mean, I just used the word "red," and there are many shades of red; painters would say, "What red?" I remember when I was decorating my house, it's like, "I want to paint the walls white." How many shades of white are there? This is too many. And they have weird names, which doesn't help. We will overestimate the similarity of our universe. And I think it's a really interesting question, how much they do indeed diverge.

16:33

You will probably remember this famous dress, this photo of a dress half the world saw as blue and black, and the other half saw as white and gold.

16:42

AS: You're a white and gold person? DB: Yeah, yeah.

16:47

AS: I'm a blue and black person. I was right, the real dress is actually blue and black. (Laughter)

16:52

AS: Never mind...

16:55

DB: We could argue about that.

16:57

AS: We couldn't. It really is blue and black. I talked to the dress designer. The actual one is blue and black. There's no argument there. But the thing that made that so weird is that it's not that we vaguely see it as one color or the other, we really see that blueness and blackness or whiteness and goldness as really existing in the world. And that was an interesting lever into a recognition of how different our perceptual universes might be. And in fact, a study we're doing at Sussex over the next year or two, we're trying to characterize the amount of perceptual diversity that is just there to be discovered. We're usually only aware of it at the extremes, people call things like neurodiversity, where people have experiences that are so different, they manifest in different behaviors. But I think there's this, sort of, big dark matter of individual diversity in perception that we know very little about, but it's there.

17:56

DB: I'm glad we could put to rest a major internet debate and come down firmly on the blue and black side of things. Daniella wants to know, "Could you explain how memory is involved in this perception of a self?"

18:13

AS: Just as there are many different aspects of selfhood, there are many different kinds of memory, too. I think colloquially, in everyday language, when we talk about memory, we often talk about autobiographical memory or episodic memory, like "What did I have for breakfast?"

"When did I last go for a walk?" These kinds of things. "When did I last have the pleasure of talking to David?" These are the memories of things that pertain to me as a continuous individual over time. That's one way in which memory plays into self, and that part of memory can go away, and self remains -- back to the earlier point. There's a famous case I talk about in the book, of a guy called Clive Wearing, who had a brain disease, an encephalopathy, which basically obliterated his ability to lay down new autobiographical memories. He lost his hippocampus, which is a brain region very important for this function. His wife described it as him living in a permanent present tense, of between seven to 30 seconds. And then, everything was new. It's very, very difficult to put yourself in the shoes of somebody like that. But other aspects of his self remained.

19:30

But then, there are all sorts of other aspects of memory that probably also play into what it is to be you or to be me. We have semantic memory. We just know things, like we know what the capital of France is, who the president is, I hope so, I don't know. Sometimes, that's a good thing. Sometimes, that's not a good thing. And all of these things that get encoded in memory shape our self too. And then finally, there's perceptual memory. It's not that experience is like a video recording that we can replay, but everything we experience changes the way we perceive things in the future, and the way we perceive things is also, in my view, part of what it is to be a self.

20:16

Actually, I just want to say, one of the really interesting questions here, and one of the things we're working on -- Imagine a typical day. You go through your typical day, you're experiencing a continuous stream of inputs. Now you blink, of course, and so on, but more or less, there's this continuous stream of inputs. Yet when we remember a day, it's usually in chunks, these autobiographical chunks: "I did this, I did that, I did the other, this happened." So a really important question is, "How does this chunking process happen?" "How does the brain extract meaningful episodes from a relatively continuous flow of data?" And it's kind of disturbing, how little of any given day we remember. So it's a very selective process, and that's something that I think is going to be useful not only for basic neuroscience, but, for instance, in helping people with memory loss and impairments, because you could, for instance, have a camera, and then, you could predict what aspects of their day would constitute a memory, and that can be very, very useful for them and for their carers.

21:24

DB: The brain clearly has a good editor. You call us, people, "feeling machines" in your book. Care to expand on that?



21:35

AS: Yeah, that's right. Well, we're not cognitive computers, we are feeling machines. And I think this is true at the level of making decisions, but for me, it's really at the heart of how to understand life, mind and consciousness. And this, really, is the idea that -- In consciousness science, we tended to think things like vision -- Vision as being the royal road to understanding consciousness. Vision is easy to study, and we're very visual creatures. But fundamentally, brains evolved and develop and operate from moment to moment to keep the body alive, always in light of this deep physiological imperative to help the organism persist in remaining an organism, in remaining alive. And that fundamental role of brains, that's what, in my view, gave rise to any kind of perception. In order to regulate something, you need to be able to predict what happens to it. It's this whole apparatus of prediction and prediction error that undergirds all of our perceptual experiences, including the self, has its origin in this role that's tightly coupled to the physiology of the body. And that's why, I think, we're feeling machines, we're not just computers that happen to be implemented on meat machines.

22:59

DB: Thank you, Anil, for chatting with us today.

23:02

AS: Really enjoyed it.

23:03

AS: Thanks a lot, David. DB: Thank you.