Big Biology Episode Transcript

Ep 109: Nothing in biology makes sense except through time (with Kevin Mitchell)

SPEAKERS

Kevin Mitchell, Art Woods, Marty Martin

Art Woods

Hey Marty, here's a question that's both trite and deeply philosophical – did you choose to make this podcast?

Marty Martin

Uh, yeah, of course. You know our origin story: a few too many beers at a bar in Tampa one evening in 2017 and, bang, launching a podcast started to sound like a good idea.

Art Woods

Sure, but did you really have a choice in making that choice? Or was it predestined, an inevitable outcome of clock-like mechanisms stretching back into your earlier life, and perhaps even before.

Marty Martin

Oh, you're asking whether I have free will.

Art Woods

Yep, age-old question, pondered for at least several thousand years.

Marty Martin

One that's now thrashed out primarily in college dorm rooms or dense tomes evacuated by philosophers.

Art Woods

That said, the topic of free will may be on your mind, too, but because of a book that's got a lot of press since it appeared a few weeks ago.

Marty Martin

That book, by Robert Sapolsky, a long-time observer of primates and their hormones and the workings of the mind, is called Determined: A Science of Life without Free Will.

Art Woods

As the title suggests, the book lays out a detailed case against free will. Sapolsky argues essentially that all of the choices we make –

Marty Martin

from the mundane to the grand –

Art Woods

reflect physiological states and neural processes that are simply outside of our control.

Marty Martin

And those neural processes are shaped by our histories –

Aart Woods

how we interacted with our parents, first dates, first breakups, our current physiological states, our genes–

Marty Martin

Sapolsky would say that all of these things determine when we are stressed or hungry or bedazzled by a new romantic partner, so in a very real sense we don't really have any control and all of the choices we make are pre-ordained.

Art Woods

Well, today we're talking today to someone whose ideas sit on the other side of the fence…Kevin Mitchell. Kevin is an Associate Professor of Developmental Neurobiology and Genetics at Trinity College Dublin, and he also just published his own book called Free Agents: How Evolution Gave Us Free Will. It reaches entirely different conclusions.

Marty Martin

Kevin's main thesis is – absolutely, humans have free will because all organisms have agency. And free will is just the name we give to our version of what all living system must have to stay alive.

Art Woods

Today, we talk with Kevin about several of the key arguments for and against free will, and their origins in physics and philosophy, including some of the very same points raised by Sapolsky in his book.

Marty Martin

But we spend the bulk of our time talking about the different ways that organisms escape from physical and chemical laws.

Art Woods

These include the uncertainties inherent in quantum effects, which can bubble up from below to affect physiological processes, including how neurons and nervous systems function.

Marty Martin

And perhaps most importantly that causality in living systems should not be understood as simple causal loops, but a stretched out spiral through time. This means that causation in some systems can flow from the top-down.

Art Woods

By top-down here, Kevin means that organisms themselves are prime causes of what happens at lower levels of organization, not the other way around.

Marty Martin

Part of what gives organisms top-down control is that they extract meaning from information coming in from their local environments.

Art Woods

They then use this meaning to limit the actions they take in the future, not because elementary particles tell them to but because organismal parts get rearranged into new forms that influence their choices.

Marty Martin

Kevin makes a strong case that organisms have evolved agency precisely because it was so important for exploring the world and making the right choices.

Art Woods

He says: “The functional organization of a system can have causal power over how things unfold.” So smile, maybe you do have free will after all!

Marty Martin

I’m Marty Martin

Art Woods

I’m Art Woods

Marty Martin

Exercise your free will to listen to the rest of this episode of Big Biology!

Intro Music

Art Woods

So Kevin, it's super great to have you on the show today. Thanks for joining us. We're here to talk primarily about your new book, which should drop just before the episode airs. And it's called free agents, how evolution gave us free will, Marty and I were just super engrossed by the book. And it fits in a super interesting way with some of the topics that he and I have been discussing a lot. And that listeners will realize we've talked about some on air, we're going to go a different direction today and delve into some other interesting areas that underlie this idea of agency and free will. But maybe let's just start briefly by you telling us what you mean by this idea of agency? And what sorts of organisms have agency?

Kevin Mitchell

Yeah, so it's a super complicated question. And there's not a sort of a bright line, I think. So agency, you know, for me, is just the ability of organisms to do things fundamentally, that there's sort of property that it's funny, if you look in, you know, so Introductory Biology textbooks, you open the textbook, and it gives you this list of characteristics that living things have. And there'll be things like, you know, replication and metabolism and some sort of sentience and homeostasis, and so on.

But agency doesn't make the list, it's not really something that people pay much attention to, it hasn't been a big topic really, in I think, a lot of mainstream biology. And yet, for me, it's the defining characteristic of living things, they can act, they do work to keep themselves organized, and they can act in the world. And, you know, at some level, you can kind of ask, you can divide things like just molecular homeostasis from actual behavior. And I prefer to use agency when I'm referring to behavior. But, you know, again, it's a bit semantic, I think you can make arguments that the molecular homeostasis kind of stuff is a kind of agency as well. It's a sort of a holistic organismal level reconfiguration of internal pathways to cope with changing conditions. And you know, behavior is just another way to do that, it's another way to cope with changing conditions by moving around in the world, approaching things that are good, avoiding things that are bad, and generally making happen what you want to happen.

Marty Martin

So it's striking that, going back to what you said just a second ago, you open up this introductory textbook to see the defining characteristics of life and agency isn't one of them. Why do you think that is? I mean, it seems pretty stark, that living things are different than non-living.

Kevin Mitchell

I mean for me, the fact that living things can do things, right. They're not just places where things happen. That's the key element, right? So other things in the universe, you know, planets, electrons, atoms, they don't do anything. Things happen to them, or near them, or in them or involving them. But living beings are causal entities unto themselves as a whole thing. And I don't know if it's a sort of a philosophical concept that seems to violate physical law somehow, you know, it introduces a new kind of causation into the universe that didn't exist before life did. Or it's so taken for granted, that it's actually not even noticed. And so I don't know which of those is true.

Marty Martin

Yeah. So you know, and it's not, it's not a secret that in evolutionary biology, there has been different, you know, the conflict historically, between creationism and scientific ways of understanding, do you think maybe it has something to do with religion, using that kind of an idea as a way to involve itself in biology?

Kevin Mitchell

It might do actually, I think you might be onto something there, because one of the aspects of agency is purpose– things doing something for a reason; goal directed behavior. And purpose is a kind of a slippery, vague term, and I think many biologists don't like it. And I think it's partly because of that history that you're talking about where it was invoked, as if the universe had a purpose. And the purpose of the universe was to create life.

And for me, that doesn't make any sense that the universe doesn't, you know, doesn't have any purpose, but it creates things that do have purpose. And I think we can naturalize that in a very simple way, which is just that living things that are organized a certain way, that act a certain way, persist. better than ones that don't. So, you know, that right there gives you some normativity that is, things are good or bad, relative to that process relative to that aim, you can call it an aim of a self. So that's partly what I was trying to do in the book actually is just kind of naturalize that concept of purpose and then once you have that, in a non-mystical, non-magical way, you can build concepts of meaning and value relative to that. Partly that history though that intellectual history might be what makes some biologists shy away from these things because they feel like they're going to the other side, as it were, getting too mystical.

Marty Martin

Yeah. And as soon as we start talking about complicated things, you know, a lot of people will tune off. So a simple message is a nice place to start, if you could do that, I mean, we're about to spend 60 minutes talking about agency and freewill, I guess it's complicated?! But before we sort of turn the page and all of those ideas into hopefully clear form, are there any non-living things that have agency or is this a property specific to life?

Kevin Mitchell

I don't think there are any non-living things yet that I think have agency. I mean, if you look at, for example, on artificial intelligence systems, right now, they can do a lot of interesting things like these large language models, for example, they kind of give an illusion, at least, of understanding things. They act like parasites from all the knowledge that they're fed, which is there in word form by humans. So their responses reflect, sort of human understanding, but they don't have understanding themselves. And I don't think they have, they definitely don't have agency in the sense that they don't do anything, right.

So I don't think we have yet any artificial systems that I would call agents, because there's nothing that sort of has that precarious nature, where what it does determines whether it continues to exist. And that makes it a locus of concern.

Art Woods

Yeah. And you can almost imagine AI is like evolving in some virtual or some real space in some way in which they had to start caring about their own survival in order to persist in those spaces, right. And that, could that lead to the emergence of agency in those systems?

Kevin Mitchell

Well, I think no, probably not by itself, I think you have to have other aspects of a functioning architecture that enable the kinds of cognition that promotes adaptive behavior, right in the service of that self survival. But, yeah, it would be really interesting to think about, you know, setting things up in silico, that are virtual agents. You know, they have some sort of master utility function, it doesn't have to be staying alive, actually, it could be anything really just to ground, to give them some normativity. And my, my feeling probably is that we might end up with something we could call a virtual agent, before we end up with something that's a physical agent, that's artificial.

Art Woods

I want to just here at the beginning to circle back to something you mentioned in passing just a few minutes ago. And that's about the different kinds of causation that might exist, that do exist, in the world and contrast something that you talked about a lot in the book, which is top-down causation versus sort of bottom-up causation, which is, you know, been been so dominant in in biology.

And I want to read a quote from chapter seven of your book: “Although they are made of physical components, “they,” meaning thw organisms, are not merely physical systems where the things that happen within them are driven by low level causes. They're organized for a purpose, and that organization constrains the physical components to enable true functionality and goal directed action.”

So it sounds like what you're saying is that there's sort of higher level things going on, that are a cause of the lower level events that underlie everything that's happening within a living body. Is that? Am I interpreting that correctly?

Kevin Mitchell

Absolutely. Yeah, absolutely. And I think I mean, it's funny when we think about causation. Since Newton, maybe, you know, we tended to think of this billiard ball kind of idea of just physical forces are the only real causes, and everything else just flows from them, whatever the organization of things, of matter is, at the moment, you just sort of could, in principle, work out all the physical forces that are aligned there and just submit that to some huge set of differential equations, and it would predict what would the next state be, right? And many people would think it predicts it exhaustively, that is, in a deterministic kind of a way, given all that information and the low level laws of physics, that's all you need to predict exactly what the next state is going to be and the next state after that, and after that, and after that. And of course, if you extrapolate that, you start to see that there's no branching timeline, right? There's only everything that was, and everything that will be, and it all exists, and it's all predictable. And there's only one future right?

Now, if that were true, and this is a this is something that was noticed already by Epicurus, the Ancient Greek philosopher. So he was arguing with Democritus, who was one of these philosophers who was advancing the idea that the world is made of atoms. And his idea was that the paths that the atoms take, were predefined, and never changed, never deviated from what they would be right. So very much a deterministic idea. And Epicurus already saw that if that were true, it would rule out the possibility of free will or agency of any kind, because the there are no possibilities in a world that looks like that. There's only one future, there's no choices, nothing matters because it couldn't ever be any other way than what it would be. And I think that that's right, I think Epicurus was correct. And so people who argue that free will or agency is compatible with the idea of pure determinism, I think they don't make a convincing case, because there are no choices.

But the other thing that Epicurus also noticed is that if there's some underdetermination, at the low levels, if the “atoms swerve,” as he said, every once in a while, then what that means is that the low level details plus the laws of physics underdetermine the future, right, they're not going to completely determine it, which means that the way that the system is organized, the macroscopic level, can constrain that possibility space. And those constraints can act as causes. And they're every bit as causal as the physical forces themselves. In fact, you don't get any physical forces, unless there is a certain organization of matter that generates them. I mean, the reason this computer works is because it's organized a certain way that constrains where the electrons are flowing. And with a function, right, I mean, it's designed in a certain way to, give some functionalities. And what I would say is that living systems are the same as that. They're organized a certain way, with certain functionalities and their organization is a is basically doing work, to constrain all of the bits to keep being in the pattern that they are, that constitutes the organism, that almost could be a definition of being a living thing is that it keeps on it keeps on being itself, and it keeps on making itself itself and constraining its bits to stop them from being in any other organization.

Art Woods

Yeah. So I'm very sympathetic to this point of view. And it's sort of intuitive for me that there can be these high-level, top-down causes on lower level aspects of the system. But let's dig into that a little bit more in the sense that you pose it in the book. And that's how do we escape from different forms of determinism, the sort of, you know, billiard ball determinism, that, you know, makes a lot of sense. And it's, I think it's a little bit hard to escape from. So let's talk about modes of escape. And if I can just sort of outline what I see is the sort of main categories of escape. One is quantum indeterminacy, and maybe we could just talk briefly about that without going too far into the details. Another is, you know, thinking about complex systems and just the fundamental unpredictability of complex systems and you know something about noise or randomness at molecular levels in in systems that maybe isn't quantum but as some other kind of noise that just gives an inherent indeterminacy to the way these physiological systems operate. So do you want to start and just maybe frame that and talk about quantum stuff first?

Kevin Mitchell

Yeah, so well, so the determinism idea, you know, comes from the notion that the lowest level that we know of so far of quantum fields and particles and so on, is, is completely deterministic in saying how a system will evolve. Which is a weird thing to say, because the Schrodinger equation, for example, which describes the evolution of systems like that, is deterministic only in the sense that it gives you a particular set of probabilities for how the system will look if you observe it. And then when you observe it, it takes on one of those probabilities, but at random, as far as we can tell. And so it seems like physics actually says, of the deterministic view that actually things are fundamentally indeterministic when they actually evolve in the real world, right?

Now, some people would say, Okay, well, that's just at the quantum level. And that doesn't bubble up to classical levels, which is the level of objects like the size of you and me, because all that quantum weirdness, just kind of, they all get entangled with each other. And then they become, it averages out all the noise and at classical levels, things behave deterministically. And it's interesting, if you look at the history of physics, that hasn't been shown, right, that's not that's not a just truly shown from experimental data. It's basically an assumption or an idealization that came into physics historically, at a certain point, and it even came into mathematics historically, at a certain point.

And I won't go into the details, cause it's a bit esoteric, but there's an idea that the numbers that describe a physical system, the state of a physical system right now, at any moment, would have to be infinitely precise. That is, there at the far decimal points of whatever numbers there are would have to be infinitely given all at once, all at the same time, in order to contain all of the information about the future states of the system. In a system, especially in a complex system that has some chaotic behavior, or sensitivity to initial conditions, or some nonlinearities, tiny differences can make a huge difference over time.

And, you know, you may be familiar with that from the sort of famous experiments that Lorenz did with the weather, right. So we had this model system for the weather. And he set it running and he went away for lunch, and he came back and it turned out, it had gone like two different ways when he did it two different times, even though the numbers were the same, but they truncated at a certain decimal point because the computer system he had couldn't deal with any more information. That imprecision just at some point got amplified, right?.

I mean, in some systems, it's not like that, right? The orbits of the planets are not like that. They don't have nonlinearities. They're not chaotic, they're, you know, just a couple of bodies interacting. And so they're, they are very deterministic at the scale that we care about. But systems like us, and especially like our brains are not deterministic in that way. And they are sensitive to initial conditions. And so first of all, that gives some openness to the way that your brain states could evolve. And secondly, then there's also the notion that we just have other sources of noise and randomness, either the quantum stuff does percolate up. So either that there's noise at that level, or there's thermal agitation, and there's things just, you know, proteins binding and unbinding, and ions diffusing around and all kinds of stuff.

Art Woods

Yeah. In your chapter, I was trying to make the connection between the quantum stuff and then this stuff about like, you know, thermal noise and random diffusion of molecules in a system and how that constitutes some kind of background noise against which the physiology is acting. And my thought was, well, like, is all of that thermal noise? Is that really random? Or is it just that we lack knowledge on, you know, the directions and momentums and, and states have all of those interacting particles. And if we had all of that information, then it would appear non-random and much more deterministic that it is. And I understand that you're saying that there's this quantum effect that is separate from that.

Kevin Mitchell 20:13

I mean, it's interesting. I've asked lots of physicists that question, you know, where does the thermal noise come from? And they'll say, “Well, it's just, you know, because there's a bunch of atoms bouncing around because they're vibrating and so on.” And I asked: “Okay, where did that come from?” And it doesn't, you won't get a satisfying answer.

And I think the answer is, I think it's not known. I'll say two things. First of all, I do take the quantum indeterminacy, to be fundamental. It just is that way, that underdetermination, I think, just is that way. And secondly, if you look, from the organisms point of view, it doesn't matter, right? If they've got noisy, jiggly components, then the challenge is actually flipped, right? Rather than saying, Oh, look, there's we take determinism as true, and then we have to ask, Well, where does the freedom come from? How could an organism have choice? In a world where possibilities don't exist?

But if you flip it, and you realize, actually, no, why take that, as the premise physics doesn't say that, you know, physics says indeterminism is true. Take that as your premise, now you've got another challenge, in that the organism is there, it's got these noisy components, it's got the indeterminate world that could go, the future could go all kinds of different ways. And it has to make happen, what it wants to happen. It has to corral all of that noise, and channel it in such a way that it is in control of what happens. And so, you know, thinking about those constraints, again, the idea of what constraints do is they narrow the possibility space of a system. And the system in this case is the organism and the bits of the environment it's interacting with, then I think that you can think of that as the process, which is actually what agency entails, is narrowing that possibility space, so that what happens is what you want it to happen.

Marty Martin

I want to challenge you, Kevin, to put this in as explicitly biological form that you can, and especially evolutionary, we've talked to several people on the podcast about similar kinds of things. And you have a series of fantastic chapters in the book, talking about the sort of origins and evolution of the various different senses, their integration, you knos, these kinds of things. But tell us a story about the early evolutionary history of life, the first sensory modalities that had to be engaged to keep these systems from falling apart. I mean, Nick Lane, for example, he has a decent story about this, and we've talked to him about it. But how do you see that having happened? How do we get from simple single celled organisms to lay that path to talking about brains and humans and such?

Kevin Mitchell

Yeah, I mean, I think Nick Lane has a great scheme for the origin of life, along with Bill Martin, I think they're fairly congruent. And that's the sort of vision that I lay out in the book of going from, from geochemistry to biochemistry, where you start in little rocky crevices that have some flow of free energy, basically, proton gradients from the seafloor. And then you can kind of concentrate biochemistry, basically, in small little areas where you can get more complex chemicals arising. Ultimately, so that at some stage, and this is all extremely speculative and hand wavy, they developed the ability to make this lipid membrane, and then they could be sort of free-floating, right, but they needed, they have the vestiges of that the history of that free energy source, in that that's this, that's the source of free energy that all cells use, is they generate a proton gradient themselves, and then they let the protons come back in through this protein to make ATP.

So what the cell is doing then, right is it's effectively doing thermodynamic work, taking in free energy, to keep itself organized and keep itself out of thermodynamic equilibrium with the environment. And in the process, what it's doing is it’s causally insulating itself from what's going on outside it, right? It's not in an even causal flow with everything else. There's a real barrier there. So stuff can happen outside that doesn't impinge on what's inside the cell.

Now, of course, what happens is stuff that happens outside that the cell wants to know about, right? Because it could say run out of food. So simple celled organisms have to get food, that's how they get their energy to create that, that proton gradient themselves. And so if there's just loads of food around, then great, you can just sit there and eat the food right? But conditions often will change. And in fact, they'll often change because of the activities of the living organisms themselves, right? They use up the food, they make too much waste product they divide too much. And then being able to move is a good trick.

So even, you know, simple organisms like bacteria can move around. And they do that in the service of homeostasis, right, they move when their internal state is not where they want it to be. So they'll have a, you know, sensors that can detect, say, certain sugars outside them. And then they relay a message inside that makes their flagellum this sort of outboard motor, long filament that they have rotate one way or another way. So they have this beautiful biochemical system that allows them to follow a gradient, move up a gradient of a food source. So all of that looks, first of all, purposeful, right? There's some meaning sort of embodied within the system. It's not the case that the bacterium senses the sugar and realizes or apprehends, there's a sugar out there, oh, I should make my flagellum do this, right? It's all pragmatically sort of coupled. But that meaning is the meaning of the signal is basically approach, or there's other signals that are the meaning is avoided. Right? So if you think about what organisms have to do to survive, you can think they have to know what is out in the world, and what should I do about it? I think that actually evolution solved the second problem first, the “what should I do about it?” It's just wired in as these control policies, very pragmatic kind of situation. And the “what is it?” only came later with the evolution of nervous systems?

Art Woods

Yeah. Can I ask, you know, if you imagine a bacterium sensing a gradient of food? And I think you're saying it has some agency in making the decision about whether to go up that gradient or not. And is the question, is that deterministic? Or is there sort of, you know, the true range of possibilities of following that gradient or not following that gradient? And does that emerge from some underlying noise in the system?

Kevin Mitchell

Yeah, you can think of this system and people have worked it out in great detail, that you have these receptors that sit in the membrane, you've got some proteins inside that are that transduce a signal. There's no physical force, there's no energy being there. It's just they bind something on the outside, and they change their conformation on the inside. So it's really an informational kind of causation.

Art Woods

It's a kind of abstraction of what's going on outside.

Kevin Mitchell

Yeah. And the system then is configured in such a way that when this happens, such and such will be the response, right? And you can, you can in the lab, you can isolate this pathway, such that the only thing that these things are exposed to is this sugar, and you're tightly controlling everything else. And that makes it look like a really linear, reflexive mechanistic kind of pathway. Where you would say: “Well, look, the bacterium is not doing anything, it's just being pushed around by its internal parts,” right? The parts do like this, that moves the bacterium that way, the bacterium is not in charge of anything. It's not exercising any real time agency.

But I think that's a forced perspective of the reductive nature of the experiments, because of course, bacteria in the world encounter all kinds of things, right, there's loads of signals, they got receptors for lots of different things out in the world. And the responses to them are very context-dependent. They depend on cell crowding, and temperature and osmolarity. They depend on the internal state of the cell. They're integrated all at once, and they're integrated through time. In fact, that's exactly how the cell can follow a gradient. It doesn't do it by a spatial comparison, say between the front of the end. It does it by a temporal comparison, how much is there now versus how much was there a minute ago. So it's carrying its history, its recent history with it, as well as its ancient history in the sense that that's why it's configured in the way that it is. And basically, we have all these relations at once. And they're all very context dependent.

And I think if you decompose it, you miss the logic of what the system is doing. First of all, it's proactive. They’re never just sitting there, waiting for a response. They’re not a passive stimulus response machine.They're endogenously active, they're accommodating to new information, but lots of sources of information at the same time. And then they are in a collective sense, doing what you can call it a basic kind of cognition. What is the state of the world? What is my own state? What's the best thing to do? That's effectively, that's a reasonable definition or description of cognition. It feels to me like even in that simplest behavior that we know about, in the simplest organisms we know about, I think that's justifiably called agency. And I think we're right to think that the agent, the organism, is a cause of things. It's not just being pushed around by stuff in the world. And it's not just being pushed around by its own parts.

Marty Martin

Would you consider the parts of those organisms? This is sort of circling back to the old question, but are the components of the bacteria, are any of those agential?

Kevin Mitchell

I don't think so. I mean, you can say, they're sensitive to, like the receptor proteins, are sensitive to something in the world, right? The flagellar motor proteins are sensitive to the state of the internal bits. I don't think we get anything by calling that agential.

Now, where this gets interesting is when you have multicellular creatures, because they are made of individual cells, which as we've just been talking about, can be extremely agential, and they often have their own agendas. And what's interesting when you get a multicellular organism, which is of course, where life went, is that you get all of these cells that they either come together to form a multicellular entity, like dictyostelium slime mold, individual cells fused together to form this multicellular slug that moves around as a unit. Or you know, in most animals, they start life as a single cell, and the embryo divides and the cells just stay stuck together. The individual is the whole thing. And the cells, I think, retain some agency, but they have to cede a lot of that agency to the collective. And of course, you know, when it fails, it, that's when you get cells going rogue and causing cancer and following their own agendas instead of the agenda of the whole organism.

Art Woods

You mentioned multicellular organisms. And so I think that's maybe a good segue to thinking about a few different branches of life and move from things that have biochemical networks that are showing agency and making decisions, maybe to things that have nervous systems and brains, and that are using those to make decisions. And so maybe let's talk about, just briefly, about the evolution of nervous systems and sort of what that opens up in the agency space.

Kevin Mitchell

Yeah, so when you get multicellular organisms, they now have some different challenges, right. So the fitness of the organism, the, you know, what natural selection cares about is that the whole thing survives, that it passes on its genes through just the germline cells, but all the cells have to be on the same team, basically, for the organism to survive. So one of the things that multicellular organisms are going to move around, is that they need to be able to coordinate their parts. And there's lots of ways they can move around, they can squish around, they can walk, they can fly, swim, whatever it is, but it's all quite coordinated. And so the earliest kinds of multicellular creatures that we know about, probably had some systems of sort of contractile cells like sponges do, for example. So they don't move around in the world, but they moved their bits in concerted ways. But then you got other things that can move around in the world. And it turns out that when you want to coordinate all the bits, and you want to do it fast, that actually neurons and muscles are a really good way to do that. So neurons and muscles were kind of invented from sort of from sheets of just myoepithelial cells that had some electrical conductivity, just with their neighbors. Now, you could get neurons which can have electrical conductivity over long distances in really specific ways, right, it's not just a wave, it's very, very specific control.

So if you think about something like a hydra, these little marine sort of shape like a cigar with a bunch of tentacles at the top, they look like Sideshow Bob from The Simpsons actually. Those things have what's called a nerve net, it's a sort of a diffuse. It's actually more complicated than it looked at first, it's probably like three or four different nerve nets that are intercalated with each other. But they don't have a brain. There's no condensation of neurons into a particular area. But they use that to move their body. They have a bunch of different behaviors they can do, they wave their tentacles, and they catch prey and so on. And so it's the first kind of glimmerings, I guess, or developments of the nervous system as a control system. And what we want to think about, again, is what is the point of behavior? It's to manage homeostasis, or allostasis, if you want to call it that, where it's not, it's not remaining, always the same, but within some range, right? So that involves control, right, and picking one thing to do at a time.

So if we move to something, that's a little more sophisticated like a C. elegans nematode, so this little worm, they have, they're like a thousand cells. 300 of those are neurons, and they're laid out along the body walls such that they can control muscles, so the worm can move forwards or move backwards. But there's also a condensation in the head. And they can sense various things–they sense touch, and smells. And when they sense those things, then they do some sort of internal integration. And then they send a signal to the command neurons that say, move forward or move backwards. So I think what you have in that design is a few things. You've got the sensors and the motors as you had in a bacterium. Now you've got a few more layers in between them. So now you've started to separate sensation from action. And you've got these levels of neurons that are integrating those signals, right. And so what they care about is what's the total amount of good stuff over here, or bad stuff over there. And then, and then they have these command neurons. So when they're choosing between these actions, they're competing with each other, right? They're set up in a circuitry, if we've one of them is active, it's inhibiting the other ones, right?

It actually involves not doing all those other things, and just doing this one, one thing out of that, that said, and for C. elegans, it's a very small set, there's a few things that can do. You know, for as organisms get more complex, that set gets bigger, but they still have to do one thing and inhibit everything else. So I think you start to see some of the design principles there that get scaled up as things get more complicated.

Marty Martin

So how then does meaning get instantiated to try to draw these things together? What does? Where's the meeting come from for the C. elegans?

Kevin Mitchell

Yeah, so it's really, that's super interesting, because now we've gotten to a stage where a little bit we're a little bit beyond these just pragmatic couplings, right? We've got some internal patterns. Now, even in a C. elegans, they're pretty tightly coupled to action. It's not obligate. Right. It's not reflexive. It's quite integrative. But they're still, you know, if they're thinking about anything, they're not thinking about much, right? They, I mean, they can't be because their sensory umwelt, like what they can detect is quite limited. It's limited, in fact, to things that they can touch. Right? So mechanical sensation, or odorants, that they literally have to bind the actual thing that they're detecting.

Art Woods

Right, they're not seeing, they're not hearing, they're not.

Kevin Mitchell

Exactly. And so I think what you get with, with the evolution of vision and hearing is a whole next level of processing that has to happen in order for the organism to make adaptive use of that information. So vision and hearing are distant senses, but they're not sensing directly, the objects that are causing the, whatever this disturbance is in the electromagnetic field, or the vibrations in the air or the water, right? The organism doesn't care about the vibrations, it doesn't care about the photons, what it wants to know is what's out in the world?

Art Woods

Right. What's going to eat me? What am I going to eat?

Kevin Mitchell

That's what it needs to know, right? That's the information it needs to get around in the world. So it needs to make sense of its sensory data. And that requires these levels of processing, eventually inferring the presence of objects out in the world. And then it's reporting that to the rest of the brain.

So there are internal patterns and parts of your, you know, higher level parts of your brain, visual brain, which report shapes or objects, or particular types of objects, like a face, for example, or faces in general. This is a philosophically loaded term, but I think it's perfectly reasonable to call those representations. And what I mean is that patterns of neural activity, that mean something to the organism. And the important thing is that it's the meaning that is going to be used to inform, to inform behavior, right? It's not, it's not the low level details of neural firing. They're actually arbitrary and incidental. And they'll change over time. In fact, the nervous system is set up precisely not to care about those, those little details, which is good, because as we were saying before, it's got noisy components. And so being able to do robust cognition, where the objects of cognition are meaningful things to the organism, in order to be able to do that, you have to sort of average out over all this noise, you have to compress signals and generate real cognition where the meaning is what the system runs on.

Art Woods

Yeah, I really like this idea of multiple realizability that you've laid out in the book. And I especially like the computer metaphor, the software metaphor that you use for that, which is, you know, we sort of already touched on this, but, but for example, we're all talking to each other and using this sort of common platform. I'm on a PC, I know, Marty's on a Mac. But like the details, the physical details of those computers don't actually matter very much. It's whatever is instantiated in the software that's running this thing that we're talking to you right now. Right?

Kevin Mitchell

Yeah, absolutely. And I think the other I mean, that's also a really nice example of top-down causation, right? In a non mysterious, everyday kind of way, right? The reason these computer programs work, is because somebody wrote them to have certain functionalities, and those functionalities are executed by constraining where the electrons flow in the circuit boards of these computers, right? So the meaning of patterns is what's important there, the details? Not so much.

Marty Martin

All right, let me see if I can put these many things together. So are you saying that the meaning is really in the arrangements of, you know, all of these different entities, these stimuli, photons, or whatever it is, are coming entering into these systems, there's some perception, and then interpretation, integration of different types of cues to weight them in different ways, as informed by evolutionary history, which is going to be you know, the ways that the these architectures are set up, but then more short term kinds of things like, you know, the learning that's happened and changing the relationships among parts in that way. I mean, is that capturing the important parts of what you're saying? What am I leaving out?

Kevin Mitchell

It is, yeah, and you've touched on a few important things that I haven't talked about yet, which is, when we are perceiving objects in the world, we're recognizing them. And we do that by linking to prior knowledge. So as you know, infants, we explore the world, with our eyes, with our hands, we're learning about the regularities of different kinds of objects. So when we see something that looks like a teddy bear, we have an expectation that it's going to weigh this much, and that it's going to be squishy, and kind of soft, right?

So it's not just, it's not just a neutral sort of pattern. It's, I recognize what that object is, and I know what I can do with it, right? So it's not just passive, like a camera, we're seeking, we're exploring, we're figuring out what's out there, all in the service of knowing what to do next. And so in, you know, more complicated organisms, that learning, of course, is the huge resource that we can draw on. And we do that throughout our lives we learn, like what an object is, what category it may belong to. So we have a kind of a hierarchical map of reality. And so, you know, bears are, are a mammal, and mammals are an animal. And knowing that it's an animal is useful, because I can relate it to a bird or a snail, and it has some common properties, right? So we build up these kinds of schemas of objects, and then what's out in the world, the nature of things that we've encountered before, along with causal relations, like I said, "What can we do with it? What can it do to me?" If it's a teddy bear, I can pick it up, if it's a grizzly bear, it can pick me up, that's a good thing to know the difference between.

And then of course, we're also learning from sequences of events. And this is really crucial. So we have this fun of reinforcement learning, where when we've encountered some scenario, there were a bunch of options that were open to us. We did A rather than B, C, or D. And A turned out well, so the next time I'm in that scenario, I'm going to do the same thing, right? Or at least it's more likely that I'm going to do the same thing. Whereas, if A turned out badly, then I'm just not going to do that again. And instead, I'm going to explore B, C, and D, and maybe now W as well, right?

So all of those things are happening. And I think the key thing here, if you want to know, how can it be that an organism is a causal agent in the world? How can it have causal power unto itself? The reason it does that without violating any physics or getting a free lunch is because it's been paying attention. Right? It's been causally intervening around the world, seeing what happens, storing that knowledge, which has a cost, and using that then to exercise causal power in the world.To behave more adaptively based on what it knows, based on its experience.

But also then, you know what, what are my ongoing goals because behavior is not just binary choices, moment by moment, we are not just instantaneously reacting to things. We like any other organism, well, even more than any other organism have ongoing projects, right, where we have sustained behavioral activities.

Art Woods

Oh phew

Marty Martin

That's an important thing to say. Yes.

Kevin Mitchell

Right, okay. But secondly, it's not just neural mechanisms. Right. So someone could say, look, neuroscience is showing that what you're calling motivation is just this, this neural circuit pathways firing? That's not you know, that's all. Why do you need it call it a motivation? It's just electrical activity firing there. And we can reduce everything to the, to the firings of these neural circuits. And that's where the meaning and the multiple realizability comes into play. If it were the case, that the details of the neural firing was all that mattered, and fine, but you should be able to then counterfactually change the details and change the outcome. And in fact, you can often change the details and not change the outcome, until you change the pattern from A to B, because the pattern means something to the organism, right? So it's part of the organism's reasons for doing things, the neurons don't have reasons for doing things.

Art Woods

And this is actually going to be quite a practical question. I think we've been talking about a lot of abstract and very theoretical stuff. But let's say a listener really buys this idea of top-down causation. And the fact that, you know, it's the integrated whole organism that in some way is an agent, and it's making decisions about what to do in the future. What does that actually mean, for the kinds of research that you carry out? And how is it different than the typical, more reductive stuff that many of us do?

Kevin Mitchell

It's a great question. I think, actually, that there is a move to a much more holistic systems kind of biology that's happening right now, that is just driven by technology, right? So we can do single cell RNA sequencing of all the RNAs within any single cell, right? Or all the cells of an embryo as it's developing or…

Art Woods

Right, so you get a snapshot of the system much more. Yeah,

Kevin Mitchell

Exactly. Or we can record from all of the neurons in the whole organism, while it's behaving, right. I mean, it used to be we could record from two neurons in a dish. Or we could do you know, pharmacology and look at behavioral patterns over time and stuff like that. But now we can really watch a living organism moving around, while we're recording its whole brain, I mean, for simple organisms.

My feeling is that what has to happen then is that actually this sort of philosophy that people apply the concepts that people apply, almost follow the technology, right, when we could only do reductive, controlled experiments, I think people had this very sort of physics-ey, mechanistic view of things because it fit the data and the way that they could think about it. Tt wasn't necessary to make any theoretical commitments to go along with the methodological restrictions, but I think people tended to do that. Whereas now I think what's happening is that people are sort of rediscovering things like, you know, cybernetics, and control theory and process philosophy and so on. Right, these are the concepts that we're going to need to understand the kinds of data that we're getting. So that now we can see not just what's happening in the system, but what the system is doing, right? What's the logic of what the system is doing?

So I have great hope that Systems Thinking generally is going to reemerge. Now that it's practical, it's useful now, right? And, in fact, we have to have it. And fortunately, there are ways that they're not just useful for us, you know, we can look at the data we can say for neural populations. When we analyze all this population data, what we see is that we can reduce that down to a couple of patterns. And that is, for us, useful statistical technique. But also, it seems to be what the organism is doing. The organism can't make sense of all that high dimensional data, it also has to do some kind of compression.

Art Woods

Right, it's abstracting it down to a few choices.

Kevin Mitchell

And exactly and this is where the meaning comes in. Right? If one population is looking at another one. It doesn't matter about all this sea of noise, it matters whether it settles into A or B. That's what it cares about.

Marty Martin

That's a great answer, I guess, in the sense that I agree with you. But one thing that I'm surprised to have heard you barely mentioned you through dynamic in there towards the end. But I found one of the coolest elements of your book, to be the slinky model. And I don't want to spill the beans on that too much, because I think maybe we can try to connect that to free will, you know, the subject of your book, and let's start mapping to that. But to put a fine point on it here, I'm going to challenge you to at least come up with a hook sentences for this grant proposal. If we're going to embrace systems. As you know what I as an Assistant Professor, I'm going to put out as my first grant proposal, and I'm going to embrace things change through time. And we'll talk a little bit more if that's not already obvious. We'll talk a little bit more about that in a second, what do we measure? What are you going to do? Like, just without spilling your own beans about the grant proposals you're going to write? What are the things that somebody should go for?

Kevin Mitchell

Sure. Okay, so depends. Let's say you are measuring a single cell, right? That a, you know, an E. coli that's adapting to different things, or you're measuring neurons and an animal that's adapting to different things. So one of the things I would do is move from the simple, reductive, highly controlled experiments, which gives you very sort of linear data to more complex things. So right, rather than challenging your bacteria with one sugar thing and measuring its gene expression, let it loose, let it explore much of you know, put it in a biofilm, put it in something more realistic. You know, and I think people are doing that I think people are realizing in neuroscience, that they need to have more naturalistic ecological settings, you can't just ask a rat to turn left or right in a maze over and over and over again, and expect to learn really much about actual behavior.

Art Woods

The rat gets pissed, I mean.

Kevin Mitchell

Exactly! It's like, "again?" "For real?" So people are moving to that. And again, it's enabled by these new technologies, where, you know, you can record from a bunch of neurons, but in a freely moving animal, because there's these mini scopes. And these, you know, calcium reporters that show you all the activity and so on fantastic technology. That means now you can track the same animal, you know, over weeks, in a big complex kind of enclosure, when it's encountering other members of its species. And it's, it's in kind of different states through time, and it has different sorts of goals, and so on. And then we can see, you know, kind of track the responsiveness to, to different things. So, I think I think people are moving that way, yeah.

Art Woods

Well, let's move in this last section to grappling with the idea of free will which we really haven't approached in depth yet. And let me just ask, so what is free will? And is free will the same thing as agency?

Kevin Mitchell

Yeah. So obviously, free will is just loaded with baggage right, philosophically and metaphysically, as a term. I tried to approach it cautiously. And actually, what I mean is, some people define it in a really absolutist way, like, you're only free if your actions were not subject to any prior causes whatsoever. Right? That you have to be completely free to have free will. And that just breaks down immediately. Right? The minute you start thinking about that, it becomes incoherent because it would mean you're, you're behaving for no reason. No information, right? With no goal,that's a prior cause. And it will just be random behavior, right? Nothing that is a self that's persisting through time can behave like that. It would just be dead. So that's a non-starter for me. And it's just, there's no point going down that route.

The other view is that it's only free will if "you" in some mystical sense, right, the material kind of you, that sits inside your head somehow, is the thing that's in charge. It's not okay, if neuroscientists show it's your brain that's active when you're making these decisions, right? Then you can say: "No, it's just your brain doing it, it just tells you afterwards." And it's a weirdly duelist idea where somehow the self, the mental bits are, like, supposed to be a separate entity from what's going on in your brain. Whereas I think you can just think of them as yourself as being entailed by what's going on in your brain, not produced by it's not some extra thing. Selfhood is not something you have, it's not something you are, it's something you do, right? It's that sort of activity.

So, I like to just start with a phenomenon, and what we're talking about is that human beings, like other animals seem to be able to do things. They seem to be able to control their behavior. They make decisions, they select actions. And of course, we feel this ourselves in our own lives, and that we spend most of our time going around thinking about what to do, you know, deciding what should I do now? What should I do, then? Or wondering why somebody else did something or said something, right we're thinking about our own reasons, we're wondering about other people's reasons, and so on. So I like to call free will just the manifestation of agency that happen to exhibit, but it has some characteristics that are unique to humans that I think, elevated above what, what's there and other animals.

So first of all, our cognitive time horizon is almost infinite. Right? So we talked about worms inhabiting the here and now, right, because they only can sense things right around them. So they only can think about thing on that kind of timescale, right? When we got vision and hearing, well, you can see something coming from a mile off, right, it pays to be able to think about things that in time are a mile off. And so we developed, we animals, developed cognitive resources that allowed planning and prediction of things over longer time frames. And then that allows sustained behavior over a longer time frames. So you can have goal directedness, for something the next day, or a week from now, or a year from now, and so on.

I also think we abstract, a lot more. So we can think about abstract things, not just that bear, we can think about bears in general, not just that animal, but animals in general, and so on. And we can build up these webs of causal relations and learn lots of things, right. So we're doing sophisticated cognition, of the same type that other animals are doing. But fancier right.

Now, the really fancy bit, I would argue is that we have an extra level or two, where we're not just thinking at these extra levels, they're not just getting information about objects in the world, or beliefs or our current internal states, they're getting information about our thoughts about those things. Now we can think about our thoughts, and we can reason about our reasons. Right? So it's not just we have reasons, we can do reasoning. For example, I tried something it didn't work out well.Was that because it was the wrong thing to do clearly? Or did I have the wrong information? Or was I unlucky? Right? So we will learn from things where we're quite confident about what was going on, and our information was good, and so on. And we don't learn as much from things where we have no idea what's going on. All of that is adaptive, just as a control system, kind of thing that metacognition is adaptive.

Where it really gets us is when we have the evolution of language, and culture. And we can tell each other what our reasons are, right? I can say, hey, I want to go over here. And you can say why? And I can say, well, there's food over there and look, yeah, great, let's go. Or we can we can coordinate action that gives us collective action. And of course, yeah, as a hyper social species, we can think about well, what's Art thinking about? Well hey Art’s thinking about what Marty is thinking about. And Marty’s thinking about what I'm thinking about. So you can you can build up these complex kinds of pictures that I think enable human culture but are at the same time supported by those cultural kind of interactions in a way that ultimately I think maybe lets us transcend the immediate biological imperatives, and as a species, and as individuals within within that species, think of things like the stuff that we're talking about right now and think of goals that are, you know, decades, decades often the future.

Art Woods

Yeah, yeah that's beautiful. Yeah. Let me ask you about a particular strain of anti free will thought. And this is espoused by a number of people, among them, Sam Harris. And the idea is that a lot of what's happening is happening in our unconscious brains. And that perhaps, you know, most of the sort of causes or the perceptions that we have are things coming out of our unconscious, and decisions are made in the unconscious without any input from us, in some way. And our conscious brain is then constructing a story that sort of makes sense of all of these things that are coming out of our unconscious. How do we escape from that as an anti free will idea?

Kevin Mitchell

It's tricky. It's tricky. Let me say a few things. First of all, I think there are cases, particularly pathological cases, where people, you know, they've had a stroke or the, you know, they had their cerebral cortex, you know, cut through the middle, things like that, where they really don't know what's going on. And they will make up stories about why they did certain things.

Now, some people kind of extrapolate from those scenarios to say, "Well, that just shows we never know what's going on. And we always just make up stories." And to me, that just simply doesn't follow. It's like saying, because optical illusions sometimes fool our visual system, none of our vision is reliable, right? Well, if it wasn't, if it wasn't a good guide to what's out in the world, we wouldn't have it and we'd be dead, right? The same thing is true, right? If our thinking about our reasons, and other people's reasons, wasn't adaptive, then it wouldn't be there. Right? We wouldn't expend so much energy, just telling ourselves stories for no good reason. So just on the experimental, just the empirical evidence that informs that view that you just articulated, I don't think it's as strong as, as the proponents of that view actually make out.

But they do touch on something really important, which is the idea that, first of all, we're driven by our subconscious. And secondly, that we're driven by the configuration of our brain at a current moment that we had no hand in, right? So let me take those two separately. First of all, the idea that we're driven by our subconscious– Great!

Art Woods

Drive away!

Kevin Mitchell

Absolutely. A load of stuff is done with our subconscious, which is brilliant, okay, we have learned we've done loads of work, we've been through lots of experiences, we've learned loads of stuff, we know what to do. And we can offload that thinking to these habitual sort of subconscious processes that guide a lot of our behavior without us having to spend loads of effort and, and crucially, take loads of time to deliberate between things from first principles every time we encounter a scenario. Okay, so it's super, super adaptive to have your subconscious sort of, you know, driving lots of stuff on autopilot.

And again, it doesn't mean we're always on autopilot. It's just a fallacious reasoning to say, because we're sometimes like that means we're always like that. And so of course, we can engage deliberative control when we need to. Daniel Kahneman in his book Thinking Fast and Slow refers to this as system one and system two, which is, you know, it's too stark, a dichotomy. But it is a nice heuristic for thinking about the idea that most of the time you've got automatic habitual behavior, but sometimes we can engage system two when we need to, like when there's a detour on your normal route, you know, that you're driving home? And you have to think, right, you were on autopilot. Now you have to think about it.

So that's the first aspect that just because some of our thinking is subconscious doesn't mean all of it is and in fact, it's good that most of it is. And then the second bit is really tricky. I think the second bit is actually for me a stronger kind of challenge to our views of free will than any issues of physical determinism or reductionism. I think those can be handled. This one which I call biological fatalism, is a little trickier, right? Because it's true, right? We do have a genetic makeup that gives us some psychological predispositions. We are shaped by our experiences, and we're shaped by just the way our brain happened to develop. And we're shaped, of course, you know, by all of the circumstances that we've been through and the circumstances that we happen to be in any moment. So we absolutely don't have all the degrees of freedom in the world, right? We have a constrained set of options available to us that are informed by these different things. Now, the question is, are they determined by those things? Right? Are they fully determined just by say, your personality traits, which is fully determined by your genes? So the answer is no, right? Your personality traits are partly influenced by your genes. And your personality traits partly influence your behavior in patterns through time, but what actually happens is still up to us.

But the interesting thing then what's the consequence of that? What's the implication? The implication is that as we're doing things in the world, we're learning things. We're shaping our own brains. We're shaping our own character, right, we're adapting to the world. So we have our personality dispositions, we're learning particular situational sort of things that make us feel good or don't feel good. So we're judging outcomes of our actions, given that we're more or less risk averse, say, fine, right? So we're adapting to the world. And we're doing that by shaping our own selves, through time and building our character, right? We're building policies, and attitudes, and more sophisticated dispositions, and heuristics and habits. And we're adopting commitments, and plans and projects, that all of which constrain our actions and guide and manage our behavior through time. And then at any moment, within all of that context, we're surveying the possible space that's still open, and making some decisions within that.

The upshot is that I think it's just not right to say the configuration of our brains right now or at any moment, is just because of things that happen to us. As if we were passive in that process. There were loads of things we chose. It's the accumulated effects of our agency, the exercise of our agency through time that collectively shape our character. I just don't accept the idea that moment by moment, we're just configured a certain way that we had nothing to do with.

Marty Martin

Kevin, so I'm trying to tie what you're saying. Now, back to the traits that you use to distinguish human free will from agency. And you just made a good case for that idea. But one thing that wasn't explicitly in your list, although you could lump it into one of the categories, I guess, but you know, in light of what you're saying, now, I want to just hear your specific take on it. What's creativity?

Kevin Mitchell

Absolutely. So let's think of scenarios that that an organism can be in, it can be in a very, very familiar scenario where it has a habitual behavior, it's all teed up ready to go, it doesn't have to think about it, it encounters this scenario or some kind of cue and does X, right? It can be in anything from there to an extremely unfamiliar scenario, where now it has to think what should I do here? Right? Now, maybe it can draw on some analogies with other types of behaviors, it's been in before and say, "Okay, so maybe I'll try what I did there."

Now, the way organisms, mammals at least, manage that is that they'll think of a few things that they could do, they will subject them to this evaluation system, which is basically effectively making predictions of the outcomes. So if I did A, given what I know about the world here, what do I think the outcome is likely to be? And then it would that be good or bad, right? So imagine, so there's some competition between those various things. It'll eventually select one option, inhibit all the other ones, and just do X, whatever that was. But it doesn't stop there. Right? Because it monitors. Well, bow's that turning out for me? Right? Am I achieving my goals? Am I being frustrated? Or is that satisfactory? If the organism is being frustrated, it may change course, because of course, you know, a behavior is not an instantaneous thing. It's something that takes some time. So during that time, it may say, you know what, this is just not working out, I have to do something different. And kind of go back to the cerebral cortex and say, “Hey, give me some new options here.” Right back to the drawing board, let's think a little more broadly.

And, and that's this, this touches upon an old model of free will from William James, who had what he called the two stage model where, you know, he basically said, “Yeah, you know, ideas do pop into your head, your conscious self is not controlling that necessarily. However, they're then subjected to this selection, where you then pick one.” So I think he said something like, “Our thoughts come to us freely. But our actions go from us willfully.” Right, so, so for me that, that feels like that makes a lot of sense, right? The organism in that case, is using the randomness, right, it's using the noisiness as a resource to draw on, when its behavior is, options are not so clear. And that's good. That's an exploratory kind of problem solving.

And of course, you can also do the same thing when you have a scenario where you just don't care, flip a coin, who cares? That's an efficient thing to do. In fact, what's important is that you do something quickly. Right? That's more important than this inconsequential decision.

Just to sum up, there's a kind of a range of decision making processes that we use in different sorts of scenarios. It's not always deliberative. It's not always habitual. And we can move between those, right? And we, in a sense, can control which mode we're operating in, in really interesting ways. But they're much richer than the kind of binary decision scenarios that are often proffered, when people are talking about the problem of free will.

Art Woods

So we're running out of time. And yet, I want to ask one more large question. And that is about consciousness. So do you explicitly have to have consciousness in order to have free will?

Kevin Mitchell

So I don't know consciousness is really tricky. It brings all kinds of other questions in along with it, right. I think what we can say is that when we're exercising the kind of rational deliberative control over our actions that I think people mean, when they say “free will,” those are normally conscious, right? If all of the activity was subconscious, it might support the idea that we are not really in charge when “we” is this metaphysical sense of ourselves, right?

But here, you're mixing up, not you, but people are mixing up two ideas, right? So one is like, myself, am I doing things? Am I an agent in charge of things, and we just talked about all the systems that allow agents to be in charge of, you know, selecting their own actions for things to be up to them. And then there's this other kind of idea, which is the sense of self, the sense of agency, which in humans involves this conscious kind of system, and actually can be impaired in in weird ways, in psychiatric conditions, and so on, and with drugs, and all that kind of stuff. So for me the sense of self, what it feels like to be doing things, whether you feel like you're the agent, you're the author of your own thoughts, and so on, your own actions. Is a somewhat separate level, from whether you're actually consciously deliberating about them. So I don't know if it's necessary. But I will say that when I say free will, what I'm talking about is conscious cognitive control. And that's why I don't use the term free will, just again, because of all the metaphysical baggage, when I'm referring to other organisms. I'm happy to say agency. And then it's really interesting to ask, well, what does consciousness get you? It's very speculative. And I don't think anybody really knows the answer to that.

Art Woods

If you had to say, which range of animals have consciousness and draw a line between the conscious and the not conscious, do you think that's possible or even useful to do?

Kevin Mitchell

It depends. Many people use consciousness just almost as a synonym for sentience. Which is, to mean two things. One, that the animal is responsive to things in the world. But also that it's something it's like to be that there's some inner experience, right? There's some sort of subjective experience. And I'm perfectly happy to say that all animals have some kind of subjective experience. It's not like ours, it's like theirs. What they may not have is this meta reflective level, where they're having a subjective experience about their own thoughts. Right? They may be having subjective experiences, about their sensations, and their inner states, and beliefs. But that's different from this meta level of having a sensation about your own thoughts that this reflexive, recursive kind of a thing? That's what I think sets us apart.

Art Woods

That's really interesting! And it feels to me like that almost is an empirical assertion about, you know, what consciousness is and is not that would be testable in the end. I mean, maybe, maybe in the end, we could sort of figure out ways of assessing, you know, what's going on inside the brains of other animals to know whether they're thinking about their own thoughts?

Kevin Mitchell

I think it would, it would be possible. And, and I think it's also possible, just to look at the design of nervous systems. And if we're going to allow that, say, a dog has a mind or a frog or a fish or a C. elegans has a mind, what could be on their mind? Right. And so it may be impossible to say that they are thinking about their own thoughts. But it may be possible to say that they couldn't be thinking about their own thoughts, because they just don't have enough neural levels to do that. Right.

So I think in a C elegans, you can say: "Look, it can't be thinking about things that are far away from it, because it has no way to detect them. And it can't be thinking about things in the far future, because it has no way to control that. And it just doesn't have enough levels of its of its organization, or they're not organized in the right way." But I think it's a super, super interesting question. I think it is an empirical question. And I guess the final point I would make on that is that it's not just like degrees of consciousness, or indeed degrees of agency. I mean, I draw an evolutionary line from some you know, the origin of life along the line that goes to humans in this book, but I don't explore the line that goes to insects or octopus or birds or other things, right, which may have different qualities of consciousness and experience, they must have in fact, and different qualities of agency, and so on. So I think there's much richer sort of tableau of, of mentality of cognition and agenetial behavioral control, certainly than I've considered in in this book.

Marty Martin

Wow yeah we opened the consciousness can there at the end. There's so many more questions I have for you, Kevin, and we really appreciate your time. So thank you so much.

Art Woods

Thanks so much, Kevin.

Kevin Mitchell

Well, thank you guys. It's yeah, it's been a pleasure. Thanks a lot.

Marty Martin

Thanks for listening! If you like what you hear, let us know via X, Facebook, Instagram, or leave a review wherever you get your podcasts. And if you don't, we'd love to know that too. Write to us at info@bigbiology.org!

Art Woods

Thanks to Steve Lane, who manages the website, and Molly Magid for producing the episode.

Marty Martin

Thanks also to Dayna de la Cruz for her amazing social media work. Keating Shahmehri produces our awesome cover art.

Art Woods

Thanks to the College of Public Health at the University of South Florida and the National Science Foundation for support.

Marty Martin

Music on the episode is from Podington Bear and Tieren Costello.