

**Lex Fridman Podcast #430 - Charan Ranganath: Human Memory, Imagination, Deja Vu,  
and False Memories**

Published - May 25, 2024

Transcribed by - [thepodtranscripts.com](https://thepodtranscripts.com)

**Lex Fridman**

The following is a conversation with Charan Ranganath, a psychologist and neuroscientist at UC Davis specializing in human memory. He's the author of *Why We Remember: Unlocking Memory's Power To Hold On To What Matters*. This is the Lex Fridman Podcast. To support it, please check out our sponsors in the description. And now, dear friends, here's Charan Ranganath. Danny Kahneman describes the experiencing self and the remembering self and that happiness and satisfaction you gained from the outcomes of your decisions do not come from what you've experienced, but rather from what you remember of the experience. So can you speak to this interesting difference that you write about in your book of the experiencing self and the remembering self?

**Charan Ranganath**

Danny really impacted me. I was an undergrad at Berkeley and I got to take a class from him long before he won the Nobel Prize or anything and it was just a mind-blowing class. But this idea of the remembering self and the experiencing self, I got into it because it's so much about memory even though he doesn't study memory. So we're right now having this experience, right? And people can watch it presumably on YouTube or listen to it on audio, but if you're talking to somebody else, you could probably describe this whole thing in 10 minutes, but that's going to miss a lot of what actually happened. And so the idea there is that the way we remember things is not the replay of the experience, it's something totally different. And it tends to be biased by the beginning and the end, and he talks about the peaks, but there's also the best parts, the worst parts, etc. And those are the things that we remember. And so when we make decisions, we usually consult memory and we feel like our memory is a record of what we've experienced, but it's not. It's this kind of very biased sample, but it's biased in an interesting and I think biologically relevant way.

**Lex Fridman**

So in the way we construct a narrative about our past, you say that it gives us an illusion of stability. Can you explain that?

**Charan Ranganath**

Basically I think that a lot of learning in the brain is driven towards being able to make sense. I mean really memory is all about the present and the future. The past is done. So biologically speaking, it's not important unless there's something from the past that's useful. And so what our brains are really optimized for is to learn about the stuff from the past that's going to be most useful and understanding the present and predicting the future. And so cause-effect relationships for instance, that's a big one. Now my future is completely unpredictable in the sense that you could in the next 10 minutes pull a knife on me and slit my throat.

**Lex Fridman**

I was planning on it.

**Charan Ranganath**

Exactly. But having seen some of your work and just generally my expectations about life, I'm not expecting that. I have a certainty that everything's going to be fine and we're going to have a great time talking today, but we're often right. It's like, okay, so I go to see a band on stage, I know they're going to make me wait, the show's going to start late and then they come on. There's a very good chance there's going to be an encore. I have a memory, so to speak for that event before I've even walked into the show. There's going to be people holding up their camera phones to try to take videos of it now because this is kind of the world we live in. So that's like everyday fortune-telling that we do though. It's not real, it's imagined. And it's amazing that we have this capability and that's what memory is about. But it can also give us the illusion that we know everything that's about to happen. And I think what's valuable about that illusion is when it's broken, it gives us the information. So I mean, I'm sure being in AI about information theory and the idea is the information is what you didn't already have. And so those prediction errors that we make based on, we make a prediction based on memory and the errors are where the action is.

**Lex Fridman**

The error is where the learning happens.

**Charan Ranganath**

Exactly. Exactly.

**Lex Fridman**

Well, just to linger on Danny Kahneman and just this whole idea of experiencing self versus remembering self, I was hoping you can give a simple answer of how we should live life based on the fact that our memories could be a source of happiness or could be the primary source of happiness, that an event when experienced bears its fruits the most when it's remembered over and over and over and over. And maybe there is some wisdom in the fact that we can control to some degree how we remember it, how we evolve our memory of it, such that it can maximize the long-term happiness of that repeated experience.

**Charan Ranganath**

Well first I'll say I wish I could take you on the road with me because that was such a great description.

**Lex Fridman**

Can I be your opening act?

**Charan Ranganath**

Oh my God, no, I'm going to open for you, dude. Otherwise, it's like everybody leaves after you're done. Believe me, I did that in Columbus, Ohio once. It wasn't fun. The opening acts drank our bar tab. We spent all this money going all the way there and there was only the -

Everybody left after the opening acts were done and there was just that stoner dude with the dreadlocks hanging out. And then next thing you know, we blew our savings on getting a hotel room.

**Lex Fridman**

So we should as a small tangent, you're a legit touring act?

**Charan Ranganath**

When I was in grad school, I played in a band and yeah, we traveled, we would play shows. It wasn't like we were in a hardcore touring band, but we did some touring and had some fun times and yeah, we did a movie soundtrack.

**Lex Fridman**

Nice.

**Charan Ranganath**

Henry, Portrait of a Serial Killer. So that's a good movie. We were on the soundtrack for the sequel, Henry 2, Mask of Sanity, which is a terrible movie.

**Lex Fridman**

How's the soundtrack? It's pretty good?

**Charan Ranganath**

It's badass. At least that one part where the guy throws up the milkshake is my song.

**Lex Fridman**

We're going to have to see. We're going to have to see it.

**Charan Ranganath**

All right, we're getting back to life advice.

**Lex Fridman**

And happiness, yeah.

**Charan Ranganath**

One thing that I try to live by, especially nowadays and since I wrote the book, I've been thinking more and more about this is, how do I want to live a memorable life? I think if we go back to the pandemic, how many people have memories from that period, aside from the trauma of being locked up and seeing people die and all this stuff. I think it's one of these things where we were stuck inside looking at screens all day, doing the same thing with the same people. And so I don't remember much from that in terms of those good memories that you're talking about. When I was growing up, my parents worked really hard for us and

we went on some vacations, but not very often. And I really try to do now vacations to interesting places as much as possible with my family because those are the things that you remember. So I really do think about what's going to be something that's memorable and then just do it even if it's a pain in the ass because the experiencing self will suffer for that but the remembering self will be like, "Yes, I'm so glad I did that."

### **Lex Fridman**

Do things that are very unpleasant in the moment because those can be reframed and enjoyed for many years to come. That's probably good advice. Or at least when you're going through, it's a good way to see the silver lining of it.

### **Charan Ranganath**

Yeah, I mean I think it's one of these things where if you have people who you've gone through since you said it, I'll just, since you've gone through shit with someone -

### **Lex Fridman**

Yeah.

### **Charan Ranganath**

- and it's a, that's bonding experience often, I mean that can really bring you together. I like to say it's like there's no point in suffering unless you get a story out of it. So in the book I talk about the power of the way we communicate with others and how that shapes our memories. And so I had this near-death experience, at least that's how I remember it, on this paddleboard where just everything that could have gone wrong did go wrong, almost. So many mistakes were made. And ended up at some point just basically away from my board, pinned in a current in this corner, not a super good swimmer, and my friend who came with me, Randy, who's a computational neuroscientist, and he had just been pushed down past me so he couldn't even see me. And I'm just like, "If I die here, I mean no one's around. It's like you just die alone." And so I just said, "Well, failure is not an option." And eventually I got out of it and froze and got cut up and I mean the things that we were going through were just insane. But short version of this is my wife and my daughter and Randy's wife, they gave us all sorts of hell about this because they were just ready to send out a search party. So they were giving me hell about it. And then I started to tell people in my lab about this and then friends and it just became a better and better story every time. And we actually had some photos of just the crazy things like this generator that was hanging over the water and we're ducking under this zig of these metal gratings and I'm going flat and it was just nuts. But it became a great story. And it was definitely, Randy and I were already tight, but that was a real bonding experience for us. And I learned from that that it's like I don't look back on that enough actually because I think we often, at least for me, I don't necessarily have the confidence to think that things will work out, that I'll be able to get through certain things. But my ability to actually get something done in that moment is

better than I give myself credit for, I think. And that was the lesson of that story that I really took away.

**Lex Fridman**

Well, actually just for me, you're making me realize now it's not just those kinds of stories, but even things like periods of depression or really low points, to me at least it feels like a motivating thing that the darker it gets, the better the story will be if you emerge on the other side. That to me feels like a motivating thing. So maybe if people listening to this and they're going through some shit, as we said, one thing that could be a source of light is that it'll be a hell of a good story when it's all over, when you emerge on the other side. Let me ask you about decisions. You already talked about it a little bit, but when we face the world and we're making different decisions, how much does our memory come into play? Is it the kind of narratives that we've constructed about the world that are used to make predictions that's fundamentally part of the decision-making?

**Charan Ranganath**

Absolutely. Yeah. So let's say after this, you and I decided we're going to go for a beer. How do you choose where to go? You're probably going to be like, "Oh yeah, this new bar opened up near me. I had a great time there. They had a great beer selection." Or you might say, "Oh, we went to this place and it was totally crowded and they were playing this horrible EDM or whatever." And so right there, valuable source of information. And then you have these things like where you do this counterfactual stuff, "Well, I did this previously." But what if I had gone somewhere else and said, "Maybe I'll go to this other place because I didn't try it the previous time"? So there's all that kind of reasoning that goes into it too. I think even if you think about the big decisions in life. It's like you and I were talking before we started recording about how I got into memory research and you got into AI and it's like we all have these personal reasons that guide us in these particular directions. And some of it's the environment and random factors in life, and some of it is memories of things that we want to overcome or things that we build on in a positive way. But either way, they define us.

**Lex Fridman**

And probably the earlier in life the memories happen, the more defining, the more defining power they have in terms of determining who we become.

**Charan Ranganath**

I mean, I do feel like adolescence is much more important than I think people give credit for. I think that there is this kind of a sense the first three years of life is the most important part, but the teenage years are just so important for the brain. And so that's where a lot of mental illness starts to emerge. Now we're thinking of things like schizophrenia as a neurodevelopmental disorder because it just emerges during that period of adolescence and early adulthood. And I think the other part of it is that I guess I was a little bit too firm in saying that memory determines who we are. It's really the self is an evolving construct. I

think we kind of underestimate that. And when you're a parent, you feel like every decision you make is consequential in forming this child and it plays a role, but so do the child's peers. And so do - there's so much, I mean that's why I think the big part of education I think that's so important is not the content you learn - I mean, think of how much dumb stuff we learned in school. But a lot of it is learning how to get along with people and learning who you are and how you function. And that can be terribly traumatizing even if you have perfect parents working on you.

### **Lex Fridman**

Is there some insight into the human brain that explains why we don't seem to remember anything from the first few years of life?

### **Charan Ranganath**

Yeah. Yeah. In fact, actually I was just talking to my really good friend and colleague, Simona Getty, who studies the neuroscience of child development and so we were talking about this. And so there are a bunch of reasons I would say. So one reason is there's an area of the brain called the hippocampus, which is very, very important for remembering events or episodic memory. And so the first two years of life, there's a period called infantile amnesia. And then the next couple years of life after that, there's a period called childhood amnesia. And the difference is that basically in the lab and even during childhood and afterwards, children basically don't have any episodic memories for those first two years. The next two years it's very fragmentary and that's why they call it childhood amnesia, so there's some, but it's not long. So one reason is that the hippocampus is taking some time to develop, but another is the neocortex of the whole folded stuff of gray matter all around the hippocampus is developing so rapidly and changing. And a child's knowledge of the world is just massively being built up, so I'm going to probably embarrass myself, but it's like if you showed you trained a neural network and you give it the first couple of patterns or something like that, and then you bombard it with another years worth of data, try to get back those first couple of patterns. It's like everything changes. And so the brain is so plastic, the cortex is so plastic during that time, and we think that memories for events are very distributed across the brain. Imagine you're trying to get back that pattern of activity that happened during this one moment, but the roads that you would take to get there have been completely rerouted. I think that's my best explanation. The third explanation is a child's sense of self takes a while to develop. And so their experience of learning might be more learning what happened as opposed to having this first-person experience of, "I remember. I was there."

### **Lex Fridman**

Well, I think somebody once said to me that kind of loosely philosophically that the reason we don't remember the first few years of life, infantile amnesia is because how traumatic it is. Basically the error rate that you mentioned when your brain's prediction doesn't match reality, the error rate in the first few years of life, your first few months certainly, is probably

crazy high. It's non-stop freaking out. The collision between your model of the world and how the world works is just so high that you want whatever the trauma of that is not to linger around. I always thought that's an interesting idea because just imagine the insanity of what's happening in a human brain in the first couple of years. You don't know anything and there's just this stream of knowledge and we're somehow, given how plastic everything is, it just kind of molds and figures it out. But it's like an insane waterfall of information.

### **Charan Ranganath**

I wouldn't necessarily describe it as a trauma and we can get into this whole stages of life thing, which I just love. Basically those first few years there are, I mean think about it, a kid's internal model of their body is changing. It's just learning to move. I mean, if you ever have a baby, you'll know that the first three months they're discovering their toes. It's just nuts. So everything is changing. But what's really fascinating is, and I think this is one of those, this is not at all me being a scientist, but it's one of those things that people talk about when they talk about the positive aspects of children is that they're exceptionally curious and they have this kind of openness towards the world. And so that prediction error is not a negative traumatic thing. I think it's a very positive thing because it's what they use, they're seeking information. One of the areas that I'm very interested in is the prefrontal cortex. It's an area of the brain that, I mean, I could talk all day about it, but it helps us use our knowledge to say, "Hey, this is what I want to do now. This is my goal, so this is how I'm going to achieve it," and focus everything towards that goal. The prefrontal cortex takes forever to develop in humans. The connections are still being tweaked and reformed into late adolescence, early adulthood, which is when you tend to see mental illness pop up. So it's being massively reformed. Then you have about 10 years maybe of prime functioning of the prefrontal cortex, and then it starts going down again and you end up being older and you start losing all that frontal function. So I look at this and you'd say, "Okay," you sit around episodic memory talks. While they always say children are worse than adults at episodic memory, older adults or worse than young adults at episodic memory. And I always would say, "God, this is so weird. Why would we have this period of time that's so short when we're perfect or optimal?" And I like to use that word optimal now because there's such a culture of optimization right now. And it's like I realize I have to redefine what optimal is because for most of the human condition, I think we had a series of stages of life where you have basically adults saying, "Okay", young adults saying, "I've got a child and I'm part of this village and I have to hunt and forage and get things done." I need a prefrontal cortex so I can stay focused on the big picture and the long haul goals. Now I'm a child, I'm in this village, I'm kind of wandering around and I've got some safety, and I need to learn about this culture because I know so little. What's the best way to do that? Let's explore. I don't want to be constrained by goals as much. I want to really be free, play and explore and learn. So you don't want a super tight prefrontal cortex. You don't even know what the goals should be yet. If you're trying to design a model that's based on a bad goal, it's not going to work well. So then you go late in life and you say, "Oh, why don't you have a great prefrontal cortex then?" But I think, I mean if you go back and you think how many species actually stick



around naturally long after their childbearing years are over, after the reproductive years are over? With menopause, from what I understand, menopause is not all that common in the animal world. So why would that happen? And so I saw Alison Gopnik said something about this so I started to look into this, about this idea that really when you're older in most societies, your job is no longer to form new episodic memories, it's to pass on the memories that you already have, this knowledge about the world, what we call semantic memory, to pass on that semantic memory to the younger generations, pass on the culture. Even now in indigenous cultures, that's the role of the elders. They're respected, they're not seen as people who are past it and losing it. And I thought that was a very poignant thing, that memory is doing what it's supposed to throughout these stages of life.

**Lex Fridman**

So it is always optimal in a sense.

**Charan Ranganath**

Yeah.

**Lex Fridman**

It's just optimal for that stage of life

**Charan Ranganath**

Yeah. And for the ecology of the system. So I looked into this and it's like another species that has menopause is orcas. Orca pods are led by the grandmothers. So it's not the young adults, not the parents or whatever, the grandmothers. And so they're the ones that pass on the traditions to I guess the younger generation of orcas. And if you look from what little I understand, different orca pods have different traditions. They hunt for different things. They have different play traditions, and that's a culture. And so in social animals, evolution I think is designing brains that are really around, it's obviously optimized for the individual but also for kin. And I think that the kin are part of this when they're a part of this intense social group, the brain development should parallel that, the nature of the ecology.

**Lex Fridman**

Well, it's just fascinating to think of the individual orca or human throughout its life in stages doing a kind of optimal wisdom development. So in the early days, you don't even know what the goal is, and you figure out the goal and you optimize for that goal and you pursue that goal. And then all the wisdom you collect through that, then you share with the others in the system, the other individuals. And as a collective, then you kind of converge towards greater wisdom throughout the generations. So in that sense, it's optimal. Us humans and orcas got something going on. It works.

**Charan Ranganath**

Well, yeah. Apex predators.

**Lex Fridman**

I just got a megalodon tooth, speaking of apex predators.

**Charan Ranganath**

Oh, man.

**Lex Fridman**

Just imagine the size of that thing. Anyway, how does the brain forget and how and why does it remember? So maybe some of the mechanisms. You mentioned the hippocampus, what are the different components involved here?

**Charan Ranganath**

So we could think about this on a number of levels. Maybe I'll give you the simplest version first, which is we tend to think of memories as these individual things and we can just access them, maybe a little bit like photos on your phone or something like that. But in the brain, the way it works is you have this distributed pool of neurons and the memories are kind of shared across different pools of neurons. And so what you have is competition, where sometimes memories that overlap can be fighting against each other. So sometimes we forget because that competition just wipes things out. Sometimes we forget because there aren't the biological signals which we can get into, I would promote long-term retention. And lots of times we forget because we can't find the cue that sends us back to the right memory, and we need the right cue to be able to activate it. So for instance, in a neural network there is no - you wouldn't go and you'd say, "This is the memory." It's like the whole network, I mean, the whole ecosystem of memories is in the weights of the neural network. And in fact, you could extract entirely new memories depending on how you feed.

**Lex Fridman**

You have to have the right query, the right prompt to access that whatever the part you're looking for.

**Charan Ranganath**

That's exactly right. That's exactly right. And in humans, you have this more complex set of ways memory works. There's, as I said, the knowledge or what you call semantic memory, and then there's these memories for specific events, which we call episodic memory. And so there's different pieces of the puzzle that require different kinds of cues. So that's a big part of it too, is just this kind of what we call retrieval failure.

**Lex Fridman**

You mentioned episodic memory, you mentioned semantic memory, what are the different separations here? What's working memory, short-term memory, long-term memory, what are the interesting categories of memory?

**Charan Ranganath**

Yeah. And so memory researchers, we love to cut things up and say, "is memory one thing or is it two things? There's two things or there's three things?" And so, one of the things that, and there's value in that, and especially experimental value in terms of being able to dissect things. In the real world, it's all connected. Speak to your question, working memory is a term that was coined by Alan Battley. It's basically thought to be this ability to keep information online in your mind right in front of you at a given time, and to be able to control the flow of that information, to choose what information is relevant, to be able to manipulate it and so forth. And one of the things that Alan did that was quite brilliant was he said, "There's this ability to kind of passively store information, see things in your mind's eye or hear your internal monologue," but we have that ability to keep information in mind. But then we also have this separate what he called a central executive, which is identified a lot with the prefrontal cortex. It's this ability to control the flow of information that's being kept active based on what it is you're doing. Now, a lot of my early work was basically saying that this working memory, which some memory researchers would call short-term memory is not at all independent from long-term memory. That is that a lot of executive function requires learning, and you have to have synaptic change for that to happen. But there's also transient forms of memory. So one of the things I've been getting into lately is the idea that we form internal models of events. The obvious one that I always use is birthday parties. So you go to a child's birthday party, once the cake comes out and you just see a candle, you can predict the whole frame set of events that happens later. And up until that point where the child blows out the candle, you have an internal model in your head of what's going on. And so if you follow people's eyes, it's not actually on what's happening, it's going where the action's about to happen, which is just fascinating. So you have this internal model, and that's a kind of a working memory product, it's something that you're keeping online that's allowing you to interpret this world around you. Now, to build that model though, you need to pull out stuff from your general knowledge of the world, which is what we call semantic memory. And then you'd want to be able to pull out memories for specific events that happened in the past, which we call episodic memory. So in a way, they're all connected, even though it's different. The things that we're focusing on and the way we organize information in the present, which is working memory, will play a big role in determining how we remember that information later, which people typically call long-term memory.

**Lex Fridman**

So if you have something like a birthday party and you've been to many before, you're going to load that from disk into working memory, this model, and then you're mostly operating on the model. And if it's a new task, you don't have a model so you're more in the data collection?

**Charan Ranganath**

Yes. One of the fascinating things that we've been studying, and we're not at all the first to do this, Jeff Sachs was a big pioneer in this, and I've been working with many other people,

Ken Norman, Leyla, Devachi and Wade. Columbia has done some interesting stuff with this, is this idea that we form these internal models at particular points of high prediction error or points of, I believe also points of uncertainty, points of surprise or motivationally significant periods. And those points are when it's maximally optimal to encode an episodic memory. So I used to think, "Oh, well, we're just encoding episodic memories constantly. Boom, boom, boom, boom, boom." But think about how much redundancy there is in all that. It's just a lot of information that you don't need. But if you capture an episodic memory at the point of maximum uncertainty, for the singular experience, it's only going to happen once, but if you capture it at the point of maximum uncertainty or maximum surprise, you have the most useful point in your experience that you've grabbed. And what we see is that the hippocampus and these other networks that are involved in generating these internal models of events, they show a heightened period of connectivity or correlated activity during those breaks between different events, which we call event boundaries. These are the points where you looked surprised or you cross from one room to another and so forth. And that communication is associated with a bump of activity in the hippocampus and better memory. And so if people have a very good internal model, throughout that event you don't need to do much memory processing, you're in a predictive mode. And so then at these event boundaries you encode, and then you retrieve and you're like, "Okay, wait a minute. What's going on here? Ranganath is now talking about orcas, what's going on?" And maybe you have to go back and remember reading my book to pull out the episodic memory to make sense of whatever it is I'm babbling about. And so there's this beautiful dynamics that you can see in the brain of these different networks that are coming together and then deaffiliating at different points in time that are allowing you to go into these modes. And so to speak to your original question, to some extent, when we're talking about semantic memory and episodic memory and working memory, you can think about it as these processes that are unfolding as these networks come together and pull apart,

### **Lex Fridman**

Can memory be trained and improved? This beautiful connected system that you've described, what aspect of it is a - you've described. What aspect of it is a mechanism that can be improved through training?

### **Charan Ranganath**

I think improvement, it depends on what your definition of optimal is. What I say in the book is that you don't want to remember more, you want to remember better, which means focusing on the things that are important. That's what our brains are designed to do. If you go back to the earliest quantitative studies of memory by Ebbinghaus, what you see is that he was trying so hard to memorize this arbitrary nonsense, and within a day, he lost about 60% of that information. He was basically using a very, very generous way of measuring it. As far as we know, nobody has managed to violate those basics of having people forget most of their experiences. If your expectation is that you should remember everything and that's what your optimal is, you're already off because this is just not what human brains are

designed to do. On the other hand, what we see over and over again is that, basically, one of the cool things about the design of the brain is it's always less is more. Less is more. I've seen estimates that the human brain uses something like 12 to 20 watts in a day. That's just nuts, the low power consumption. It's all about reusing information and making the most of what we already have. That's why basically, again, what you see biologically is neuromodulators, for instance, these chemicals in the brain like norepinephrine, dopamine, serotonin. These are chemicals that are released during moments that tend to be biologically significant, surprise, fear, stress, et cetera. These chemicals promote lasting plasticity, essentially, some mechanisms by which the brain can, say, prioritize the information that you carry with you into the future. Attention is a big factor as well, our ability to focus our attention on what's important, and so there's different schools of thought on training attention, for instance. One of my colleagues, Amishi Jha, she wrote a book called Peak Mind and talks about mindfulness as a method for improving attention and focus. She works a lot with military like Navy SEALs and stuff to do this kind of work with mindfulness meditation. Adam Gazzaley, another one of my friends and colleagues, has worked on training through video games actually as a way of training attention. So it's not clear to me, one of the challenges, though, in training is you tend to overfit to the thing that you're trying to optimize. If I'm looking at a video game, I can definitely get better at paying attention in the context of the video game, but you transfer it to the outside world, that's very controversial.

### **Lex Fridman**

The implication there is that attention is a fundamental component of remembering something, allocating attention to it, and then attention might be something that you could train, how you allocate attention and how you hold attention on a thing.

### **Charan Ranganath**

I can say that, in fact, we do in certain ways. If you are an expert in something, you are training attention. We did this one study of expertise in the brain. People used to think, let's say, if you're a bird expert or something, people will go, "If you get really into this world of birds, you start to see the differences and your visual cortex is tuned up, and it's all about plasticity of the visual cortex." Vision researchers love to say everything is visual, but it's like we did this study of working memory and expertise. One of the things that surprised us were the biggest effects as people became experts in identifying these different kinds of just crazy objects that we made up, as they developed this expertise of being able to identify what made them different from each other and what made them unique, we were actually seeing massive increases in activity in the prefrontal cortex. This fits with some of the studies of chess experts and so forth that it's not so much that you learn the patterns passively. You learn what to look for. You learn what's important and what's not. You can see this in any kind of expert professional athlete. They're looking three steps ahead of where they're supposed to be, so that's a kind of a training of attention. Those are also what you'd call expert memory skills. If you take the memory athletes, I know that's something we're

both interested in, so these are people who train in these competitions and they'll memorize a deck of cards in a really short amount of time. There's a great memory athlete, her name I think is pronounced Yǎnjāa Wintersoul. I think she's got a giant Instagram following. She had this YouTube video that went viral where she had memorized an entire Ikea catalog. How do people do this? By all accounts, from people who become memory athletes, they weren't born with some extraordinary memory, but they practice strategies over and over and over again. The strategy that they use for memorizing a particular thing, it can become automatic, and you can just deploy it in an instant. Again, one strategy for learning the order of a deck of cards might not help you for something else that you need like remembering your way around Austin, Texas. But it's going to be these, whatever you're interested in, you can optimize for that. That's just a natural byproduct of expertise.

**Lex Fridman**

There's a certain hacks. There's something called the Memory Palace that I played with. I don't know if you're familiar with that -

**Charan Ranganath**

Yeah. Yeah.

**Lex Fridman**

- whole technique, and it works. It's interesting. So another thing I recommend for people a lot is I use Anki a lot every day. It's an app that does spaced repetition. Medical students use this a lot to remember a lot of different things.

**Charan Ranganath**

Yeah. Yeah. Oh, yeah. Okay. We can come back to this, but yeah, go ahead.

**Lex Fridman**

Sure. It's the whole concept of spaced repetition. When the thing is fresh, you have to remind yourself of it a lot and then, over time, you can wait a week, a month, a year before you have to recall the thing again. That way, you essentially have something like note cards that you can have tens of thousands of and can only spend 30 minutes a day and actually be refreshing all of that information, all of that knowledge. It's really great. For Memory Palace, it's a technique that allows you to remember things like the Ikea catalog by placing them visually in a place that you're really familiar with, like, "I'm really familiar with this place," so I can put numbers or facts or whatever you want to remember you can walk along that little palace and it reminds you. It's cool. There's stuff like that that I think memory athletes could use, but I think also regular people can use. One of those things that I have to solve for myself is how to remember names. I'm horrible at it. I think it's because when people introduce themselves, I have the social anxiety of the interaction where I'm like, "I know I should be remembering that," but I'm freaking out internally about social interaction in general, and so therefore, I forget immediately, so I'm looking for good tricks for that.

**Charan Ranganath**

I feel like we've got a lot in common because when people introduce themselves to me, it's almost like I have this just blank blackout for a moment, and then I'm just looking at them like, "What happened?" I look away or something. What's wrong with me? I'm totally with you on this. The reason why it's hard is that there's no reason we should be able to remember names, because when you say you're remembering a name, you're not really remembering a name. Maybe in my case, you are, but, most of the time, you're associating a name with a face and an identity, and that's a completely arbitrary thing. Maybe in the olden days, somebody named Miller, it's like they're actually making flour or something like that. For the most part, it's like these names are just utterly arbitrary, so you have no thing to latch on to. It's not really a thing that our brain does very well to learn meaningless, arbitrary stuff. So what you need to do is build connections somehow, visualize a connection, and sometimes it's obvious or sometimes it's not. I'm trying to think of a good one for you now, but the first thing I think of is Lex Luthor -

**Lex Fridman**

That's great.

**Charan Ranganath**

- that I can think of. Yeah, so I think with Lex Luthor -

**Lex Fridman**

Doesn't Lex Luthor wear a suit, I think?

**Charan Ranganath**

I know he has a shaved head, though, or he's bald, which you're not. I'd trade hair with you any day -

**Lex Fridman**

Right.

**Charan Ranganath**

- but for something like that. If I can come up with something, I could say, "Okay, so Lex Luthor is this criminal mastermind," then I'd just imagine you -

**Lex Fridman**

We talked about stabbing or whatever earlier about [inaudible 00:41:07] -

**Charan Ranganath**

Yeah. Yeah. Exactly. Right?

**Lex Fridman**

- all just connected and that's it.

**Charan Ranganath**

Yeah. Yeah, but I'm serious though that these kinds of weird association is now I'm building a richer network. One of the things that I find is you can have somebody's name that's just totally generic like John Smith or something, no offense to people with that name, if I see a generic name like that, but I've read John Smith's papers academically and then I meet John Smith at a conference, I can immediately associate that name with that face 'cause I have this pre-existing network to lock everything in to. You can build that network, and that's what the method of loci or the Memory Palace technique is all about is you have a pre-existing structure in your head of your childhood home or this mental palace that you've created for yourself. So now you can put arbitrary pieces of information in different locations in that mental structure of yours and then you can walk through the different path and find all the pieces of information you're looking for. The method of loci is a great method for just learning arbitrary things because it allows you to link them together and get that cue that you need to pop in and find everything.

**Lex Fridman**

We should maybe linger on this Memory Palace thing just to make it obvious, 'cause when people were describing to me a while ago what this is, it seems insane. You literally think of a place like a childhood home or a home that you're really visually familiar with and you literally place in that three-dimensional space facts or people or whatever you want to remember, and you just walk in your mind along that place visually and you can remember, remind yourself of the different things. One of the limitations is there is a sequence to it. You can't just go upstairs right away or something. You have to walk along the room. It's really great for remembering sequences, but it's also not great for remembering individual facts out of context. The full context of the tour, I think, is important, but it's fascinating how the mind is able to do that. When you ground these pieces of knowledge into something that you remember well already, especially visually, it's fascinating. I think you do that for any kind of sequence. I'm sure she used something like this for the Ikea catalog, something of this nature.

**Charan Ranganath**

Oh, yeah, absolutely. Absolutely. I think the principle here is, again, I was telling you this idea that memories can compete with each other. Well, I like to use this example, and maybe someday I'll regret this, but I've used it a lot recently. Imagine if this were my desk, it could be cluttered with a zillion different things. Imagine it's just cluttered with a whole bunch of yellow Post-it notes and on one of them I put my bank password on it. Well, it's going to take me forever to find it. It's just going to be buried under all these other Post-it notes. If it's hot pink, it's going to stand out and I find it really easily. That's one way in which if things are distinctive, if you've processed information in a very distinctive way, then you can have a



memory that's going to last. That's very good, for instance, for name/face associations. If I get something distinctive about you that's it like you've got a very short hair, and maybe I can make the association with Lex Luthor that way or something like that. If I get something very specific, that's a great cue. But the other part of it is what if I just organized my notes so that I have my finances in one pile and I have my reminders, my to-do list in one pile and so forth so I organize them. Well, then, I know exactly if I'm going for my bank password, I could go to the finance pile. The method of loci works or Memory Palaces work because they give you a way of organizing. There's a school of thought that says that episodic memory evolved from this knowledge of space and basically there's primitive abilities to figure out where you are, and so people explain the method of loci that way. Whether or not the evolutionary argument is true, the method of loci is not at all special. If you're not a good visualizer, stories are a good one. So a lot of memory athletes will use stories and they'll go, like if you're memorizing a deck of cards, they have a little code for the different, the King and the Jack and the 10 and so forth. They'll make up a story about things that they're doing and that'll work. Songs are a great one. I can still remember there's this obscure episode of the TV show Cheers. They song about Albania that he uses to memorize all these facts about Albania. I could still sing that song to you as just as I saw it on the TV show.

### **Lex Fridman**

So you mentioned space repetition. So do you like this process? Maybe can you explain it?

### **Charan Ranganath**

Oh, yeah. If I am trying to memorize something, let's say if I have an hour to memorize as many Spanish words as I can, if I just try to do half-an-hour and then later in the day I do half-an-hour, I won't retain that information as long as if I do half-an-hour today and half-an-hour one week from now. So doing that extra spacing should help me retain the information better. Now, there's an interesting boundary condition, which is, it depends on when you need that information. So many of us, for me, I can't remember so much from college and high school 'cause I crammed 'cause I just did everything at the last minute. Sometimes I would literally study in the hallway right before the test, and that was great because what would happen is I just had that information right there. So actually, not spacing can really help you if you need it very quickly, but the problem is is that you tend to forget it later on. But on the other hand, if you space things out, you get a benefit for later on retention. So there's many different explanations. We have a computational model of this. It's currently under revision. But in our computer model, what we say is that maybe a good way of thinking about this is this conversation that you and I are having, it's associated with a particular context, a particular place in time. So all of these little cues that are in the background, these little guitar sculptures that you have and that big light umbrella thing, all these things are part of my memory for what we're talking about, the content. So now later on, you're sitting around, and you're at home drinking a beer and you're thinking, "God, what a strange interview that was," right? So now you're trying to remember it, but the context is

different. So your current situation doesn't match up with the memory that you pulled up, there's error. There's a mismatch between what you've pulled up and your current context. So in our model, what you start to do is you start to erase or alter the parts of the memory that are associated with a specific place and time, and you heighten the information about the content. So if you remember this information in different times in different places, it's more accessible at different times in different places because it's not overfitted in an AI way of thinking about things. It's not overfitted to one particular context. But that's also why the memories that we call upon the most also feel like they're just things that we read about almost. You don't vividly reimagine them, right? It's like they're just these things that just come to us, like facts, right?

**Lex Fridman**

Yeah.

**Charan Ranganath**

It's a little bit different than semantic memory, but it's like basically these events that we have recalled over and over and over again, we keep updating that memory so it's less and less tied to the original experience. But then we have those other ones, which it's like you just get a reminder of that very specific context. You smell something, you hear a song, you see a place that you haven't been to in a while, and boom, it just comes back to you. That's the exact opposite of what you get with spacing, right?

**Lex Fridman**

That's so fascinating. So with space repetition, one of its powers is that you lose attachment to a particular context, but then it loses the intensity of the flavor of the memory.

**Charan Ranganath**

Mmmhmm.

**Lex Fridman**

That's interesting. That's so interesting.

**Charan Ranganath**

Yeah, but at the same time, it becomes stronger in the sense that the content becomes stronger.

**Lex Fridman**

So it's used for learning languages, for learning facts, for that generic semantic information type of memories.

**Charan Ranganath**

Yeah, and I think this falls into a category. We've done other modeling. One of these is a published study in PLOS Computational Biology where we showed that another way, which is, I think, related to the spacing effect is what's called the testing effect. So the idea is that if you're trying to learn words, let's say in Spanish or something like that, and this doesn't have to be words, it could be anything, you test yourself on the words. That act of testing yourself helps you retain it better over time than if you just studied it. So from traditional learning theories, some learning theories, anyway, this seems weird, why would you do better giving yourself this extra error from testing yourself rather than just giving yourself perfect input that's a replica of what it is that you're trying to learn? I think the reason is is that you get better retention from that error, that mismatch that we talked about. So what's happening in our model, it's actually conceptually similar to what happens with backprop in AI or neural networks. So the idea is that you expose, "Here's the bad connections, and here's the good connections." So we can keep the parts of the cell assembly that are good for the memory and lose the ones that are not so good. But if you don't stress test the memory, you haven't exposed it to the error fully. So that's why I think this is a thing that I come back to over and over again, is that you will retain information better if you're constantly pushing yourself to your limit. If you are feeling like you're coasting, then you're actually not learning, so it's like -

**Lex Fridman**

You should always be stress testing the memory system.

**Charan Ranganath**

Yeah, and feel good about it. Even though everyone tells me, "Oh, my memory is terrible," in the moment they're overconfident about what they'll retain later on. So it's fascinating. So what happens is when you test yourself, you're like, "Oh, my God, I thought I knew that, but I don't." So it can be demoralizing until you get around that and you realize, "Hey, this is the way that I learn. This is how I learned best." It's like if you're trying to star in a movie or something like that, you don't just sit around reading the script. You actually act it out, and you're going to botch those lines from time to time, right?

**Lex Fridman**

You know what? There's an interesting moment, you probably have experienced this. I remember a good friend of mine, Joe Rogan, I was on his podcast, and we were randomly talking about soccer, football, somebody I grew up watching Diego Armando Maradona, one of the greatest soccer players of all time. We were talking about him and his career and so on, and Joe asked me if he's still around. I said, "Yeah." I don't know why I thought, "Yeah," because that was a perfect example of memories. He passed away. I tweeted about it, how heartbroken I was, all this kind of stuff a year before. I know this, but in my mind, I went back to the thing I've done many times in my head of visualizing some of the epic runs he had on goal and so on. So for me, he's alive. Part of also the conversation when you're talking to

Joe, there's stress and the focus is allocated. The attention is allocated in a particular way. But when I walked away, I was like, "In which world was Diego Maradona still alive?" 'Cause I was sure in my head that he was still alive. It's a moment that sticks with me. I've had a few like that in my life where it just - obvious things just disappear from mind, and it's cool. It shows actually the power of the mind in the positive sense to erase memories you want erased maybe, but I don't know. I don't know if there's a good explanation for that.

### **Charan Ranganath**

One of the cool things that I found is that some people really just revolutionize a field by creating a problem that didn't exist before. It's why I love science is engineering is like solving other people's problems and science is about creating problems. I'm just much more like I want to break things and create problems, not necessarily move fast, though. But one of my former mentors, Marcia Johnson, who in my opinion is one of the greatest memory researchers of all time, she comes up young woman in the field in this mostly guy field. She gets into this idea of how do we tell the difference between things that we've imagined and things that we actually remember? How do we tell, I get some mental experience, where did that mental experience come from? It turns out this is a huge problem because essentially our mental experience of remembering something that happened, our mental experience of thinking about something, how do you tell the difference? They're both largely constructions in our head, and so it is very important. The way that you do it is, it's not perfect, but the way that we often do it and succeed is by, again, using our prefrontal cortex and really focusing on the sensory information or the place in time and the things that put us back into when this information happened. If it's something you thought about, you're not going to have all of that vivid detail as you do for something that actually happened, but it doesn't work all the time. But that's a big thing that you have to do. But it takes time. It's slow, and it's again, effortful, but that's what you need to remember accurately. But what's cool, and I think this is what you alluded to about how that was an interesting experience is, imagination is exactly the opposite. Imagination is basically saying, "I'm just going to take all this information from memory, recombine it in different ways and throw it out there." So for instance, Dan Schachter and Donna Addis have done cool work on this. Demis Hassabis did work on this with Eleanor McGuire in UCL, and this goes back actually to this guy, Frederic Bartlett, who is this revolutionary memory researcher, Bartlett. He actually rejected the whole idea of quantifying memory. He said, "There's no statistics in my book." He came from this anthropology perspective and short version of the story is he just asked people to recall things. You give people stories and poems, ask people to recall them. What we found was people's memories didn't reflect all of the details of what they were exposed to, and they did reflect a lot more - they were filtered through this lens of prior knowledge; the cultures that they came from, the beliefs that they had, the things they knew. So what he concluded was that he called remembering an imaginative construction, meaning that we don't replay the past, we imagine how the past could have been by taking bits and pieces that come up in our heads. Likewise, he wrote this beautiful paper on imagination saying when we imagine something and create something, we're creating it from these specific experiences that

we've had and combining it with our general knowledge. But instead of trying to focus it on being accurate and getting out one thing, you're just ruthlessly recombining things without any necessary goal in mind, or at least that's one kind of creation.

### **Lex Fridman**

So imagination is fundamentally coupled with memory in both directions.

### **Charan Ranganath**

I think so. It's not clear that it is in everyone, but one of the things that's been studied is some patients who have amnesia, for instance, they have brain damage, say, to the hippocampus. If you ask them to imagine things that are not in front of them, imagine what could happen after I leave this room, they find it very difficult to give you a scenario what could happen. Or if they do, it would be more stereotyped like, "Yes, this would happen, this would..." But it's not like they can come up with anything that's very vivid and creative in that sense. It's partly 'cause when you have amnesia, you're stuck in the present because to get a very good model of the future, it really helps to have episodic memories to draw upon, and so that's the basic idea. In fact, one of the most impressive things when people started to scan people's brains and ask people to remember past events, what they found was there was this big network of the brain called the default mode network. It gets a lot of press because it's thought to be important. It's engaged during mind wandering. If I ask you to pay attention to something, it only comes on when you stop paying attention, so people, "Oh, it's just this kind of daydreaming network." I thought, "This is just ridiculous research. Who cares?" But then what people found was when people recall episodic memories, this network gets active. So we started to look into it, and this network of areas is really closely functionally interacting with the hippocampus. So in fact, some would say the hippocampus is part of this default network. If you look at brain images of people or brain maps of activation, so to speak, of people imagining possible scenarios of things that could happen in the future or even things that couldn't really be very plausible, they look very similar. To the naked eye, they look almost the same as maps of brain activation when people remember the past. According to our theory, and we've got some data to support this, we've broken up this network in various sub pieces, is that basically it's taking apart all of our experiences and creating these little Lego blocks out of them. Then you can put them back together if you have the right instructions to recreate these experiences that you've had, but you could also reassemble them into new pieces to create a model of an event that hasn't happened yet, and that's what we think happens when our common ground that we're establishing in language requires using those building blocks to put together a model of what's going on.

### **Lex Fridman**

Well, there's a good percentage of time I personally live in the imagined world. I do thought experiments a lot. I take the absurdity of human life as it stands and play it forward in all kinds of different directions. Sometimes it's rigorous thoughts, thought experiments,

sometimes it's fun ones. So I imagine that that has an effect on how I remember things. I suppose I have to be a little bit careful to make sure stuff happened versus stuff that I just imagined happened. Some of my best friends are characters inside books that never even existed. There's some degree to which they actually exist in my mind. Like these characters exist, authors exist, Dostoevsky exists, but also Brothers Karamazov.

**Charan Ranganath**

I love that book. One of the few books I've read. One of the few literature books that I've read, I should say. I read a lot in school that I don't remember, but Brothers Karamazov, I remember. Alyosha -

**Lex Fridman**

They exist, and I have almost conversations with them, it's interesting. It's interesting to allow your brain to play with ideas of the past of the imagined and see it all as one.

**Charan Ranganath**

Yeah, there was actually this famous mnemonist, he's like back then the equivalent of a memory athlete, except he would go to shows and do this, that was described by this really famous neuropsychologist from Russia named Luria. So this guy was named Solomon Shereshevsky, and he had this condition called synesthesia that basically created these weird associations between different senses that normally wouldn't go together. So that gave him this incredibly vivid imagination that he would use to basically imagine all sorts of things that he would need to memorize, and he would just imagine, just create these incredibly detailed things in his head that allowed him to memorize all sorts of stuff. But it also really haunted him by some reports that basically it was like he was at some point, and again, who knows if the drinking was part of this, but he at some point had trouble differentiating his imagination from reality. This is interesting because it's like that's what psychosis is in some ways is first of all, you're just learning connections from prediction errors that you probably shouldn't learn. The other part of it is that your internal signals are being confused with actual things in the outside world. Right?

**Lex Fridman**

Well, that's why a lot of this stuff is both feature and bug. It's a double-edged sword.

**Charan Ranganath**

Yeah, it might be why there's such an interesting relationship between genius and psychosis.

**Lex Fridman**

Yeah. Maybe they're just two sides of the same coin. Humans are fascinating, aren't they?

**Charan Ranganath**

I think so, sometimes scary, but mostly fascinating.

**Lex Fridman**

Can we just talk about memory sport a little longer? There's something called the USA Memory Championship. What are these athletes like? What does it mean to be elite level at this? Have you interacted with any of them or reading about them, what have you learned about these folks?

**Charan Ranganath**

There's a guy named Henry Roediger who's studying these guys. There's actually a book by Joshua Foer called *Moonwalking with Einstein*, where he talks about, he actually, as part of this book, just decided to become a memory athlete. They often have these life events that make them go - athlete, they often have these life events that make them go, "Hey, why don't I do this?" So there was a guy named Scott Hagwood who I write about, who thought that he was getting chemo for cancer. And so he decided, because chemo, there's a well-known thing called chemo brain where people become, they just lose a lot of their sharpness. And so he wanted to fight that by learning these memory skills. So he bought a book, and this is the story you hear in a lot of memory athletes is they buy a book by other memory athletes or other memory experts, so to speak. And they just learn those skills and practice them over and over again. They start by winning bets and so forth. And then they go into these competitions. And the competitions are typically things like memorizing long strings of numbers or memorizing orders of cards and so forth. So they tend to be pretty arbitrary things, not things that you'd be able to bring a lot of prior knowledge. But they build the skills that you need to memorize arbitrary things.

**Lex Fridman**

Yeah, that's fascinating. I've gotten a chance to work with something called n-back tasks. So there's all these kinds of tasks, memory recall tasks that are used to kind of load up the quote-unquote, working memory.

**Charan Ranganath**

Yeah, yeah.

**Lex Fridman**

The psychologist used it to test all kinds of stuff to see how well you're good at multitasking. We used it in particular for the task of driving. If we fill up your brain with intensive working memory tasks, how good are you at also not crashing, that kind of stuff. So it's fascinating, but again, those tasks are arbitrary and they're usually about recalling a sequence of numbers in some kind of semi-complex way. Do you have any favorite tasks of this nature in your own studies?

**Charan Ranganath**

I've really been most excited about going in the opposite direction and using things that are more and more naturalistic. And the reason is that we've moved in that direction because what we found is that memory works very, very differently when you study memory in the way that people typically remember. And so it goes into a much more predictive mode. And you have these event boundaries, for instance, and you have - but a lot of what happens is this kind of fascinating mix that we've been talking about, a mix of interpretations and imagination with perception. And the new direction we're going in is understanding navigation in our memory [inaudible 01:06:44] places. And the reason is that there's a lot of work that's done in rats, which is very good work. They have a rat and they put it in a box and the rat goes chases cheese in a box. You'll find cells in the hippocampus that fire when a rat is in different places in the box. And so the conventional wisdom is that the hippocampus forms this map of the box. And I think that probably may happen when you have absolutely no knowledge of the world, right? But I think one of the cool things about human memory is we can bring to bear our past experiences to economically learn new ones. And so for instance, if you learn a map of an IKEA, let's say if I go to the IKEA in Austin, I'm sure there's one here. I probably could go to this IKEA and find my way to where the wine glasses are without having to even think about it because it's got a very similar layout, even though IKEA is a nightmare to get around. Once I learned my local IKEA, I can use that map everywhere. Why form a brand new one for a new place? So that kind of ability to reuse information really comes into play when we look at things that are more naturalistic tasks. And another thing that we're really interested in is this idea of what if instead of basically mapping out every coordinate in a space, you form a pretty economical graph that connects basically the major landmarks together? And being able to use that as emphasizing the things that are most important, the places that you go for food and the places that are landmarks that help you get around. And then filling in the blanks for the rest, because I really believe that cognitive maps or mental maps of the world, just like our memories for events are not photographic. I think they're this combination of actual verifiable details and then a lot of inference that you make.

**Lex Fridman**

What have you learned about this kind of spatial mapping of places? How do people represent locations?

**Charan Ranganath**

There's a lot of variability, I think that - and there's a lot of disagreement about how people represent locations. In a world of GPS and physical maps, people can learn it from basically what they call a survey perspective, being able to see everything. And so that's one way in which humans can do it that's a little bit different. There's one way which we can memorize routes. I know how to get from here to, let's say if I walk here from my hotel, I could just rigidly follow that route back, right? And there's another more integrative way, which would be what's called a cognitive map. Which would be kind of a sense of how everything relates



to each other. And so there's lots of people who believe that these maps that we have in our head are isomorphic with the world, that are these literal coordinates that follow Euclidean space. And as you know, Euclidean mathematics is very constrained, right? And I think that we are actually much more generative in our maps of space so that we do have these bits and pieces. And we've got a small task, it's right now, not yet - we need to do some work on it for further analyses. But one of the things we're looking at is these signals called ripples in the hippocampus, which are these bursts of activity that you see that are synchronized with areas in the neocortex, in the default network actually. And so what we find is that those ripples seem to increase at navigationally important points when you're making a decision or when you reach a goal. This speaks to the emotion thing, right? Because if you have limited choices, if I'm walking down a street, I could really just get a mental map of the neighborhood with a more minimal kind of thing by just saying, "Here's the intersections and here's the directions I take to get in between them." And what we found in general in our MRI studies is basically the more people can reduce the problem, whether it's space or any kind of decision-making problem, the less the hippocampus encodes. It really is very economical towards the points of most highest information, content and value.

### **Lex Fridman**

So can you describe the encoding in the hippocampus and the ripples you were talking about? What's the signal in which we see the ripples?

### **Charan Ranganath**

Yeah, so this is really interesting. There are these oscillations, right? So there's these waves that you basically see. And these waves are points of very high excitability and low excitability. And at least during - they happen actually during slow-wave sleep too. So the deepest stages of sleep, when you're just zonked out, right? You see these very slow waves, where it's very excitable and then very unexcitable, it goes up and down. And on top of them you'll see these little sharp wave ripples. And when there's a ripple in the hippocampus, you tend to see a sequence of cells that resemble a sequence of cells that fire when an animal is actually doing something in the world. So it almost is like a little, people call it replay, I think it's a little bit - I don't like that term, but it's basically a little bit of a compressed play of the sequence of activity in the brain that was taking place earlier. And during those moments, there's a little window of communication between the hippocampus and these areas in the neocortex. And so that I think helps you form new memories, but it also helps you, I think, stabilize them, but also really connect different things together in memory. And allows you to build bridges between different events that you've had. And so this is one of at least our theories of sleep, and its real role in helping you see the connections between different events that you've experienced.

### **Lex Fridman**

So during sleep is when the connections are formed?

**Charan Ranganath**

The connections between different events, right?

**Lex Fridman**

Yeah.

**Charan Ranganath**

So it's like you see me now, you see me next week, you see me a month later. You start to build a little internal model of how I behave and what to expect of me. And we think sleep, one of the things it allows you to do is figure out those connections and connect the dots and find the signal in the noise.

**Lex Fridman**

So you mentioned fMRI. What is it? And how is it used in studying memory?

**Charan Ranganath**

This is actually the reason why I got into this whole field of science is when I was in grad school, fMRI was just really taking off as a technique for studying brain activity. And what's beautiful about it is you can study the whole human brain. And there's lots of limits to it, but you can basically do it in a person without sticking anything into their brains, and very non-invasive. For me being in an MRI scanner is like being in the womb, I just fall asleep. If I'm not being asked to do anything, I get very sleepy. But you can have people watch movies while they're being scanned or you can have them do tests of memory, giving them words and so forth to memorize. But what MRI is itself is just this technique where you put people in a very high magnetic field. Typical ones we would use would be 3 Tesla to give you an idea. So a 3 Tesla magnet, you put somebody in, and what happens is you get this very weak but measurable magnetization in the brain. And then you apply a radio frequency pulse, which is basically a different electromagnetic field. And so you're basically using water, the water molecules in the brain as a tracer, so to speak. And part of it in fMRI is the fact that these magnetic fields that you mess with by manipulating these radio frequency pulses and the static field, and you have things called gradients, which change the strength of the magnetic field in different parts of the head. So we tweak them in different ways, but the basic idea that we use in fMRI is that blood is flowing to the brain. And when you have blood that doesn't have oxygen on it, it's a little bit more magnetizable than blood that does because you have hemoglobin that carries the oxygen, the iron basically in the blood that makes it red. And so that hemoglobin, when it's deoxygenated actually has different magnetic field properties than when it has oxygen. And it turns out when you have an increase in local activity in some part of the brain, the blood flows there. And as a result you get a lower concentration of hemoglobin that is not oxygenated, and then that gives you more signal. So I gave you, I think I sent you a GIF, as you like to say.

**Lex Fridman**

Yeah, we had off-record intense argument about if it's pronounced GIF or GIFF, but we shall set that aside as friends.

**Charan Ranganath**

We could have called it a stern rebuke perhaps, but -

**Lex Fridman**

Rebuke, yeah. I drew a hard line, it is true the creator of GIF said it's pronounced GIFF, but that's the only person that pronounces GIFF. Anyway, yes, you sent a GIF of -

**Charan Ranganath**

This would be basically a whole - a movie of fMRI data. And so when you look at it, it's not very impressive, it looks like these very pixelated maps of the brain, but it's mostly kind of white. But these tiny changes in the intensity of those signals that you probably wouldn't be able to visually perceive, about 1% can be statistically very, very large effects for us. And that allows us to see, "Hey, there's an increase in activity in some part of the brain when I'm doing some task like trying to remember something." And I can use those changes to even predict, is a person going to remember this later or not? And the coolest thing that people have done is to decode what people are remembering from the patterns of activity from - because maybe when I'm remembering this thing, I'm remembering the house where I grew up. I might have one pixel that's bright in the hippocampus and one that's dark. And if I'm remembering something more like the car that I used to drive when I was 16, I might see the opposite pattern where a different pixel is bright. And so all that little stuff that we use to think of noise, we can now think of almost like a QR code for memory, so to speak. Where different memories have a different little pattern of bright pixels and dark pixels. And so this really revolutionized my research. So there's fancy research out there where people really - not even that - by your standards, this would be Stone Age, but applying machine learning techniques to do decoding and so forth. And now there's a lot of forward encoding models and you can go to town with this stuff, right? And I'm much more old school of designing experiments where you basically say, "Okay, here's a whole web of memories that overlap in some way, shape or form." Do memories that occurred in the same place have a similar QR code? And do memories that occurred in different places have a different QR code? And you can just use things like correlation coefficients or cosine distance to measure that stuff, right? Super simple, right? And so what happens is you can start to get a whole state space of how a brain area is indexing all these different memories. It's super fascinating because what we could see is this little separation between how certain brain areas are processing memory for who was there. And other brain areas are processing information about where it occurred, or the situation that's kind of unfolding. And some are giving you information about what are my goals that are involved and so forth. And the hippocampus is just putting it all together into these unique things that just are of about when and where it happened.

**Lex Fridman**

So there's a separation between spatial information concepts, literally there's distinct, as you said, QR codes for these?

**Charan Ranganath**

So to speak. Let me try a different analogy too, that might be more accessible for people. Which would be, you've got a folder on your computer, right? I open it up, there's a bunch of files there. I can sort those files by alphabetical order. And now things that both start with letter A are lumped together, and things that start with Z versus A are far apart, right?

**Lex Fridman**

Mmhmm.

**Charan Ranganath**

And so that is one way of organizing the folder, but I could do it by date. And if I do it by date, things that were created close together in time are close, and things that are far apart in time are far. So you can think of how a brain area or a network of areas contributes to memory by looking at what the sorting scheme is. And these QR codes that we're talking about that you get from fMRI allow you to do that. And you can do the same thing if you're recording from massive populations of neurons in an animal. And you can do it for recording local potentials in the brain. So little waves of activity in let's say a human who has epilepsy and they stick electrodes in their brain to try to find seizures. So that's some of the work that we're doing now. But all of these techniques basically allow you to say, "Hey, what's the sorting scheme?" And so we've found that some networks of the brain sort information in memory according to who was there. So I might have - we've actually shown in one of my favorite studies of all time that was done by a former postdoc, Zach Reagh. And Zach did the study where we had a bunch of movies with different people in my labs that are two different people. And you filmed them at two different cafes and two different supermarkets. And what you could show is in one particular network, you could find the same kind of pattern of activity, more or less, a very similar pattern of activity. Every time I saw Alex in one of these movies, no matter where he was, right? And I could see another one that was a common pattern that happened every time I saw this particular supermarket nugget. And it didn't matter whether you're watching a movie or whether you're recalling the movie, it's the same kind of pattern that comes up, right?

**Lex Fridman**

It's so fascinating.

**Charan Ranganath**

It is fascinating. And so now you have those building blocks for assembling a model of what's happening in the present, imagining what could happen, and remembering things

very economically from putting together all these pieces. So that all the hippocampus has to do is get the right kind of blueprint for how to put together all these building blocks.

**Lex Fridman**

These are all beautiful hints at a super interesting system that makes me wonder on the other side of it how to build it. But it's fascinating the way it does the encoding is really, really fascinating. Or I guess the symptoms, the results of that encoding are fascinating to study from this. Just as a small tangent, you mentioned sort of the measuring local potentials with electrodes versus fMRI.

**Charan Ranganath**

Oh yeah.

**Lex Fridman**

What are some interesting limitations, possibilities of fMRI? The way you explained it is brilliant with blood and detecting the activations or the excitation because blood flows to that area. What's the latency of that? What's the blood dynamics in the brain that - how quickly can the tasks change and all that kind of stuff?

**Charan Ranganath**

Yeah, it's very slow. To the brain, 50 milliseconds, it's an eternity. Maybe not 50 mil - maybe let's say half a second, 500 milliseconds, just so much back and forth stuff happens in the brain in that time, right? So in fMRI, you can measure these magnetic field responses about six seconds after that burst of activity would take place. All these things, it's like is it a feature or is it a bug? Right? So one of the interesting things that's been discovered about fMRI is it's not so tightly related to the spiking of the neurons. So we tend to think of the computation, so to speak, as being driven by spikes, meaning there's just a burst of it's either on or it's off and the neurons going up or down. But sometimes what you can have is these states where the neuron becomes a little bit more excitable or less excitable. And so fMRI is very sensitive to those changes in excitability. Actually, one of the fascinating things about fMRI is where does that - how is it we go from neural activity to essentially blood flow to oxygen? All this stuff. It's such a long chain of going from neural activity to magnetic fields. And one of the theories that's out there is most of the cells in the brain are not neurons, they're actually these support cells called glial cells. And one big one is astrocytes, and they play this big role in regulating, kind of being a middle man, so to speak, with the neurons. So if, for instance, one neuron's talking to another, you release a neurotransmitter like let's say glutamate. And that gets another neuron, starts getting active after you release it in the gap between the two neurons called the synapse. So what's interesting is if you leave that, imagine you're just flooded with this liquid in there, right? If you leave it in there too long, you just excite the other neuron too much and you can start to basically get seizure activity. You don't want this, so you got to suck it up. And so actually what happens is these astrocytes, one of their functions is to suck up the glutamate from the synapse.

And that is a massively - and then break it down and then feed it back into the neuron so that you could reuse it. But that cycling is actually very energy intensive. And what's interesting is at least according to one theory, they need to work so quickly that they're working on metabolizing the glucose that comes in without using oxygen. Kind of like anaerobic metabolism, so they're not using oxygen as fast as they're using glucose. So what we're really seeing in some ways may be in fMRI, not the neurons themselves being active, but rather the astrocytes which are meeting the metabolic demands of the process of keeping the whole system going.

### **Lex Fridman**

It does seem to be that fMRI is a good way to study activation. So with these astrocytes, even though there's a latency, it's pretty reliably coupled to the activations.

### **Charan Ranganath**

Oh, well, this gets me to the other part. So now let's say for instance, if I'm just kind of I'm talking to you, but I'm kind of paying attention to your cowboy hat, right? So I'm looking off to the - or I'm thinking about the [inaudible 01:26:12], even if I'm not looking at it. What you'd see is that there'd be this little elevation in activity in areas in the visual cortex, which process vision around that point in space, okay? So if then something happened like a suddenly a light flashed in that part of - right in front of your cowboy hat, I would have a bigger response to it. But what you see in fMRI is even if I don't see that flash of light, there's a lot of activity that I can measure because you're kind of keeping it excitable [inaudible 01:26:46] that in and of itself, even though I'm not seeing anything there that's particularly interesting, there's still this increase in activity. So it's more sensitive with fMRI. So is that a feature or is it a bug? People who study spikes in neurons would say, "Well, that's terrible, we don't want that." Likewise, it's slow, and that's terrible for measuring things that are very fast. But one of the things that we found in our work was when we give people movies and when we give people stories to listen to, a lot of the action is in the very, very slow stuff. Because if you're thinking about a story, let's say you're listening to a podcast or something, you're listening to Lex Fridman Podcast, right? You're putting this stuff together and building this internal model over several seconds. Which is basically we filter that out when we look at electrical activity in the brain because we're interested in this millisecond scale, it's almost massive amounts of information, right? So the way I see it is every technique gives you a little limited window into what's going on. fMRI has huge problems, people lie down in the scanner. There's parts of the brain where - I will show you in some of these images where you'll see kind of gaping holes because you can't keep the magnetic field stable in those spots. You'll see parts where it's like there's a vein, and so it just produces big increase and decrease in signal or respiration that causes these changes. There's lots of artifacts and stuff like that. Every technique has its limits. If I'm lying down in an MRI scanner, I'm lying down. I'm not interacting with you in the same way that I would in the real world. But at the same time, I'm getting data that I might not be able to get otherwise. And so different techniques give you different kinds of advantages.

**Lex Fridman**

What kind of big scientific discoveries, maybe the flavor of discoveries have been done throughout the history of the science of memory, the studying of memory? What kind of things have been understood?

**Charan Ranganath**

Oh, there's so many, it's really so hard to summarize it. I think it's funny because it's like when you're in the field, you can get kind of blasé about this stuff. But then once I started write the book, I was like, "Oh my God, this is really interesting. How did we do all this stuff?" I would say that some of the - from the first study, it's just showing how much we forget is very important. Showing how much schemas, which is our organized knowledge about the world increase our ability to remember information, just massively increase in [inaudible 01:29:25] of expertise. Showing how experts like chess experts can memorize so much in such a short amount of time because of the schemas they have for chess. But then also showing that those lead to all sorts of distortions in memory.

**Lex Fridman**

Mmhmm.

**Charan Ranganath**

The discovery that the act of remembering can change the memory, it can strengthen it, but it can also distort it if you get misinformation at the time. And it can also strengthen or weaken other memories that you didn't even recall. So just this whole idea of memory as an ecosystem I think was a big discovery. I could go, this idea of breaking up our continuous experience into these discrete events, I think was a major discovery.

**Lex Fridman**

So the discreteness of our encoding of events?

**Charan Ranganath**

Maybe, yeah, and again, there's controversial ideas about this, right? But it's like, yeah, this idea that - and this gets back to just this common experience of you walk into the kitchen and you're like, "Why am I here?" And you just end up grabbing some food from the fridge. And you go back and you're like, "Oh, wait a minute, I left my watch in the kitchen. That's what I was looking for." And so what happens is that you have a little internal model of where you are, what you're thinking about. And when you cross from one room to another, those models get updated. And so now when you're in the kitchen, have to go back and mentally time travel back to this earlier point to remember what it was that you went there for. And so these event boundaries turns out in our research, and again, I don't want to make it sound like we've figured out everything. But in our research, one of the things that we found is that basically, as people get older, the activity in the hippocampus at these event boundaries tends to go down, but independent of age. If I give you outside of the scanner,

you're done with the scanner, I just scan you while you're watching a movie, just watch it. You come out, I give you a test of memory for stories. What happens is you find this incredible correlation between the activity in the hippocampus at these singular points in time, these event boundaries. And your ability to just remember a story outside of the scanner later on. So it's marking this ability to encode memories, just these little snippets of neural activity. So I think that's a big one. There's all sorts of work in animal models that I can get into. Sleep, I think there's so much interesting stuff that's being discovered in sleep right now. Being able to just record from large populations of cells and then be able to relate that - [inaudible 01:32:03], I think the coolest thing gets back to this QR code thing, because what we can do now is I can take fMRI data while you're watching a movie. Let's do better than that. Let me get fMRI data while you use a joystick to move around in virtual reality. So you're in the metaverse, whatever. But it's kind of a crappy metaverse because there's only so much metaversing you can do in an MRI scanner. So you're doing this crappy metaversing. So now, I can take a rat, record from its hippocampus and prefrontal cortex and all these areas with these really new electrodes that get massive amounts of data. And have it move around on a trackball in virtual reality in the same metaverse that I did, and record that rat's activity. I can get a person with epilepsy who we have electrodes in their brain anyway, to try to figure out where the seizures are coming from. And if it's a healthy part of the brain, record from that person, right? And I can get a computational model. And one of the brand new members in my lab, Tyler Brown is just doing some great stuff. He relates computer vision models and looks at the weaknesses of computer vision models and relates to what the brain does well.

**Lex Fridman**

Mmhmm. Nice.

**Charan Ranganath**

And so you can actually take a ground truth code for the metaverse, basically, and you can feed in the visual information, let's say the sensory information or whatever that's coming in to a computational model that's designed to take real world inputs, right? And you could basically tie them all together by virtue of the state spaces that you're measuring in neural activity, in these different formats and these different species, and in the computational model. Which is I just find that mind-blowing. And you could do different kinds of analyses on language and basically come up with - basically it's the guts of LLMs, right? You could do analyses on language and you could do analyses on sentiment analyses of emotions and so forth. Put all this stuff together, it's almost too much. But if you do it right and you do it in a theory-driven way as opposed to just throwing all the data at the wall and see what sticks, that to me is just exceptionally powerful.



**Lex Fridman**

So you can take fMRI data across species and across different types of humans or conditions of humans, and construct models that help you find the commonalities or the core thing that makes somebody navigate through the metaverse, for example?

**Charan Ranganath**

Yeah. Yeah, more or less. There's a lot of details, but yes, I think - and not just fMRI, but you can relate it to, like I said, recordings from large populations of neurons that could be taken in a human or even in a non-human animal, that is where you think it's an anatomical homologue. So that's just mind-blowing to me.

**Lex Fridman**

What's the similarities in humans and mice? That's what Smashing Pumpkins, we're all just rats in a cage. Is that Smashing Pumpkins?

**Charan Ranganath**

Despite all of your rage.

**Lex Fridman**

Is that Smashing Pumpkins? I think [inaudible 01:35:17].

**Charan Ranganath**

Despite all of your rage at GIFs, you're still just a rat in a cage.

**Lex Fridman**

Oh yeah. All right, good callback. Anyway -

**Charan Ranganath**

Good callback, see these memory retrieval exercises I'm doing are actually helping you build a lasting memory of this conversation.

**Lex Fridman**

And it's strengthening the visual thing I have of you with James Brown on stage just become stronger and stronger by the second. Anyway -

**Charan Ranganath**

[inaudible 01:35:43].

**Lex Fridman**

But animal studies work here as well.

**Charan Ranganath**

Yeah, yeah. Okay. So I think I've got great colleagues who I talk to who study memory in mice. And one of the valuable things in those models is you can study neural circuits in an enormously targeted way, because you can study neural circuits in an enormously targeted way because you could do these genetic studies, for instance, where you can manipulate particular groups of neurons, and it's just getting more and more targeted to the point where you can actually turn on a particular kind of memory, just by activating a particular set of neurons that was active during an experience. So, there's a lot of conservation of some of these neural circuits across evolution in mammals, for instance. And then some people would even say that there's genetic mechanisms for learning that are conserved, even going back far, far before. But let's go back to the mice in humans question. There's a lot of differences. So, for one thing, the sensory information is very different. Mice and rats explore the world largely through smelling, olfaction, but they also have vision that's kind of designed to catch death from above. So, it's like a very big view of the world. And we move our eyes around in a way that focuses on particular spots in space where you get very high resolution from a very limited set of spots in space. So, that makes us very different in that way. We also have all these other structures as social animals that allow us to respond differently. There's language, there's - you name it - there's obviously gobs of differences. Humans aren't just giant rats. There's much more complexity to us. Timescales are very important. So, primate brains and human brains are especially good at integrating and holding on to information across longer and longer periods of time. Also, finally, it's like our history of training data, so to speak, is very, very different than - human's world is very different than a wild mouse's world. And a lab mouse's world is extraordinarily impoverished relative to an adult human. Yeah.

**Lex Fridman**

But still, what can you understand by studying mice? I mean, just basic, almost behavioral stuff about memory?

**Charan Ranganath**

Well, yes, but that's very important. So, you can understand, for instance, how do neurons talk to each other? That's a really big, big question. Neural computation, in and of itself - you think it's the most simple question, right? Not at all. I mean, it's a big, big question, and understanding how two parts of the brain interact, meaning that it's not just one area, speaking it's not like Twitter where one area of the brain's shouting and then another area of the brain's just stuck listening to this crap. It's like they're actually interacting on the millisecond scale. How does that happen and how do you regulate those interactions, these dynamic interactions? We're still figuring that out. But that's going to be coming largely from model systems that are easier to understand. You can do manipulations, like drug manipulations, to manipulate circuits, and use viruses and so forth, and lasers to turn on circuits that you just can't do in humans. So, I think there's a lot that can be learned from mice. There's a lot that can be learned from non-human primates. And then there's a lot that

you need to learn from humans. And I think unfortunately, some of the people in the National Institutes of Health think you can learn everything from the mouse. It's like, "Why study memory in humans when I could study learning in a mouse?" And just like, "Oh my God, I'm going to get my funding from somewhere else."

**Lex Fridman**

Well, let me ask you some random fascinating question.

**Charan Ranganath**

Yeah, sure.

**Lex Fridman**

How does déjà vu work?

**Charan Ranganath**

So, déjà vu, it's actually one of these things I think that some of the surveys suggest that 75% of people report having a déjà vu experience one time or another. I don't know where that came from, but I've polled people in my class and most of them say they've experienced déjà vu. It's this kind of sense that I've experienced this moment sometime before, I've been here before. And actually there's all sorts of variants of this. The French have all sorts of names for various versions of this, [foreign language 01:40:12]. I don't know. It's like all these different vus. But déjà vu is the sense that it can be almost disturbing intense sense of familiarity. So, there was a researcher named Wilder Penfield - actually, this goes back even earlier to some of the earliest, like Hughlings Jackson was this neurologist who did a lot of the early characterizations of epilepsy. And one of the things he notices in epilepsy patients, some group of them right before they would get a seizure, they would have this intense sense of déjà vu. So, it's this artificial sense of familiarity, it's a sense of having a memory that's not there. What was happening was there was electrical activity in certain parts of these brains, so the guy Penfield, later on when he was trying to look for how do we map out the brain to figure out which parts we want to remove and which parts don't we, he would stimulate parts of the temporal lobes of the brain and find you could elicit the sense of déjà vu. Sometimes you'd actually get a memory that a person would re-experience just from electrically stimulating some parts. Sometimes they just have this intense feeling of being somewhere before. And so, one theory which I really like is that in higher order areas of the brain, they're integrating from many, many different sources of input. What happens is that they're tuning themselves up every time you process a similar input. And so that allows you to just get this kind of affluent sense that, "I'm very familiar..." You're very familiar with this place. And so just being here, you're not going to be moving your eyes all over the place because you kind of have an idea of where everything is. And that fluency gives you a sense of, "I'm here." Now, I wake up in my hotel room and I have this very unfamiliar sense of where I am. But there's a great set of studies done by Anne Cleary at Colorado State where she created these virtual reality environments. And we'll go back to the metaverse. Imagine

you go through a virtual museum, and then she would put people in virtual reality and have them go through a virtual arcade. But the map of the two places was exactly the same. She just put different skins on them. So, one looks different than the other, but they've got same landmarks, and the same places, same objects, same everything, but carpeting, colors, theme, everything's different. People will often not have any conscious idea that the two are the same, but they could report this very intense sense of déjà vu. So, it's like a partial match that's eliciting this kind of a sense of familiarity. And that's why in patients who have epilepsy, that affects memory, you get this artificial sense of familiarity that happens. And so we think that... and again, this is just one theory amongst many, but we think that we get a little bit of that feeling, it's not enough to necessarily give you déjà vu, even for very mundane things. So, it's like if I tell you the word rutabaga, your brain's going to work a little bit harder to catch it than if I give you word like apple. That's because you hear apple a lot. So, your brain's very tuned up to process it efficiently, but rutabaga takes a little bit longer and more intense. And you can actually see a difference in brain activity in areas in the temporal lobe when you hear a word just based on how frequent it is in the English language.

**Lex Fridman**

That's fascinating.

**Charan Ranganath**

We think it's tied to this basic - it's basically a by-product of our mechanism of just learning, doing this error-driven learning as we go through life to become better and better and better to process things more and more efficiently.

**Lex Fridman**

So, I guess déjà vu is just thinking extra elevated, the stuff coming together, firing for this artificial memories, as if it's the real memory. I mean, why does it feel so intense?

**Charan Ranganath**

Well, it doesn't happen all the time, but I think what may be happening is it's a partial match to something that we have, and it's not enough to trigger that sense of - that ability to pull together all the pieces. But it's a close enough match to give you that intense sense of familiarity, without the recollection of exactly what happened when.

**Lex Fridman**

But it's also a spatio-temporal familiarity. So, it's also in time. There's a weird blending of time that happens, and we'll probably talk about time because I think that's a really interesting idea how time relates to memory. But you also kind of - artificial memory brings to mind this idea of false memories that comes in all kinds of contexts. But how do false memories form?

### **Charan Ranganath**

Well, I like to say there's no such thing as true or false memories. It's like Johnny Rotten from the Sex Pistols, he had a saying that's like, "I don't believe in false memories any more than I believe in false songs." And so the basic idea is that we have these memories that reflect bits and pieces of what happened, as well as our inferences and theories. So, I'm a scientist and I collect data, but I use theories to make sense of that data. And so, a memory is kind of a mix of all these things. Where memories can go off the deep end and become what we would call conventionally as false memories are sometimes little distortions where we filled in the blanks, the gaps in our memory, based on things that we know, but don't actually correspond to what happened. So, if I were to tell you that a story about this person who's worried that they have cancer or something like that, and then they see a doctor and the doctor says, "Well, things are very much like you would've expected or what you were afraid of," or something. When people remember that, they'll often remember, "Well, the doctor told the patient that he had cancer." Even if that wasn't in the story because they're infusing meaning into that story. So, that's a minor distortion. But what happens is that sometimes things can really get out of hand where people have trouble telling the difference between things that they've imagined versus things that happen. But also, as I told you, the act of remembering can change the memory. And so what happens then is you can actually be exposed to some misinformation. And so Elizabeth Loftus was a real pioneer in this work, and there's lots of other work that's been done since. But basically, it's like if you remember some event, and then I tell you something about the event, later on, when you remember the event, you might remember some original information from the event as well as some information about what I told you. And sometimes, if you're not able to tell the difference, that information that I told you gets mixed into the story that you had originally. So, now I give you some more misinformation or you're exposed to some more information somewhere else, and eventually your memory becomes totally detached from what happened. And so sometimes you can have cases where people - this is very rare, but you can do it in lab too, or a significant - not everybody, but a chunk of people will fall for this, where you can give people misinformation about an event that never took place. And as they keep trying to remember that event more and more, what happens is they start to imagine, they start to pull up things from other experiences they've had, and eventually they can stitch together a vivid memory of something that never happened because they're not remembering an event that happened. They're remembering the act of trying to remember what happened, and basically putting it together into the wrong story.

### **Lex Fridman**

It's fascinating because this could probably happen at a collective level. This is probably what successful propaganda machines aim to do, this creating false memory across thousands, if not millions of minds.

**Charan Ranganath**

Yeah, absolutely. I mean, this is exactly what they do. And so, all these kind of foibles of human memory get magnified when you start to have social interactions. There's a whole literature on something called social contagion, which is basically when misinformation spreads like a virus, like you remember the same thing that I did, but I give you a little bit of wrong information, then that becomes part of your story of what happened. Because once you and I share a memory, I tell you about something I've experienced and you tell me about your experience at the same event, it's no longer your memory or my memory, it's our memory. And so now the misinformation spreads. And the more you trust someone or the more powerful that person is, the more of a voice they have in shaping that narrative. And there's all sorts of interesting ways in which misinformation can happen. There's a great example of when John McCain and George Bush Jr. were in a primary, and there were these polls where they would do these, I guess they were not robocalls, but real calls where they would poll voters, but they actually inserted some misinformation about McCain's beliefs on taxation, I think, or maybe it was something about illegitimate children or - I don't really remember. But they included misinformation in the question that they asked, "How do you feel about the fact that he wants to do this?" Or something. And so people would end up becoming convinced he had these policy things or these personal things that were not true, just based on the polls that were being used. So, it was a case where, interestingly enough, the people who were using misinformation were actually ahead of the curve relative to the scientists who were trying to study these effects in memory.

**Lex Fridman**

Yeah, it's really interesting. So, it's not just about truth and falsehoods, like us as intelligent, reasoning machines, but it's the formation of memories where they become visceral. You can rewrite history. If you just look throughout the 20th century, some of the dictatorships with Nazi Germany, with the Soviet Union, effective propaganda machines can rewrite our conceptions of history, how we remember our own culture, our upbringing, all this kind of stuff. And you could do quite a lot of damage in this way. And then there's probably some kind of social contagion happening there. Certain ideas that, maybe initiated by the propaganda machine, can spread faster than others. You could see that in modern day, certain conspiracy theories, there's just something about them that they are really effective at spreading. There's something sexy about them to people to where something about the human mind eats it up and then uses that to construct memories as if they almost were there to witness whatever the content of the conspiracy theory is. It's fascinating. Because you feel like you remember a thing, I feel like there's a certainty. It emboldens you to say stuff. It's not just you believe in ideas, true or not, it's at the core of your being that you feel like you were there to watch the thing happen.

**Charan Ranganath**

Yeah, I mean there's so much in what you're saying. I mean, one of the things is that people's sense of collective identity is very much tied to shared memories. If we have a shared

narrative of the past, or even better, if we have a shared past, we will feel more socially connected with each other, and I will feel part of this group. They're part of my tribe, if I remember the same things in the same way. And you brought up this weaponization of history, and it really speaks to, I think, one of the parts of memory, which is that if you have a belief, you will find, and you have a goal in mind, you'll find stuff in memory that aligns with it, and you won't see the parts in memory that don't. So, a lot of the stories we put together are based on our perspectives. And so let's just zoom out for the moment from misinformation to take something even more fascinating, but not as scary. I was reading Thanh Viet Nguyen, but he wrote a book about the collective memory of the Vietnam War. He is a Vietnamese immigrant who was flown out after the war was over. And so he went back to his family to get their stories about the war, and they called it the American War, not the Vietnam War. And that just kind of blew my mind, having grown up in the US and having always heard about it as the Vietnam War. But of course they call it the American War, because that's what happened. America came in. And that's based on their perspective, which is a very valid perspective. And so that just gives you this idea of the way we put together these narratives based on our perspectives. And I think the opportunities that we can have in memory is if we bring groups together from different perspectives and we allow them to talk to each other and we allow ourselves to listen. I mean, right now you'll hear a lot of just jammering, people going, "Blah, blah, blah," about free speech, but they just want to listen to themselves. I mean, it's like, let's face it, the old days before people were supposedly woke, they were trying to ban 2 Live Crew. Just think about Lenny Bruce got canceled for cursing. Jesus Christ. It's like this is nothing new. People don't like to hear things that disagree with them. But if you're in a - I mean, you can see two situations in groups with memory. One situation is you have people who are very dominant, who just take over the conversation. And basically what happens is the group remembers less from the experience and they remember more of what the dominant narrator says. Now, if you have a diverse group of people, and I don't mean diverse in necessarily the human resource sense of the word, I mean diverse in any way you want to take it, but diverse in every way, hopefully. And you give everyone a chance to speak and everyone's being appreciated for their unique contribution, you get more accurate memories and you get more information from it. Even two people who come from very similar backgrounds, if you can appreciate the unique contributions that each one has, you can do a better job of generating information from memory. And that's a way to inoculate ourselves, I believe, from misinformation in the modern world. But like everything else, it requires a certain tolerance for discomfort. And I think when we don't have much time, and I think when we're stressed out and when we are just tired, it's very hard to tolerate discomfort.

### **Lex Fridman**

And I mean, social media has a lot of opportunity for this because it enables this distributed one-on-one interaction that you're talking about, where everybody has a voice, but still our natural inclination, you see this on social media, as there's a natural clustering of people and opinions and you just form these kind of bubbles. To me personally, I think that's a

technology problem that could be solved if there's a little bit of interaction, kind, respectful, compassionate interaction with people that have a very different memory, that respectful interaction will start to intermix the memories and ways of thinking to where you're slowly moving towards truth. But that's a technology problem because naturally, left our own devices, we want to cluster up in a tribe.

### **Charan Ranganath**

Yeah, and that's the human problem. I think a lot of the problems that come up with technology aren't the technology itself, as much as the fact that people adapt to the technology in maladaptive ways. I mean, one of my fears about AI is not what AI will do, but what people will do. I mean, take text messaging. It's a pain in the to text people, at least for me. And so what happens is the communication becomes very Spartan and devoid of meaning. It's this very telegraphic. And that's people adapting to the medium. I mean, look at you. You've got this keyboard that's got these dome shaped things, and you've adapted to that to communicate. That's not the technology adapting to you, that's you adapting to the technology. And I think one of the things I learned when Google started to introduce autocomplete in emails, I started to use it. And about a third of the time I was like, "This isn't what I want to say." A third of the time, I'd be like, "This is exactly what I wanted to say." And a third of the time I was saying, "Well, this is good enough. I'll just go with it." And so what happens is it's not that the technology necessarily is doing anything so bad, as much as it's just going to constrain my language because I'm just doing suggested to me. And so this is why I say, kind of like my mantra for some of what I've learned about everything in memory, is to diversify your training data, basically, because otherwise you're going to - so, humans have this capability to be so much more creative than anything generative AI will put together, at least right now, who knows where this goes? But it can also go the opposite direction where people could become much, much less creative, if they just become more and more resistant to discomfort and resistant to exposing themselves to novelty, to cognitive dissonance, and so forth.

### **Lex Fridman**

I think there is a dance between natural human adaptation of technology and the people that design the engineering of that technology. So, I think there's a lot of opportunity to create, like this keyboard, things that on net are a positive for human behavior. So, we adapt and all this kind of stuff. But when you look at the long arc of history across the years and decades, has humanity been flourishing? Are humans creating more awesome stuff, are humans happier? All that kind of stuff. And so there, I think technology, on net, has been, and I think, maybe hope, will always be, on net, a positive thing.

### **Charan Ranganath**

Do you think people are happier now than they were 50 years ago or 100 years ago?



**Lex Fridman**

Yes, yes.

**Charan Ranganath**

I don't know about that.

**Lex Fridman**

I think humans in general like to reminisce about the past, "The times were better."

**Charan Ranganath**

That's true.

**Lex Fridman**

And complain about the weather today or complain about whatever today, because there's this kind of complainy engine, there's so much pleasure in saying, "Life sucks," for some reason.

**Charan Ranganath**

That's why I love punk rock.

**Lex Fridman**

Exactly. I mean, there's something in humans that loves complaining, even about trivial things. But complaining about change, complaining about everything. But ultimately, I think, on net, every measure, things are getting better, life is getting better.

**Charan Ranganath**

Oh, life is getting better. But I don't know that necessarily that attracts people's happiness, right? I mean, I would argue that maybe, who knows, I don't know this, but I wouldn't be surprised if people in hunter-gatherer societies are happier. I mean, I wouldn't be surprised if they're happier than people who have access to modern medicine and email and cellphones.

**Lex Fridman**

Well, I don't think there's a question whether you take hunter-gatherer folks and put them into modern day and give them enough time to adapt, they would be much happier. The question is, in terms of every single problem they've had, is now solved. There's now food, there's guaranteed survival, and shelter and all this kind of stuff. So, what you're asking is a deeper sort of biological question, do we want to be - Werner Herzog and the movie Happy People: Life in the Taiga, do we want to be busy 100% of our time hunting, gathering, surviving, worried about the next day? Maybe that constant struggle ultimately creates a more fulfilling life. I don't know. But I do know this modern society allows us to, when we're sick, to find medicine, to find cures, when we're hungry, to get food, much more than we did

even a hundred years ago. And there's many more activities that you could perform, all creative, all these kinds of stuff that enables the flourishing of humans at the individual level. Whether that leads to happiness, I mean, that's a very deep philosophical question. Maybe struggle, deep struggle is necessary for happiness.

### **Charan Ranganath**

Or maybe cultural connection. Maybe it's about functioning in social groups that are meaningful, and having time. But I do think there's an interesting memory related thing, which is that if you look at things like reinforcement learning for instance, you're not learning necessarily every time you get a reward, if it's the same reward, you're not learning that much. You mainly learn if it deviates from your expectation of what you're supposed to get. So, it's like you get a paycheck every month from MIT or whatever, and you probably don't even get excited about it when you get the paycheck. But if they cut your salary, you're going to be pissed. And if they increase your salary, "Oh good, I got a bonus." And that adaptation and that ability that basically you learn to expect these things, I think, is a major source of - I guess it's a major way in which we're kind of more, in my opinion, wired to strive and not be happy, to be in a state of wanting. And so people talk about dopamine, for instance, being this pleasure chemical. And there's a lot of compelling research to suggest it's not about pleasure at all. It's about the discomfort that energizes you to get things, to seek a reward. And so you could give an animal that's been deprived of dopamine a reward and, "Oh yeah, I enjoy it. It's pretty good." But they're not going to do anything to get it. And just one of the weird things in our research is I got into curiosity from a postdoc in my lab, Matthias Gruber, and one of the things that we found is when we gave people a question, like a trivia question that they wanted the answer to, that question, the more curious people were about the answer, the more activity in these dopamine-related circuits in the brain, we would see. And again, that was not driven by the answer per se, but by the question. So, it was not about getting the information, it was about the drive to seek the information. But it depends on how you take that. If you get this uncomfortable gap between what you know and what you want to know, you could either use that to motivate you and energize you, or you could use it to say, "I don't want to hear about this. This disagrees with my beliefs. I'm going to go back to my echo chamber."

### **Lex Fridman**

Yeah, I like what you said that maybe we're designed to be in a kind of constant state of wanting, which by the way, is a pretty good either band name or rock song name, state of wanting.

### **Charan Ranganath**

That's like a hardcore band name. Yeah, yeah, yeah.

### **Lex Fridman**

Yeah. It's pretty good.

**Charan Ranganath**

But I also like the hedonic treadmill.

**Lex Fridman**

Hedonic treadmill is pretty good.

**Charan Ranganath**

Yeah, yeah. We could use that for our techno project, I think.

**Lex Fridman**

You mean the one we're starting?

**Charan Ranganath**

Yeah, exactly.

**Lex Fridman**

Okay, great. We're going on tour soon. This is our announcement.

**Charan Ranganath**

We could build a false memory of a show, in fact, if you want. Let's just put it all together so we don't even have to do all the work to play the show. We can just create a memory of it and it might as well happen because the remembering itself is in charge anyway.

**Lex Fridman**

So, let me ask you about - we talked about false memories, but in the legal system, false confessions. I remember reading 1984 where, sorry for the dark turn of our conversation, but through torture, you can make people say anything and essentially remember anything. I wonder to which degree, there's truth to that, if you look at the torture that happened in the Soviet Union, for confessions, all that kind of stuff. How much can you really get people to force false memories, I guess?

**Charan Ranganath**

Yeah. I mean, I think there's a lot of history of this actually, in the criminal justice system. You might've heard the term "the third degree." If you actually look it up historically, it was a very intense set of beatings and starvation and physical demands that they would place at people to get them to talk. And there's certainly a lot of work that's been done by the CIA in terms of enhanced interrogation techniques. And from what I understand, the research actually shows that they just produce what people want to hear, not necessarily the information that is being looked for. And the reason is that - I mean, there's different reasons. One is people just get tired of being tortured and just say whatever. But another part of it is that you create a very interesting set of conditions where there's an authority figure telling you something that, "You did this, we know you did this. We have witnesses

saying you did this." So, now you start to question yourself. Then they put you under stress. Maybe they're not feeding you, maybe they're making you be cold or exposing you to music that you can't stand or something, whatever it is, right? It's like they're creating this physical stress. And so stress starts to down-regulate the prefrontal cortex. You're not necessarily as good at monitoring the accuracy of stuff. Then they start to get nice to you and they say, "Imagine, okay, I know you don't remember this, but maybe we can walk you through how it could have happened." And they feed you the information. And so you're in this weakened mental state, and you're being encouraged to imagine things by people who give you a plausible scenario. And at some point, certain people can be very coaxed into creating a memory for something that never happened. And there's actually some pretty convincing cases out there where you don't know exactly the truth. There's a sheriff, for instance, who came to believe that he had a false memory - I mean, that he had a memory of doing sexual abuse based on essentially, I think it was - I'm not going to tell the story because I don't remember it well enough to necessarily accurately give it to you, but people could look this stuff up. There are definitely stories out there like this where people confess to crimes that they just didn't do, and - out there like this, where people confess to crimes that they just didn't do and objective evidence came out later on. There's a basic recipe for it, which is you feed people the information that you want them to remember, you stress them out. You have an authority figure pushing this information on them, or you motivate them to produce the information you're looking for. That pretty much over time gives you what you want.

### **Lex Fridman**

It's really tragic that centralized power can use these kinds of tools to destroy lives. Sad. Since there's a theme about music throughout this conversation, one of the best topics for songs is heartbreak. Love in general, but heartbreak. Why and how do we remember and forget heartbreak? Asking for a friend.

### **Charan Ranganath**

Oh, God, that's so hard to - asking for a friend. I love that. It's such a hard one. Part of this is we tend to go back to particular times that are the more emotionally intense periods, and so that's a part of it. Again, memory is designed to capture these things that are biologically significant, and attachment is a big part of biological significance for humans. Human relationships are super important and sometimes that heartbreak comes with massive changes in your beliefs about somebody say if they cheated on you or something like that, or regrets and you kind of ruminate about things that you've done wrong. There's really so many reasons though, but I've had this. My first pet I had, we got it for a wedding present. It was a cat. Got it after, but it died of FIP when it was four years old. I just would see her everywhere around the house. We got another cat, then we got a dog. Dog eventually died of cancer, and the cat just died recently. So we got a new dog because I kept seeing the dog around and I was just so heartbroken about this, but I still remember the pets that died. It just comes back to you. I mean, it's part of this. I think there's also something about attachment that's just so crucial that drives again, these things that we want to remember

and that gives us that longing sometimes. Sometimes it's also not just about the heartbreak, but about the positive aspects of it. The loss comes from not only the fact that the relationship is over, but you had all of these good things before that you can now see in a new light. Part of one of the things that I found from my clinical background that really I think gave me a different perspective on memory is so much of the therapy process was guided towards reframing and getting people to look at the past in a different way, not by imposing changing people's memories or not by imposing an interpretation, but just offering a different perspective and maybe one that's kind of more optimized towards learning and an appreciation maybe, or gratitude, whatever it is that gives you a way of taking - I think you said it in the beginning, right? Where you can have this kind of dark experiences and you can use it as training data to grow in new ways, but it's hard.

### **Lex Fridman**

I often go back to this moment, this show Louis with Louis CK, where he's all heartbroken about a breakup with a woman he loves, and an older gentleman tells him that that's actually the best part, that heartbreak, because you get to intensely experience how valuable this love was. He says the worst part is forgetting it. It is actually when you get over the heartbreak, that's the worst part. I sometimes think about that because having the love and losing it, the losing it is when you sometimes feel it the deepest, which is an interesting way to celebrate the past and relive it. It sucks that you don't have a thing, but when you don't have a thing, it's a good moment to viscerally experience the memories of something that you now appreciate even more.

### **Charan Ranganath**

So you don't believe that an owner of a lonely heart is much better than an owner of a broken heart? You think an owner of a broken heart is better than the owner of a lonely heart?

### **Lex Fridman**

Yes, for sure. I think so. I think so. I'm going to have to day by day. I don't know. I'm going to have to listen to some more Bruce Springsteen to figure that one out.

### **Charan Ranganath**

Well, it's funny because it's like after I turned 50, I think of death all the time. I just think that I have probably a fewer years ahead of me than I'm behind me. I think about one thing, which is what are the memories that I want to carry with me for the next period of time? And also, about just the fact that everything around me could be - I know more people who are dying for various reasons. I'm not Lot. I'm not that old, but it's something I think about a lot. I'm reminded of how I talked to somebody who's a Buddhist and I was like, "The whole of Buddhism is renouncing attachment." In some way, the idea of Buddhism is like staying out of the world of memory and staying in the moment. They talked about how do you renounce attachments to the people that you love? They're just saying, "Well, I appreciate that I have

this moment with them and knowing that they will die makes me appreciate this moment that much more.” You said something similar in your daily routine that you think about things this way, right?

**Lex Fridman**

Yeah, I meditate on mortality every day, but I don’t know, at the same time, that really makes you appreciate the moment and live in the moment. I also appreciate the full deep rollercoaster of suffering involved in life, the little and the big too. I don’t know. The Buddhist removing yourself from the world or the Stoic removing yourself from the world, the world of emotion, I’m torn about that one. I’m not sure.

**Charan Ranganath**

This is where Hinduism and Buddhism, or at least some strains of Hinduism and Buddhism, differ. Hinduism, if you read the Bhagavad Gita, the philosophy is not one of renouncing the world because the idea is that not doing something is no different than doing something. What they argue, and again, you could interpret in different ways, positive and negative, but the argument is that you don’t want to renounce action, but you want to renounce the fruits of the action. You don’t do it because of the outcome. You do it because of the process, because the process is part of the balance of the world that you’re trying to preserve. Of course you could take that different ways, but I really think about that from time to time in terms of letting go of this idea of does this book sell or trying to impress you and get you to laugh at my jokes or whatever, and just be more like I’m sharing this information with you and getting to know you or whatever it is. It’s hard, because we’re so driven by the reinforcer, the outcome.

**Lex Fridman**

You’re just part of the process of telling the joke, and if I laugh or not, that’s up to the universe to decide.

**Charan Ranganath**

Yep. It’s my dharma.

**Lex Fridman**

How does studying memory affect your understanding of the nature of time? We’ve been talking about us living in the present and making decisions about the future, standing on the foundation of these memories and narratives about the memories that we’ve constructed. It feels like it does weird things to time.

**Charan Ranganath**

Yeah, and the reason is that in some sense, I think especially the farther we go back, there’s all sorts of interesting things that happen. Your sense of if I ask how different does one hour ago feel from two hours ago? You’d probably say pretty different. But if I ask you, okay, go

back one year ago versus one year and one hour ago, it's the same difference in time. It won't feel very different. There's this kind of compression that happens as you look back farther in time. It is kind of like why when you're older, the difference between somebody who's 50 and 45 doesn't seem as big as the difference between 10 and five or something. When you're 10 years old, everything seems like it's a long period of time. Here's the point is that - one of the interesting things that I found when I was working on the book actually was during the pandemic, I just decided to ask people in my class when we were doing the remote instruction. One of the things I did was I would pull people. I just asked people, "Do you feel like the days are moving by slower or faster or about the same?" Almost everyone in the class said that the days were moving by slower. Then I would say, "Okay, so do you feel like the weeks are passing by slower, faster, or the same?" The majority of them said that the weeks were passing by faster. According to the laws of physics, I don't think that makes any sense, but according to memory, it did because what happened was people were doing the same thing over and over in the same context. Without that change in context, their feeling was that they were in one long monotonous event. Then at the end of the week, you look back at that week and you say, "Well, what happened? I have no memories of what happened," so the week just went by without even my noticing it. That week went by during the same amount of time as an eventful week where you might've been going out hanging out with friends on vacation or whatever. It's just that nothing happened because you're doing the same thing over and over. I feel like memory really shapes our sense of time, but it does so in part because context is so important for memory.

### **Lex Fridman**

That compression you mentioned, it's an interesting process because when I think about when I was 12 or 15, I just fundamentally feel like the same person. It's interesting what that compression does. It makes me feel like we're all connected, not just amongst humans and spatially, but in terms back in time. There's a kind of eternal nature, like the timelessness I guess, to life. That could be also a genetic thing just for me. I don't know if everyone agrees to this view of time, but to me it all feels the same.

### **Charan Ranganath**

You don't feel the passage of time?

### **Lex Fridman**

No, I feel the passage of time in the same way that your students did from day to day. There's certain markers that let you know that time has passed, you celebrate birthdays and so on, but the core of who I am and who others I know are, or events, that compression of my understanding of the world removes time because time is not useful for the compression. The details of that time, at least for me, is not useful to understanding the core of the thing.

### **Charan Ranganath**

Maybe what it is that you really like to see connections between things. This is really what motivates me in science actually too. It's like when you start recalling the past and seeing the connections between the past and present, now you have this web of interconnected memories. I can imagine in that sense there is this kind of the present is with you. What's interesting about what you said too that struck me is that your 16-year-old self was probably very complex. By the way, I'm the same way, but it's like it really is the source of a lot of darkness for me. When you can look back at, let's say you hear a song that you used to play before you would go do a sports thing or something like that, you might not think of yourself as an athlete, but once you mentally time travel to that particular thing, you open up this little compartment of yourself that wasn't there before that didn't seem accessible before. Dan Schacter's lab did this really cool study where they would ask people to either remember doing something altruistic or imagine doing something altruistic, and that act made them more likely to want to do things for other people. That act of mental time travel can change who you are in the present. We tend to think of, this goes back to that illusion of stability, and we tend to think of memory in this very deterministic way that I am who I am because I have this past, but we have a very multi-faceted past and can access different parts of it and change in the moment based on whatever part we want to reach for.

### **Lex Fridman**

How does nostalgia connect into this desire and pleasure associated with going back?

### **Charan Ranganath**

My friend Felipe de Brigard wrote this, and it just blew my mind, where the word nostalgia was coined by a Swiss physician who was actually studying traumatized soldiers. He described nostalgia as a disease. The idea was it was bringing these people extraordinary unhappiness because they're remembering how things used to be. I think it's very complex. As people get older, for instance, nostalgia can be an enormous source of happiness. Being nostalgic can improve people's moods in the moment, but it just depends on what they do with it because what you can sometimes see is nostalgia has the opposite effect of thinking those were the good old days, and those days are over. It's like America used to be so great, and now it sucks. My life used to be so great when I was a kid and now it's not. You're selectively remembering the things that - I mean, we don't realize how selective our remembering self is. I lived through the 70s. It sucked. Partly it sucked more for me, but I would say that even otherwise, it's like there's all sorts of problems going on, gas lines, people were worried about Russia, nuclear war, blah, blah, blah. It's just this idea that people have about the past can be very useful if it brings you happiness in the present, but if it narrows your worldview in the present, you're not aware of those biases that you have, it can be toxic either at a personal level or at a collective level.



**Lex Fridman**

Let me ask you both a practical question and an out there question. Let's start with a more practical one. What are your thoughts about BCIs, brain computer interfaces, and the work that's going on with Neuralink? We talked about electrodes and different ways of measuring the brain, and here Neuralink is working on basically two-way communication with the brain. The more out there question will be like, where's this go? More practically in the near term, what do you think about Neuralink?

**Charan Ranganath**

I can't say specifics about the company because I haven't studied it that much, but I think there's two parts of it. One is, they're developing some really interesting technology I think with these surgical robots and things like that. BCI though has a whole lot of innovation going on. I am not necessarily seeing any scientific evidence from Neuralink, and maybe that's just because I'm not looking for it, but I'm not seeing the evidence that they're anywhere near where the scientific community is. There's lots of startups that are doing incredibly innovative stuff. One of my colleagues, Sergey Stavisky is just a genius in this area, and they're working on it. I think speech prosthetics like they're incorporating, decoding techniques with AI and movement prosthetics. This is just the rate of progress is just enormous. Part of the technology is having good enough data and understanding which data to use and what to do with it. Then the other part of it then is the algorithms for decoding it and so forth. I think part of that has really resulted in some real breakthroughs in neuroscience as a result. There's lots of new technologies like Neuropixels for instance, that allow you to harvest activity from many, many neurons from a single electrode. I know Neuralink has some technologies that are also along these lines, but again, because they do their own stuff, the scientific community doesn't see it. I think BCI is much, much bigger than Neuralink and there's just so much innovation happening. I think the interesting question which we may be getting into is, I was talking to Sergey a while ago about a lot of language is not just what we hear and what we speak, but also our intentions and our internal models. And so, are you really going to be able to restore language without dealing with that part of it? He brought up a really interesting question, which is the ethics of reading out people's intentions and understanding of the world as opposed to the more concrete parts of hearing and producing movements.

**Lex Fridman**

Just so we're clear, because you said a few interesting things, when we talk about language and BCIs, what we mean is getting signal from the brain and generating the language, say you're not able to actually speak, it's as a kind of linguistic prosthetic. It's able to speak for you exactly what you want it to say. Then the deeper question is, well, saying something isn't just the letters, the words that you're saying, it's also the intention behind it, the feeling behind all that kind of stuff. Is it ethical to reveal that full shebang, the full context of what's going on in our brain? That's really interesting. That's really interesting. Our thoughts, is it ethical for anyone to have access to our thoughts? Because right now the resolution is so

low that we're okay with it, even doing studies and all this kind of stuff. If neuroscience has a few breakthroughs to where you can start to map out the QR codes for different thoughts, for different kinds of thoughts, maybe political thoughts, the McCarthyism, what if I'm getting a lot of them communist thoughts, or however we want to categorize or label it? That's interesting. That's really interesting. I think ultimately this always - the more transparency there is about the human mind, the better it is. There could be always intermediate battles with how much control does a centralized entity have, like a government and so on. What is the regulation? What are the rules? What's legal and illegal? If you talk about the police, whose job is to track down criminals and so on, and you look at all the history, how the police could abuse its power to control the citizenry, all that kind of stuff. People are always paranoid and rightfully so. It's fascinating. It's really fascinating. We talk about freedom of speech, freedom of thought, which is also a very important liberty at the core of this country and probably humanity. It starts to get awfully tricky when you start to be able to collect those thoughts. What I wanted to actually ask you is do you think for fun and for practical purposes, we would be able to modify memories? How far away we are from understanding the different parts of the brains, everything we've been talking about, in order to figure out how can we adjust this memory at the crude level from unpleasant to pleasant? You talked about we can remember the mall and the location, the people. Can we keep the people and change the place? This kind of stuff, how difficult is that?

**Charan Ranganath**

In some sense we know we can do it, just behaviorally.

**Lex Fridman**

Behaviorally, yes.

**Charan Ranganath**

I can just tell you under certain conditions anyway, it can give you the misinformation and then you can change the people, the places and so forth. On the crude level, there's a lot of work that's being done on a phenomenon called reconsolidation, which is the idea that essentially when I recall a memory, what happens is that the connections between the neurons and that cell assembly that give you the memory are going to be more modifiable. Some people have used techniques to try to, for instance, with fear memories, to reduce that physical visceral component of the memory when it's being activated. Right now, I think as an outsider looking at the data, I think it's mixed results. Part of it is, and this speaks to the more complex issue, is that you need somebody to actually fully recall that traumatic memory in the first place. In order to actually modify it, then what is the memory? That is the key part of the problem. If we go back to reading people's thoughts, what is the thought? People can sometimes look at us like behaviorists and go, "Well, the memory is like I've given you A and you produce B," but I think that's a very bankrupt concept about memory. I think it's much more complicated than that. One of the things that when we started studying naturalistic memory, like memory from movies, that was so hard was we had to

change the way we did the studies. If I show you a movie and I watched the same movie and you recall everything that happened, and I recall everything that happened, we might take a different amount of time to do it. We might use different words. And yet, to an outside observer, we might've recalled the same thing. It's not about the words necessarily, and it's not about how long we spent or whatever. There's something deeper that is there that's this idea, but it's like, how do you understand that thought? I encounter a lot of concrete thinking that it's like if I show a model, like the visual information that a person sees when they drive, I can basically reverse engineer driving. Well, that's not really how it works. I once saw somebody talking in this discussion between neuroscientists and AI people, and he was saying that the problem with self-driving cars that they had in cities as opposed to highways was that the car was okay at doing the things it's supposed to, but when there were pedestrians around, it couldn't predict the intentions of people. And so, that unpredictability of people was the problem that they were having in the self-driving car design. It didn't have a good enough internal model of what the people were, what they were doing, what they wanted. What do you think about that?

### **Lex Fridman**

I spent a huge amount of time watching pedestrians, thinking about pedestrians, thinking about what it takes to solve the problem of measuring, detecting the intention of a pedestrian, really, of a human being in this particular context of having to cross the street. It's fascinating. I think it's a window into how complex social systems are that involve humans. I would just stand there and watch intersections for hours. What you start to figure out is every single intersection has its own personality. There's a history to that intersection, like jaywalking, certain intersections allow jaywalking a lot more because what happens is we're leaders and followers, so there's a regular, let's say, and they get off the subway and they start crossing on a red light, and they do this every single day. Then there's people that don't show up to that intersection often, and they're looking for cues of how we're supposed to behave here. If a few people start to jaywalk and cross on a red light, they will also. They will follow. There's just a dynamic to that intersection. There's a spirit to it. If you look at Boston versus New York versus a rural town versus even Boston, San Francisco or here in Austin, there's different personalities city-wide, but there's different personalities area-wise, region-wise, and there's different personalities at different intersections. It's just fascinating. For a car to be able to determine that, it's tricky. Now, what machine learning systems are able to do well is collect a huge amount of data. For us, it's tricky because we get to understand the world with very limited information and make decisions grounded in this big foundation model that we've built of understanding how humans work. AI could literally, in the context of driving, this is where I've often been really torn in both directions. If you just collect a huge amount of data, all of that information, and then compress it into a representation of how humans cross streets, it's probably all there. In the same way that you have a Noam Chomsky who says, "No, no, no, AI can't talk, can't write convincing language without understanding language." More and more you see large language models without "understanding" can generate very convincing language. I think what the process of

compression from a huge amount of data compressing into a representation is doing is in fact understanding deeply. In order to be able to generate one letter at a time, one word at a time, you have to understand the cruelty of Nazi Germany and the beauty of sending humans to space. You have to understand all of that in order to generate, "I'm going to the kitchen to get an apple," and do that grammatically correctly. You have to have a world model that includes all of human behavior.

**Charan Ranganath**

You're thinking LLM is building that world model.

**Lex Fridman**

It has to in order to be good at generating one word at a time, a convincing sentence. In the same way, I think AI that drives a car, if it has enough data, will be able to form a world model that will be able to predict correctly what a pedestrian does. When we as humans are watching pedestrians, we slowly realize, damn, this is really complicated. In fact, when you start to self-reflect on driving, you realize driving is really complicated. There's subtle cues we take about just - this is a million things I could say, but one of them, determining who around you is an asshole, aggressive driver, potentially dangerous.

**Charan Ranganath**

Yes, I was just thinking about this. Yes. You can read it a mile - once you become a great driver, you can see it a mile away this guy's going to pull an asshole move in front of you.

**Lex Fridman**

Exactly.

**Charan Ranganath**

He's way back there, but you know it's going to happen.

**Lex Fridman**

I don't know what - because we're ignoring all the other cars, but for some reason, the asshole, like a glowing obvious symbol is just right there, even in the periphery vision because again, we're usually when we're driving just looking forward, but we're using the periphery vision to figure stuff out. It's a little puzzle that we're usually only allocating a small amount of our attention to, at least cognitive attention to. It's fascinating, but I think AI just has a fundamentally different suite of sensors in terms of the bandwidth of data that's coming in that allows you to form the representation that perform inference on using the representation you form. For the case of driving, I think it could be quite effective. One of the things that's currently missing, even though OpenAI just recently announced adding memory, and I did want to ask you how important it is, how difficult is it to add some of the memory mechanisms that you've seen in humans to AI systems?

## Charan Ranganath

I would say superficially not that hard, but then in a deeper level, very, very hard because we don't understand episodic memory. One of the ideas I talk about in the book, because one of the oldest dilemmas in computational neurosciences, what Steve Grossberg called the Stability Plasticity Dilemma, when do you say something is new and overwrite your preexisting knowledge versus going with what you had before and making incremental changes? Part of the problem with going through massive - part of the problem of things like if you're trying to design an LLM or something like that, is, especially for English, there's so many exceptions to the rules. If you want to rapidly learn the exceptions, you're going to lose the rules, and if you want to keep the rules, you have a harder time learning the exception. David Marr is one of the early pioneers in computational neuroscience, and then Jay McClelland and my colleague, Randy O'Reilly, some other people like Neil Cohen, all these people started to come up with the idea that maybe that's part of what we need. What the human brain is doing is we have this kind of actually a fairly dumb system, which just says, "This happened once at this point in time," which we call episodic memory, so to speak. Then we have this knowledge that we've accumulated from our experiences of semantic memory. Now when we encounter a situation that's surprising and violates all our previous expectations, what happens is that now we can form an episodic - expectations. What happens is that now we can form an episodic memory here, and the next time we're in a similar situation, boom. We can supplement our knowledge with this information from episodic memory and reason about what the right thing to do is. So it gives us this enormous amount of flexibility to stop on a dime and change, without having to erase everything we've already learned. And that solution is incredibly powerful, because it gives you the ability to learn from so much less information, really, and it gives you that flexibility. So one of the things I think that makes humans great is having both episodic and semantic memory. Now, can you build something like that? Computational neuroscience, people would say, "Well, yeah, you just record a moment and you just get it, and you're done." But when do you record that moment? How much do you record? What's the information you prioritize and what's the information you don't? These are the hard questions. When do you use episodic memory? When do you just throw it away? These are the hard questions we're still trying to figure out in people. Then you start to think about all these mechanisms that we have in the brain for figuring out some of these things. And it's not just one, but it's many of them that are interacting with each other. And then you just take not only the episodic and the semantic, but then you start to take the motivational survival things, right? It's just like the fight-or-flight responses that we associate with particular things, or the reward motivation that we associate with certain things, so forth. And those things are absent from AI. I frankly don't know if we want it. I don't necessarily want a self-motivated LLM, right? It's like, and then there's the problem of how do you even build the motivations that should guide a proper reinforcement learning kind of thing, for instance. So a friend of mine, Sam Gershman, I might be missing the quote exactly, but he basically said, "If I wanted to train a typical AI model to make me as much money as possible, first thing I might do is

sell my house.” So it’s not even just about having one goal or one objective, but just having all these competing goals and objectives, and then things start to get really complicated.

### **Lex Fridman**

Well, it’s all interconnected. I mean, just even the thing you’ve mentioned is the moment, if we record a moment, it is difficult to express concretely what a moment is, how deeply connected it’s to the entirety of it. Maybe to record a moment, you have to make a universe from scratch. You have to include everything. You have to include all the emotions involved, all the context, all the things that built around it, all the social connections, all the visual experiences, all the sensory experience, all of that, all the history that came before that moment is built on. And we somehow take all of that and we compress it, and keep the useful parts and then integrate it into the whole thing, into our whole narrative. And then each individual has their own little version of that narrative, and then we collide in a social way, and we adjust it. And we evolve.

### **Charan Ranganath**

Yeah. Yeah. I mean, well, even if we want to go super simple, like Tyler Bonin, who’s a postdoc, who’s collaborating with me, he actually studied a lot of computer vision at Stanford. And so, one of the things he was interested in is some people who have brain damage in areas of the brain that were thought to be important for memory, but they also seem to have some perception problems with particular kinds of object perception. And this is super controversial, and some people found this effect, some didn’t. And he went back to computer vision and he said, “Let’s take the best state-of-the-art computer vision models, and let’s give them the same kinds of perception tests that we were giving to these people.” And then he would find the images where the computer vision models would just struggle, and you’d find that they just didn’t do well. Even if you add more parameters, you add more layers on and on and on. It doesn’t help. The architecture didn’t matter. It was just there, the problem. And then, he found those were the exact ones where these humans with particular damage to this area called the perirhinal cortex, that was where they were struggling. So somehow this brain area was important for being able to do these things that were adversarial to these computer vision models. So then he found that it only happened if people had enough time, they could make those discriminations, but without enough time if they just get a glance, they’re just like the computer vision models. So then what he started to say was, “Well, maybe let’s look at people’s eyes.” So computer vision model sees every pixel all at once, and we don’t, we never see every pixel all at once. Even if I’m looking at a screen with pixels, I’m not seeing every pixel at once. I’m grabbing little points on the screen by moving my eyes around, and getting a very high resolution picture of what I’m focusing on, and kind of a lower resolution information about everything else. But I’m not necessarily choosing, but I’m directing that exploration, and allowing people to move their eyes and integrate that information gave them something that the computer vision models weren’t able to do. So somehow integrating information across time and getting less information at each step gave you more out of the process.

**Lex Fridman**

The process of allocating attention across time seems to be a really important process. Even the breakthroughs that you get with machine learning mostly has to do attention is all you need, is about attention. Transform is about attention. So attention is a really interesting one. But then, yeah, how you allocate that attention, again is at the core of what it means to be intelligent, what it means to process the world, integrate all the important things, discard all the unimportant things. Attention is at the core of it, it's probably at the core of memory too. There's so much sensory information. There's so much going on, there's so much going on. To filter it down to almost nothing and just keep those parts, and to keep those parts, and then whenever there's an error to adjust the model, such that you can allocate attention even better to new things that would resolve, maybe maximize the chance of confirming the model, or disconfirming the model that you have, and adjusting it since then. Yeah, attention is a weird one. I was always fascinated. I mean, I got a chance to study peripheral vision for a bit and indirectly study attention through that. And it's just fascinating how good humans are looking around and gathering information.

**Charan Ranganath**

Yeah. At the same time. People are terrible at detecting changes that can happen in the environment if they're not attending in the right way, if their predictive model is too strong. So you have these weird things where the machines can do better than the people. It's not that it's like, so this is the thing, is people go, "Oh, the machines can do this stuff that's just like humans." It's like, well, the machines make different kinds of mistakes than the people do, and I will never be convinced unless we've replicated human. I don't even like the term intelligence. I think it is a stupid concept, but I don't think we've replicated human intelligence, unless I know that the simulator is making exactly the same kinds of mistakes that people do, because people make characteristic mistakes. They have characteristic biases, they have characteristic heuristics that we use, and those have yet to see evidence that ChatGPT will do that.

**Lex Fridman**

Since we're talking about attention, is there an interesting connection to you between ADHD and memory?

**Charan Ranganath**

Well, it's interesting for me, because when I was a child, I was actually told, my school, I don't know if it came from a school psychologist, they did do some testing on me, I know for IQ and stuff like that, or if it just came from teachers who hated me, but they told my parents that I had ADHD. And so, this was of course in the '70s. So basically they said, "He has poor motor control and he's got ADHD," and there was social issues, so I could have been put a year ahead in school. But then they said, "Oh, but he doesn't have the social capabilities." So I still ended up being an outcast even in my own grade. So then my parents said, okay, well, they got me on a diet free of artificial colors and flavors, because that was the thing that

people talked about back then. I'm interested this topic, because I've come to appreciate now that I have many of the characteristics, if not full-blown, it's like I'm definitely, timeline is a rejection since you name it, they talk about it. It's like impulsive behavior. I can tell you about all sorts of fights I've gotten into in the past, just you name it. But yeah, so ADHD is fascinating though, because right now we're seeing more and more diagnosis of it, and I don't know what to say about that. I don't know how much of that is based on inappropriate expectations, especially for children and how much of that is based on true maladaptive kinds of tendencies. But what we do know is this, is that ADHD is associated with differences in prefrontal function, so that attention can be both more, you're more distractible, you have harder time focusing your attention on what's relevant, and so you shift too easily. But then, once you get on something that you're interested in, you can get stuck. And so, the attention is this beautiful balance of being able to focus when you need to focus, and shift when you need to shift. And so it's that flexibility plus stability again, and that's balance seems to be disrupted in ADHD. And so, as a result, memory tends to be poor in ADHD, but it's not necessarily because there's a traditional memory problem, but it's more because of this attentional issue. And people with ADHD often will have great memory for the things that they're interested in, and just no memory for the things that they're not interested in.

### **Lex Fridman**

Is there advice from your own life on how to learn and succeed from that? From just how the characteristics of your own brain with ADHD and so on, how do you learn, how do you remember information? How do you flourish in this sort of education context?

### **Charan Ranganath**

I'm still trying to figure out the flourishing per se, but education, I mean, being in science is enormously enabling of ADHD. It's like you're constantly looking for new things. You're constantly seeking that dopamine hit, and that's great. They tolerate your being late for things. Nobody's going to die if you screw up. It's nice. It's not like being a doctor or something where you have to be much more responsible and focused. You could just freely follow your curiosity, which is just great. But what I'd say is that I'm learning now about so many things, like about how to structure my activities more and basically say, okay, if I'm going to be - email is the big one that kills me right now, I'm just constantly shifting between email and my activities. And what happens is that I don't actually get the email. I just look at my email and I get stressed, because I'm like, oh, I have to think about this. Let me get back to it. And I go back to something else. And so, I've just got fragmentary memories of everything. So what I'm trying to do is set aside a timer. This is my email time, this is my writing time, this is my goofing off time. And so, blocking these things off, you give yourself the goofing off time. Sometimes I do that and sometimes I have to be flexible, and go like, okay, I'm definitely not focusing. I'm going to give myself the down time, and it's an investment. It's not like wasting time. It's an investment in my attention later on.



**Lex Fridman**

And I'm very much with Cal Newport on this. He wrote Deep Work and a lot of other amazing books. He talks about task switching as the thing that really destroys productivity. So switching, it doesn't even matter from what to what, but checking social media, checking email, maybe switching to a phone call, and then doing work and then switching. Even switching between if you're reading a paper, switching from paper to paper to paper, because curiosity and whatever the dopamine hit from the attention switch, limiting that, because otherwise your brain is just not capable to really load it in, and really do that deep deliberation I think that's required to remember things, and to really think through things.

**Charan Ranganath**

Yeah, I mean, you probably see this, I imagine in AI conferences, but definitely in neuroscience conferences, it's now the norm that people have their laptops out during talks, and conceivably they're writing notes. But in fact, what often happens if you look at people, and we can speak from a little bit of personal experience, is you're checking email, or I'm working on my own talk. But often, it's like you're doing things that are not paying attention, and I have this illusion, well, I'm paying attention and then I'm going back. And then, what happens is I don't remember anything from that day. It just kind of vanished, because what happens, I'm creating all these artificial event boundaries. I'm losing all this executive function every time I switch, I'm getting a few seconds slower and I'm catching up mentally to what's happening. And so, instead of being in a model where you're meaningfully integrating everything and predicting and generating this kind of rich model, I'm just catching up. And so yeah, there's great research by Melina Uncapher and Anthony Wagner on multitasking, that people can look up that talks about just how bad it is for memory, and it's becoming worse and worse of a problem.

**Lex Fridman**

So you're a musician. Take me through how'd you get into music? What made you first fall in love with music, with creating music?

**Charan Ranganath**

So I started playing music just when I was doing trumpet in school for school band. And I would just read music and play, and it was pretty decent at it, not great, but I was decent.

**Lex Fridman**

You go from trumpet to -

**Charan Ranganath**

Guitar?

**Lex Fridman**

- to guitar, especially the kind of music you're into.

## **Charan Ranganath**

Yeah, so basically in high school. So I kind of was a late bloomer to music, but just kind of MTV grew up with me. I grew up with MTV. And so, then you started seeing all this stuff. And then, I got into metal was kind of my early genre, and I always reacted to just things that were loud and had a beat. I mean, ADHD, right? It's like everything from Sergeant Pepper by the Beatles to Led Zeppelin II. My dad had, both my parents had both those albums, so I listened to them a lot. And then, the Police, Ghost in the Machine. But then I got into metal Def Leppard, and AC/DC, Metallica. Went way down the rabbit hole of speed metal. And that time was kind of like, oh, why don't I play guitar? I can do this. And I had friends who were doing that, and I just never got it. I took lessons and stuff like that, but it was different, because when I was doing trumpet, I was reading sheet music and I was learning by looking, there's a thing called Tablature, where it's like you see a drawing of the fretboard with numbers, and that's where you're supposed to put your - it's kind of paint by numbers. And so, I learned it in a completely different way, but I was still terrible at it and I didn't get it. It's actually taken me a long time to understand exactly what the issue was, but it wasn't until I really got into punk and I saw bands. I saw Sonic Youth, I remember especially, and it just blew my mind, because they violated the rules of what I thought music was supposed to be. I was like, this doesn't sound right. These are not power chords, and this isn't just have a shouty verse, and then a chorus part. It's not going back. This is just weird. And then it occurred to me, you don't have to write music the way people tell you it's supposed to sound. That just opened up everything for me, and I was playing in a band. I was struggling with writing music, because I would try to write whatever was popular at the time, or whatever sounded other bands that I was listening to. And somehow I kind of morphed into just grabbing a guitar and just doing stuff. And I realized a part of my problem with doing music before was, I didn't enjoy trying to play stuff that other people played. I just enjoyed music just dripping out of me and spilling out, and just doing stuff. And so, then I started to say, what if I don't play a chord? What if I just play notes that shouldn't go together and just mess around with stuff? Then I said, well, what if I don't do four beats? Go na, na, na na, one, two, three four, one two, three four, one, two, three, four. What if I go one, two, three, four, five, one, two, three, four, five? And started messing around with time signatures. Then I was playing in this band with a great musician, Brent Ritzel, who was in this band with me, and he taught me about arranging songs. And it was like, what if we take this part and instead of make it go back and forth, we make it a circle, or what if we make it a straight line, or zigzag, just make it nonlinear in these interesting ways? And then next thing you know, it's the whole world sort of opens up as, and then what I started to realize, especially so you could appreciate this as a musician, I think. So time signatures. So we are so brainwashed to think in 4/4, right? Every rock song you could think of almost is in 4/4. I know you're a Floyd fan, So think of Money by Pink Floyd, right?

## **Lex Fridman**

Yeah.

**Charan Ranganath**

You feel like it's in 4/4, because it resolves itself, but it resolves on the last note of - basically it resolves on the first note of the next measure. So it's got seven beats instead of eight where the riff is actually happening.

**Lex Fridman**

Interesting.

**Charan Ranganath**

But you're thinking in four, because that's how we are used to thinking. So the music flows a little bit faster than it's supposed to, and you're getting a little bit of prediction error every time this is happening. And once I got used to that, I was like, I hate writing at 4/4, because I was like, everything just feels better if I do it in 7/4, if I alternate between 4 and 4, and doing all this stuff. And then it's like jazz music is like that. They just do so much interesting stuff with this.

**Lex Fridman**

So playing with those time signatures allows you to really break it all open and just, I guess there's something about that where it allows you to actually have fun.

**Charan Ranganath**

Yeah, and so I'm actually very, one of the genres we used to play in was Math Rock, is what they called it. It was just like, this is so many weird time signatures.

**Lex Fridman**

What is math rock? Oh, interesting.

**Charan Ranganath**

Yeah.

**Lex Fridman**

That's the math part of rock is what, the mathematical disturbances of it or what?

**Charan Ranganath**

Yeah, I guess it would be. So instead of might go, instead of playing four beats in every measure, na-na-na-na-na-na-na-na. You go, na-na-na na-na-na na-na-na-na-na, and just do these things. And then you might arrange it in weird ways so that there might be three measures of verse, and then five measures of chorus, and then two measures. So you could just mess around with everything.

**Lex Fridman**

What does that feel like to listen to? There's something about symmetry or patterns that feel good and relaxing for us or whatever, it feels like home. And disturbing that can be quite disturbing.

**Charan Ranganath**

Yeah.

**Lex Fridman**

So is that the feeling you would have if you keep messing math rock? I mean -

**Charan Ranganath**

Yeah.

**Lex Fridman**

- that's stressing me out just listening, learning about it.

**Charan Ranganath**

So I mean, it depends. So a lot of my style of songwriting is very much in terms of repetitive themes, but messing around with structure, because I'm not a great guitarist technically, and so I don't play complicated stuff. And there's things that you can hear stuff, where it's just so complicated. But often what I find is having a melody, and then adding some dissonance to it, just enough, and then adding some complexity that gets you going just enough. But I have a high tolerance for that kind of dissonance and prediction. I think I have a theory, a pet theory, that it's like basically you can explain most of human behavior as some people are lumpers and some people are splitters. And so, it's like some people are very kind of excited when they get this dissonance and they want to go with it. Some people are just like, "No, I want to lump everything." I don't know, maybe that's even a different thing, but it's basically, it's like I think some people get scared of that discomfort, and I really -

**Lex Fridman**

Thrive on it. I love it. What's the name of your band now?

**Charan Ranganath**

The cover band I play in is a band called Pavlov's Dogs. It's a band, unsurprisingly, of mostly memory researchers, neuroscientists.

**Lex Fridman**

I love this. I love this so much.

**Charan Ranganath**

Yeah, actually one of your MIT colleagues, Earl Miller plays bass.

**Lex Fridman**

Plays bass. Do you play rhythm or leader?

**Charan Ranganath**

You could compete if you want. Maybe we could audition you.

**Lex Fridman**

For audition. Oh yeah, I'm coming for you, Earl.

**Charan Ranganath**

Earl's going to kill me. He's very precise though.

**Lex Fridman**

I'll play triangle or something. Or we're the cowbell. Yeah, I'll be the cowbell guy. What kind of songs do you guys do?

**Charan Ranganath**

So it's mostly late '70s punk and '80s New Wave and post-punk. Blondie, Ramones, Clash. I sing Age of Consent by New Order and Love Will Tear Us Apart -

**Lex Fridman**

You said you have a female singer now?

**Charan Ranganath**

Yeah, yeah, yeah. Carrie Hoffman and also Paula Crocks. And so, yeah, so Carrie does Blondie amazingly well, and we do Gigantic by the Pixies. Paula does that one.

**Lex Fridman**

Which song do you love to play the most? What kind of song is super fun for you?

**Charan Ranganath**

Of someone else's?

**Lex Fridman**

Yeah. Cover. Yeah.

**Charan Ranganath**

Cover. Okay. And it's one we do with Pavlov's Dogs. I really enjoy playing. I Want To Be Your Dog by Iggy and the Stooges.

**Lex Fridman**

That's a good song.

**Charan Ranganath**

Which is perfect, because we're Pavlov's Dogs and Pavlov, of course, was basically created learning theory. So there's that, but also it's like, I mean, Iggy in the Stooges, that song, so I play and sing on it, but it's just like it devolves into total noise, and I just fall on the floor and generate feedback. I think in the last version, it might've been that, or a Velvet Underground cover in our last show, I actually, I have a guitar made of aluminum that I got made, and I thought this thing's indestructible. So I was just moving it around, had it upside down and all this stuff to generate feedback. And I think I broke one of the tuning pegs.

**Lex Fridman**

Oh wow.

**Charan Ranganath**

So I managed, I've managed to break an all metal guitar. Go figure.

**Lex Fridman**

A bit of a big ridiculous question, but let me ask you. We've been talking about neuroscience in general. You've been studying the human mind for a long time. What do you love most about the human mind? Like, when you look at it, we look at the fMRI, just the scans and the behavioral stuff, the electrodes, the psychology aspect, reading the literature on the biology side, in your biology, all of it. When you look at it, what is most beautiful to you?

**Charan Ranganath**

I think the most beautiful, but incredibly hard to put your finger on, is this idea of the internal model, that it's like there's everything you see, and there's everything you hear, and touch, and taste, every breath you take, whatever, but it's all connected by this dark energy that's holding that whole universe of your mind together. And without that, it's just a bunch of stuff. And somehow we put that together and it forms so much of our experience, and being able to figure out where that comes from and how things are connected to me is just amazing. But just this idea of the world in front of us, we're only sampling this little bit and trying to take so much meaning from it, and we do a really good job. Not perfect, I mean, but that ability to me is just amazing.

**Lex Fridman**

Yeah, it's an incredible mystery, all of it. It's funny you said dark energy, because the same in astrophysics. You look out there, you look at dark matter and dark energy, which is this loose term assigned to a thing we don't understand, which helps make the equations work in terms of gravity and the expansion of the universe. In the same way, it seems like there's that kind of thing in the human mind that we're striving to understand.

**Charan Ranganath**

Yeah. Yeah. It's funny that you mentioned that. So one of the reasons I wrote the book *Amongst Many* is that I really felt like people needed to hear from scientists. And COVID was just a great example of this, because people weren't hearing from scientists. One of the things I think that people didn't get was the uncertainty of science and how much we don't know. And I think every scientist lives in this world of uncertainty, and when I was writing the book, I just became aware of all of these things we don't know. And so, I think of physics a lot. I think of this idea of overwhelming majority of the stuff that's in our universe cannot be directly measured. I used to think, I hate physics. Physicists get the Nobel Prize for doing whatever stupid thing. It's like there's 10 physicists out there. I'm just kidding.

**Lex Fridman**

Just strong words.

**Charan Ranganath**

Yeah, no, no, no, I'm just kidding. The physicists who do neuroscience could be rather opinionated. So sometimes I like to dish on that.

**Lex Fridman**

It's all love.

**Charan Ranganath**

It's all love. That's right. This is ADHD talking. So, but at some point, I had this aha moment where I was like, to be aware of that much that we don't know, and have a bead on it and be able to go towards it, that's one of the biggest scientific successes that I could think of. You are aware that you don't know about this gigantic section, overwhelming majority of the universe. And I think the more what keeps me going to some extent is realizing changing the scope of the problem, and figuring out, oh my God, there's all these things we don't know. And I thought I knew this, because science is all about assumptions, right? So have you ever read *The Structure of Scientific Revolutions* by Thomas Kuhn?

**Lex Fridman**

Yes.

**Charan Ranganath**

That's my only philosophy really, that I've read. But it's so brilliant in the way that they frame this idea of, he frames this idea of assumptions being core to the scientific process, and the paradigm shift comes from changing those assumptions, and this idea of finding out this whole zone of what you don't know to me is the exciting part.

**Lex Fridman**

Well, you are a great scientist and you wrote an incredible book, so thank you for doing that. And thank you for talking today. You've decreased the amount of uncertainty I have just a tiny little bit today and reveal the beauty of memory, this fascinating conversation. Thank you for talking today.

**Charan Ranganath**

Oh, thank you. It has been blast.

**Lex Fridman**

Thanks for listening to this conversation with Charan Raganath. To support this podcast, please check out our sponsors in the description. And now let me leave you with some words from Haruki Murakami. Most things are forgotten over time, even the war itself, the life and death struggle people went through is now like something from the distant past. We're so caught up in our everyday lives that events of the past are no longer in orbit around our minds. There are just too many things we have to think about every day, too many new things we have to learn. But still, no matter how much time passes, no matter what takes place in the interim, there are some things who can never assign to oblivion, memories who can never rub away. They remain with us forever, like a touchstone. Thank you for listening. I hope to see you next time.