



# Lovesick Cyborg

« [Google Decides Not to Renew a Military AI Contract](#)  
Real 'Westworld' Haptic Vests Better Than Fiction »

## 'Westworld' Science Advisor Talks Brains and AI

By Jeremy Hsu | June 7, 2018 10:17 pm



A scene from the first episode of season 2 of the HBO show "Westworld." (Credit: HBO)

One of many hats that neuroscientist David Eagleman wears in real life is science advisor for HBO's science fiction show "Westworld." The show takes place in a futuristic theme park staffed by robotic hosts who seemingly exist only to fulfill the dark and violent fantasies of wealthy human guests who want to indulge adventure and vice in a Western-style playground for adults. But as the show hints from the very first episode, the robotic hosts are not necessarily content to remain subservient human playthings for too much longer.

During season one of "Westworld," Eagleman took a break from his work as **adjunct professor** in the department of psychiatry & behavioral sciences at Stanford University to visit the show's writers and producers in Los Angeles and have an intense brainstorming session about the meaning of consciousness and the possibilities of **artificial intelligence**. As season two rolls toward its conclusion, Eagleman got on the phone to help separate the show's science fiction from science fact—and to talk about some intriguing real-world questions that may not have answers just yet.

*Warning, there will be spoilers for both seasons one and two of "Westworld" in this Q&A. This interview has been condensed and edited for clarity.*

**Lovesick Cyborg: What has been your role as science advisor for the show, and what does that entail in terms of how often you consult with the writers and producers?**

**David Eagleman:** The thing I want to emphasize is that these guys are really smart and they don't really need me. What I did is I went down to LA last season and had like an eight hour session with them, the writers room and the producers, talking about all

the issues at the heart of the show. Like what would it mean to build consciousness from pieces and parts, can a robot become sentient, what is free will, do we have it, would robots have it, these sorts of questions. And you know, these are the questions that sit right at the heart of what's happening in neuroscience. And most of them, by the way, have no clear answer. So the reason it took eight hours is we were debating all the intricacies of the questions.

What's cool in "Westworld" is they aren't even suggesting they have a clear answer. One of the things happening this season that is very intriguing is the question of free will, where the robots are like "hey, we get free will," and then one of the guys who has written the stories for the [robot] hosts will quote something to them right when they're saying it. And so they realize that they feel like they have free will, but it's not necessarily the case that any of us are as free as we think and maybe not free at all.

One thing I love about the way the writers and producers are doing this is by having those questions there remaining unanswered.

**Lovesick Cyborg: So they're just teasing at different aspects of what is human consciousness and all the things that it entails, like free will, emotion, memory, every little thing?**

**Eagleman:** Well you know a system can have things like memory, of course, without consciousness or free will. A computer has that. When you make a computer program that runs all kinds of sophisticated graphics or whatever, we know that it's just manipulating zeroes and ones and pushing electrical voltages through transistors and that's all it's doing.

So the question is, do we think that your computer program is conscious—that it is like something to be that program—or is it just a manipulation of symbols? That’s really the heart of the question. We don’t really know the answer to that in neuroscience. What we have is an existence proof which is ourselves. We know we are made of pieces and parts: almost 100 billion neurons, each one of which has about 10,000 connections. So we’re made of enormously sophisticated stuff. And yet we know we have consciousness, so it seems it should be possible to build consciousness in a machine.

But on the other hand, we don’t know if our science is too young to realize that there is other stuff involved. What we don’t have right now is a theory of consciousness, and we don’t really know what it would look like. In other words, at what point would I, say, just carry the 2 and do a triple integral, and there you have the taste of feta cheese or the beauty of sunset or the smell of cinnamon? The internal experiences ... it’s not clear how we will even come to phrase those in the language we typically use in science.

**Lovesick Cyborg: It reminds me of what an AI researcher said, which is that maybe “Westworld” is less about “this is how we can get to AI” and is more about what you’re talking about. Holding up a mirror to human consciousness and what makes us tick.**

**Eagleman:** I think it’s both. At least it would suffice to say that it’s about more than just AI. That’s the twist that season 2 has recently taken, is that it’s really about how would you upload a human so you don’t have to die.

**Lovesick Cyborg: Even though there have been a few fringe researchers batting the idea around, it doesn’t**

**seem like it's grounded in what we can do, right? But I'd be curious to get your take.**

**Eagleman:** Here's what I'd say: we simply don't know. Given that the brain is made of pieces and parts, it should be downloadable. We should be able to reconstruct it on any medium you want, whether it's beer cans and tennis balls or silicon chips. It should be that once we understand the algorithms taking place, we should be able to reproduce that. And if somehow that algorithm equates to consciousness, it should be consciousness.

If I put you on this giant other substrate and said "Hey, how are you feeling," and you say "I feel a little cold, a little hungry," I have to assume it's you experiencing consciousness. So it's actually not that fringe-y. The only thing is it's not in our capability to even try this right now. In other words, it'll be another—any number I'd give is made up of course—it'll be another 50 or 100 years before we can say "OK we have enough computational power now, we're going to try to simulate a whole brain."

But people are already working towards that direction. There's no successful project and it won't be for a while, but I think the endeavor is totally straight forward and it's just a matter of time.

**Lovesick Cyborg: Are you saying it might be more about the engineering challenge and getting the computational resources together?**

**Eagleman:** Exactly, yeah. Currently if you wanted to **simulate a full human brain**—if you took the electron microscopy of each [brain] slice and tried to reconstruct that—it would be all the computational capacity of the planet right now. So it's an enormously challenging problem.

But as I said, the neuroscience community has already started climbing that hill. We call it a mountain, because it's a big mountain. But again, there's nothing theoretically preventing the community from trying that.

Now as I said, it's an open scientific question whether, when we get to a super detailed scan of the brain and a full reconstruction of that on a computer and we press go, that will end up having consciousness. In other words, is it just something about the algorithms that you run that equals consciousness, or is there something else we need. And we don't know what that something else is yet.

One of the things going on in "Westworld"—one of the clever ideas because it's set 30 years in the future—is that if we already figured out the principles of brain operation, you could actually replicate it on something much smaller than the brain. So there's the pearl that's inside the host's head, and that takes care of all the brain operations. The idea is that all Mother Nature had to work with was cells and she managed to put together billions of cells into this device that had consciousness, but could you do it in a smaller, tighter way with better technology. So that's part of what is happening there.

**Lovesick Cyborg: I'd be curious to get your view on the technologies we refer to as artificial intelligence right now, which are obviously very far from general AI or human consciousness. How do you see that current gap between those?**

**Eagleman:** The AI right now can do **very extraordinary things** along very narrow paths. You can train a network to recognize the difference between a dog and a cat much better than a human can. But then if you try to teach a network to recognize something like

navigating a complex room or having a discussion with someone, it's a complete fail and you need to train it completely over again. The issue is that what we have is not anything like the intelligence of a three-year-old. Instead, we have these very narrow paths where it does better. The three-year-old child can do things so much better than any AI system we have currently in terms of navigating a room, and socially manipulating adults, and picking up food and putting it in her mouth, and picking up weird objects and playing with them. Just all the things that a human can do quite easily that are extremely hard for AI to do.

So I think it's unlikely we move toward a situation where we have an **AI that is just like humans**. Instead what we will get is AI that is amazing in very particular directions. But what we have with humans is what people refer to as generalized intelligence, where you can do all kinds of tasks, we're extraordinary cognitively flexible.

Like everybody in the world, I love "Westworld," but it's probably not likely we are heading toward a future where we will build robots that look like humans, because there is no point. You might as well build a robot that's better, that's on wheels, that can move faster. There's also not much reason to build an AI that can take care of all the things humans can take care of, when the real reason for **building robots and AI** is for things that you think humans aren't much good at.

Just take Google Maps being aware of every single car on every single road and telling me which way is fastest to go. That's the kind of thing that is incredible. I use it every single day and I never cease to be amazed by how wondrous it is. But it just kind of fades into our background because it's part of the world we live in.

**Lovesick Cyborg: In your involvement with “Westworld,” were there ever any heated discussion points, or parts where you had to put your foot down and say “this wouldn’t really make a lot of sense?”**

**Eagleman:** You know what’s interesting about it? Most of what we talked about are these really big issues. And we did have heated debate about this stuff, but the interesting part is there is no right answer. In other words, it’s not like I, as the neuroscientist, can put my foot down and say this is not true or this is true. So instead we had a really good heated debate about things. But it’s not like there’s a right answer to it that we know of. *Being science advisor for “Westworld” has its perks. Eagleman was able to get the show’s creators to **showcase a special vest** technology that he and his startup NeoSensory designed to provide a sixth sense for deaf and blind people.*

CATEGORIZED UNDER: **MIND & BRAIN, TECHNOLOGY, TOP POSTS**

MORE ABOUT: **AI, ARTIFICIAL INTELLIGENCE, BRAIN STRUCTURE & FUNCTION, BRAIN UPLOADS, COMPUTERS, CONSCIOUSNESS, DAVID EAGLEMAN, INTELLIGENCE, KILLER ROBOTS, MIND UPLOADS, ROBOT HOSTS, ROBOTS, WESTWORLD, WESTWORLD HOSTS CONSCIOUS**