



Neuroethics & Agency: A Discussion with Joshua May, PhD

SUMMARY KEYWORDS

people, brain, neuroscience, implant, learning, depends, problem, behavior, moral, freewill, parkinson, addiction, gut feelings, responsibility, control, true, responsible, understanding, brain implants, brain tumor

SPEAKERS

Joshua May, Anne Zimmerman

Anne Zimmerman 00:04

Welcome to the Voices in Bioethics podcast. Joshua May is a professor of philosophy at University of Alabama, and recently wrote the book *Neuroethics: Agency in the Age of Brain Science*. He joins the Voices in Bioethics podcast today to discuss neuroethics and the intersection of neuroscience and philosophy. Welcome, Josh.

Joshua May 00:25

Thanks so much for having me, Anne. It's a pleasure to be here.

Anne Zimmerman 00:27

Thank you. So you have a bunch of stories in your book, and they touch on people who behave

outside the expected social norms, some of them commit crimes or endanger people. And let's begin with talking about agency and responsibility. Autonomy is a key term in bioethics, but neurological differences and physical problems like cysts and tumors or diseases like Parkinson's can impact behavior. You have examples, including a man who strangles his wife and tosses her out a window and remains relatively calm afterward, only to find that a cyst may have altered his judgment. How do you really describe free will? And is there a definition?

Joshua May 01:08

Yeah, is there any more controversial term in philosophy than freewill? It's hard to define, I suppose in an uncontroversial way. But that's just philosophy. If I could give a pretty straightforward definition, it'd be something like the ability to make choices among options, in light of reasons – in light of one's values and preferences and goals. I think everyone would be happy with that as a general definition. And then the question is, you know, what does that amount to what is it to choose among genuine options and to choose in light of one's reasons, and a lot of people think that neuroscience is suggesting that maybe we don't have this capacity, maybe that free will is an illusion. And I think that all depends on whether we can construe freewill in those neutral terms. Or if we have to assume that we have some sort of deeper – I like to think of it as like a magic – capacity to defy the laws of nature, or to go against our brain chemistry. And certainly neuroscience doesn't show us that we have that power. If that means that we don't have freewill, then that's because we have a certain conception of freewill as requiring that kind of capacity. And I'm part of the point of the book is to suggest that maybe neuroscience is showing us we need to revise, but not reject our notion of freewill. It's a little bit different than we thought it was. But we still have this capacity to choose among options on the basis of reasons.

Anne Zimmerman 02:29

It seems like there is a big challenge to that idea of choice, and it is sort of turning into almost a cop-out to say people don't really control their actions, and maybe that's due to sickness or some type of disorder. And how do you feel about that idea that people aren't responsible for their actions?

Joshua May 02:51

I think it's right, that there's a general tendency among both neuroscientists and the ordinary reader to say, well, if somebody has a brain tumor, or something's going on in their brain, then that does provide a ready excuse for for bad behavior. Part of the goal of the book is to push back against that natural kind of reading as well. I think it's just too simplistic. And one of the things I'm trying to show is that the neuroscience really requires us to have a much more nuanced conception of neuroethics and of human agency more generally. So I think it just depends, sometimes having a brain tumor can compromise one's capacity to make choices and decisions. And sometimes it doesn't. I tell the story of Kevin, the guy who had severe epilepsy, and then a part of his brain removed, and then he started having these really deviant desires,

and ultimately was downloading child pornography and got in trouble for it. Now, it clearly seems like having part of his brain removed, played a role in him having these deviant desires. But does that mean just by that fact alone, that he doesn't have control over his choices? Not necessarily. And actually, the judge was not convinced that he lacked control, at least completely, because he downloaded some of this illicit pornography only at home and not at work. And so that was indicative that maybe he had some control. Clearly, he had some different desires and different preferences after the surgery. But it doesn't necessarily mean that a change in the brain means lack of control, it really just depends on the case. And so it's the theme of the book to say, well, we can't just say somebody has a brain tumor, somebody has some brain surgery, that means that they aren't responsible or don't have control. The same even goes for people, I think with psychiatric disorders, doesn't mean that that's going to say that they have no control, or they shouldn't be held responsible. But it doesn't mean that they shouldn't be excused. In some cases. It all just depends on the case. I know that's a terrible way to put things and sounds like you know, we're not getting much information. But it just depends on the case.

Anne Zimmerman 04:50

It just seems like there's really no bright-line test for control. And in some contexts, a capacity test or a test of legal competence might matter. But those also seem to be changing with neurological discoveries. Do you think there should be some kind of bright-line tests that someone is either in control or out of control, or in control of certain things they do and not other things?

Joshua May 05:14

I think it's difficult to have a bright-line test, I think we can have tests. And that's important for the law. And just for ordinary or ordinary everyday morality, we have to kind of think about is this person going to have less control in a way that's going to diminish their responsibility. And I think we can do that. And neuroscience can help us because sometimes it's hard to figure out what's going on in a person's mind just by looking at their behavior. And so it can be helpful to notice that say somebody has a brain tumor. This relates to another case in the book about this man who did strangle his wife. And then he seemed completely unmoved by the fact that he had done this and that he was now a suspect. And it was helpful to find out that he did have a massive brain tumor in his frontal lobe. It wasn't a tumor actually, a cyst. And you know, that can help us to figure out what's going on his mind. I think the real danger is thinking that just because someone has a cyst, it means that they lack control. That's too simple of an inference. But it's still true, that actually looking at the brain can help us understand what's going on in the mind. So it's almost the way I like to think about is it's like, the brain is a hardware and the mind is software. And so it can be helpful to figure out what's going on with the software of a computer by looking at the hardware and seeing if there's something broken in there, you know, maybe there's a motherboard, it's got some damage to it. But that doesn't mean that, you know, necessarily tells us anything about how the software is working, we have to draw those connections and make those steps. And it's really hard neuroscience right now, because we

basically don't really know how the brain generates the mind. We know a lot in certain areas of neuroscience. I spent a couple of years studying neuroscience, I had a fellowship that bought out my teaching so that I could sit in on classes and join an autism lab. And, you know, I did learn that there are many things that we know, especially about the hardware of the mind, we know about certain circuits, we know roughly what they do, but we don't really know how it'll go, we can go from that to how the mind is working. Even with some really effective treatments like deep brain stimulation for movement disorders like Parkinson's, it kind of shocked me that actually, we don't know exactly how that works. I mean, there are clearly neurons that are that are breaking down in a certain part of the brain in Parkinson's. I mean, we know the pathology very well. But then they stick electrodes down there, and you might think, okay, neurons are dying so we need to, you know, excite those neurons. But it looks like that sometimes you'd need to not excite them and do it in different areas for it to work. It's not really clear, it just you know, it works. It seems to work in many cases, and we're not really sure why. So I think it's hard to have a bright-line test, especially in terms of the neurobiology. Maybe in the future. I mean, I'm gonna predict, you know, probably 100 years, maybe more, you know, we might have a better grip on this. But at present, I think we're still in the early days of understanding how the brain actually generates things like self control, desires, and preferences and reasoning.

Anne Zimmerman 05:15

Maybe scientists are really finding explanations for too many behaviors. I mean, is it good to be analyzing the brain for these things, when behaviors also have to do with social norms and things that are not even inside the body and really have nothing to do with the person?

Joshua May 08:20

Yeah, it can be kind of scary. The more you learn about your own choices and decisions, and you realize, well, it's all just this firing of neurons and various chemicals in your brain. It does have the effect of making us feel like I'm losing a grip on myself having control here. And that's a point where I think that neuroscience is again, calling on us to revise but not reject our conception of ourselves. We do need to think of ourselves as not somehow sitting or standing above our brains and the laws of nature. We are animals, we're part of nature are part of the physical world. We don't even have to settle the issue about dualism or physicalism. Even Descartes thought that, you know, everything goes through the brain. So even if we have a separate soul, it still goes through the brain. And so of course, every action even for a staunch dualist like Descartes must have something to do with our physical bodies and brains. And so we kinda have to get over that notion that well, just because there's a physical explanation for my decisions, that that means that I'm not really making a choice. One of the ways I like to think about is in terms of analogy with a corporation, I think we make this mistake where we assume that our conscious part of our minds are what's really us. And that I think it's like the CEO of a corporation, we tend to say, you know, there's something special about Jeff Bezos, he plays a special role in Amazon. There's a sense in which we kind of tend to think that Jeff Bezos is Amazon. But that's clearly a mistake. I mean, he's not Amazon. There are many other features that corporations actually a very complex agency. And so when we think about our own human

agency, I think it's similar. Our conscious selves are part of it, but there's also all this stuff going on in my brain. And it's a mistake to distance myself from it to say, well, there's this thing that I'm not fully aware of that's going on in my brain. Because that means it's not me. That's true, neuroscience is showing that many of our choices are influenced by various unconscious factors and all of that, but it's still part of me. It's still part of the corporation that I am this complex agent. And I don't know, you know, what the boundaries are exactly. But I think it's a mistake to say, well, just because there is an explanation for my action, somehow it means that I'm not really there.

10:22

Do

Anne Zimmerman 10:22

Do you think we privilege science in a way, and that leads everybody, when they when they hear oh, actually, an MRI confirms this behavioral trait that suddenly they view it as something different, something scientific, and then they think, you know, almost the rules don't apply? Or it changes how we think about responsibility?

Joshua May 10:41

I think that's right, there's, there's something about science, it's a different conception of ourselves. But I think it's almost like looking at one side of the same coin, or maybe think about the famous Duck Rabbit illusion, you know, you can see it as a duck can see it as a rabbit. But it's not really either, or it's both. And so I think that we can think about human agency from the inside and think about our own conscious decisions, the things we're aware of, our private mental space. But then we also have this other way of viewing ourselves, which is through the lens of science, through looking at predicting or explaining our behavior and looking at what's going on in the hardware of the mind. I think it's best to think of them as two sides of the same coin. And so they can seem intention. But I think what's really exciting actually about neuroethics, and about neuroscience in general, is that it is calling for us to rethink those things and try to make connections. So maybe, you know, we've we've thought of these as two separate things. But they're clearly connections, even if they're fundamentally at some deep level different things. Even if I have a soul that's different from my body, still, we need to think about those connections that even Descartes was trying to figure out by really, you know, cutting up cadavers and trying to figure out what was going on in the brain, you know, Descartes had all these kinds of theories about this. So we need to do the same thing now, I think. But we've got more tools available. And we're trying to figure out, you know, which aspects of the brain are responsible for our decisions, for our morality, for our values. And the more we can make those connections, I think, the better we can have conception of ourselves as agents that can have some sort of freewill, even if it's a little bit different than we thought it would be.

Anne Zimmerman 12:12

It seems already in the realm of kind of lifestyle choices or actions that are not criminal, but that are just sort of, ah, there are already kind of excuses. Do you think that this looking into neuroscience more deeply will lead to more disorders or more sort of diagnoses or medicalization of behavior in general – some people blame the brain for their inability to hold down a job or to rein in excessive spending habits. Do you think neuroscience is sort of somehow contributing to having an excuse for those behavior patterns?

Joshua May 12:48

I think that's right. It's a tendency, and it's one that I tried to really resist in the book and tried to provide a framework for resisting. I mean, again, I think it just depends, and we have to have a nuanced neuroethics. So it's true that sometimes by medicalizing something, we can get a better grip on it and understand that really, it's not something you should blame someone for. We've done this a lot with addiction, in particular. And many people thought it's wonderful that we're treating addiction now as a disease. And we're getting a better understanding of it in the brain. And so that means that addicts shouldn't be, you know, moralized and shamed for their their behavior. But we should treat it as a condition that they're not responsible for. And even there, I think that we have to be careful. That can be a useful tactic in certain ways, and it's important to not over-moralize addiction. But it also just as a too simple inference. It's not true that just because someone has an addiction, and that there's something going on in the brain with it, that they don't have control over their lives in certain respects. I think it all just depends. There are aspects of addiction where we can exert control, there are aspects where it does compromise control. And sometimes it depends on the person. Their severity of addiction sometimes depends on the time of day, sometimes say, you can have more or less control, depending upon various features of the day, if you're particularly stressed, if the urges are particularly strong. And I think we have to say the same thing about psychiatric conditions or neurodivergences. It really just depends, every condition – I think we're learning this from the neuroscience – is on a spectrum. And I what I like to emphasize in the book as well is not just a spectrum for each individual person. It's a spectrum for each person across time. So one of the cases I really like is of Elyn Saks. She has schizophrenia. But she's this distinguished professor at the University of Southern California, and she's really seriously grappled with schizophrenia. She has this wonderful book and Ted Talk describing all of it. She has been put into psychiatric hospitals, she has been bound against her will, drugged against her will, and she has experienced serious psychotic breaks. But, it's only sometimes. And when she's well supported when she has a great environment, when she's not particularly stressed, then she can manage and you know, she can be a really successful academic. There are other times where you know, things are not going so well. And it especially seems like very stressful conditions in life often trigger it. And then now she has all these symptoms that are causing problems with, you know, controlling her life and really having sometimes coherent thoughts. So I think it just depends. Depends on the person, depends on the particular time of day, or their circumstances in their lives. And what I tried to do in the book is really show this means that there's a broad spectrum across all of human agency. And the same thing goes for neurotypical people, too. There are

good days, there are bad days. Sometimes, you know, our freewill is compromised, and we can't control our behavior. Sometimes, you know, neurotypical people are a little delusional and engage in wishful thinking. And, you know, they even can hallucinate. So it just depends on the day and what's going on with that individual. So putting somebody into a bucket of you know, there's something going on in the brain or there's not doesn't directly help us assess these really important questions and ethics about who's responsible and how we ought to respect people's autonomy. I think we've got to look more at the individual cases. But the optimistic upshot is that neuroscience can help us do that. And the more that we understand the brain, we can get more nuanced and figure out what is going on, in a case by case basis, I think.

Anne Zimmerman 16:09

I think there's a problem – something I see as a problem – of medicalization across psychiatry and medicine and looking at things through this medical model. When you talk about the schizophrenia, for example, society is very comfortable with saying someone either has a disorder or they don't, and then sort of treating them differently and bucketing them that way. And in the addiction setting, it's probably better to see something as a disease than a crime. But it isn't necessarily better to see it as a disease than as a normal aspect of human nature, which a lot of people might argue addiction is just this natural thing that people have. It's a behavior pattern. Do you see sort of that medical model as something that do you think neuroscience is propelling us even further into that medical model?

Joshua May 16:56

I think that's absolutely right. That the tendency is to gain more of a medical understanding of what's going on. But I think it's the same kind of danger of if we're not too nuanced about it, then we're going to have problems with the kinds of conclusions we can draw. To take an addiction, for example, I think it's really wonderful that we've gained more of an understanding of what's going on there. And we can potentially expand our possible treatments for addiction. But we don't want to just focus on what's going on in the brain and medicalizing it. There's this neuroscientists who I really like to read and counter throughout this process. And Carl Hart, who's who's at Columbia, and you know, he has really been, I think he's kind of soured on the idea of focusing too much on the brain when it comes to addiction, because he worries that if we're too narrowly focused on what's going on in the brain, then we miss out on other factors that are really important for addiction. Like poverty, like alienation, like loneliness, like other psychiatric conditions. And if we focus too much on the brain, we might ignore those things. They're not incompatible, and we really should focus on both at the same time. It's just this kind of narrow focus on one or the other, that's going to be the problem. And I think that we're going to continue to run into this problem. If we only think in terms of categories: if we think in terms of if somebody has a mental condition, or they don't; somebody has a psychiatric disorder or not. I think that's still going to make us run into these problems of "Okay, can we put it in the bucket of it's a medical problem, or it's not a medical problem." And that can be useful to some degree in some contexts, as long as we don't lose sight of the fact that that's a little bit arbitrary. The relevant analogy here, I think, is with many aspects of the law, you know, we have to have

cut offs in the law, we have to say at a certain age, people are eligible to vote. But it's not like once you turn 18, at least, you know, most place in the United States, you've suddenly become mature enough to vote and become a citizen in this democracy. It's clearly a matter of degree. There are some 16 year olds who are much more mature than a 20 year old. And so it just is a bit of an arbitrary cutoff that we have to have for practical purposes. And that's okay, as long as we don't lose sight of the fact that that is a bit arbitrary. So it might be fine to say, look, there are certain cases where we should say it's a disorder, and we should medicalize it. And that's okay, it can be useful as long as we don't lose sight of the fact that that doesn't mean that we can draw simple conclusions from it – we've got to still be nuanced in our assessment of what's going on.

Anne Zimmerman 17:25

It seems the risk is pretty high. When scholars and academic people and people in medicine and psychiatry really focus on the brain. I think it does sort of draw something away from the focus on loneliness, poverty, older adults, the way we communicate in communities. And I think there is still that element of privileging science over social science, and an interdisciplinary approach would seem better. It's again, just sort of that medical model always seems to at the end of the day, have the privilege Do you think that some of these academic hubs should make more of an effort to include social scientists in some of these studies about behavior?

Joshua May 19:56

I think that's a great way of putting it that we need to just attack it at all angles. I do kind of worry about, you know, swinging the pendulum too far the other way and saying, well, the neurobiology doesn't matter. And all we need to focus on are sort of environmental factors or maybe co-occurring mental conditions. I think we've got to do it all at the same time. It's like the Duck Rabbit again, you know, it's not either or, it's both at the same time. Sometimes we can only focus on one or the other at one time. And I think it's really hard, it's hard to attack it at all angles. In the same way, it's hard to see the illusion both as a duck and a rabbit at the same time, that may be kind of psychologically impossible for us. And I don't know that just including social scientists into neurobiological studies is going to necessarily help. I think it can be okay, if we're working in our different areas of academia, we just have to have I think some of these higher level big picture syntheses and, you know, not to toot my own horn, but that's the goal, I think of the book and some, you know, philosophical kinds of integrations of these very different areas of science. That's one of the things I discovered trying to learn about neuroscience is it's just such a massive field. I mean, it compasses, cognitive neuroscience, you know, a lot of people think of neurosciences like fMRI. One of the huge conclusions of the book I hope people take away, especially philosophers, is that neuroscience is more than fMRI. It's much more than that. It's neurology. It's Alzheimer's, it's Parkinson's. There's the medical side of things. There's the cognition side of things, there's psychiatry. It's so hard to integrate all of these very different aspects of human knowledge. But I think the best we can do is just make sure that we're not drawing over-simplified inferences about these moral implications, or these philosophical implications. And I hope the book provides some kind of framework for that. It's not just to say that, "Oh, you know, we can never draw these conclusions." But we've just got to be careful

about the steps going from the neurobiology to things like the social-environmental conditions, you can make those connections, they're just very difficult to do.

Anne Zimmerman 20:03

You also mentioned the distinction between the inability to do something in a particular situation despite understanding the overall mission. And you contextualize that to different neuroscientific theories, for example, distinguishing choosing healthy food in the moment when at the grocery store, from knowing that general concept that someone should have healthy food, that those are two really different, you know, whether it's parts of the brain or sort of functions. But we see that certain deviations from what we sort of consider normal can lead to poor decision making. How people decide what to do in the moment might explain their actions, but are they really reasoning? Or are they acting on impulse? And do you believe distinguishing reasoning from impulse informs moral decision making?

Joshua May 22:39

Yeah, I think this is another area where we have to be careful and also revise some of our intuitive conceptions about human agency. We often want to divide reasoning from impulse, and we want to think of decision-making as again, mostly our conscious minds, you know, what we're consciously aware of. But I think that our understanding of the brain is revealing that there's a lot of our decision-making that is automatic, and unconscious, a lot of it is what we might describe as impulse or gut feelings. And we often use our simplistic categories of saying, Well, you know, reasoning is good, and you know, impulse gut feelings is bad. But I think it just depends on the case. Sometimes we actually can perform a decision that's going to be influenced by a lot of automatic and unconscious processes. And yet the neuroscience is suggesting that that's actually how things normally go, that when you do have patients who have damaged parts of the brain that make them less able to have these kinds of intuitive, automatic kinds of gut feelings to guide their behavior, then they start to have problems. They can't just make simple decisions, yeah, like at the grocery store, or about what to eat or where to go for dinner. And that suggests to a lot of people now – something of a consensus in neuroscience now – that we need both automatic kind of gut feelings to guide our decisions and conscious reasoning, and that ordinary human life just involves an interplay of both of those. So it kind of makes it difficult to say, “Okay, there's reason is good, emotion is bad. And it's really just some sort of complex interplay.” It's another one of these areas where I think neuroscience is forcing us to maybe revise our thinking about this a bit, but certainly not reject it all completely. We still do have these decision-making capacities. We just have to understand that, you know, they don't always just work by sitting there, you know, consciously thinking through a problem, you know, with our head perched on our fists, like Rodin's Thinker, you know, this is kind of the foot conception things, but also the philosophers conception of reasoning that, you know, it's all this conscious deliberation, like we do in a philosophy seminar debate. But a lot of ordinary human decision-making is really just these automatic kinds of skills that we have that we're deploying on the fly. bBut they're not unreasonable, they're not untrustworthy, necessarily. It just is a part of how humans work. You know, we acquire these

kinds of skills and abilities over time and deploy them often quite automatically. It's more like the, you know, the middle managers of Amazon or, you know, these other kinds of, you know, lower level workers that are going on and playing a role in directing the agency. It's not always the conscious CEO. And I think that's okay. We just need to realize that that's kind of how the brain normally works.

Anne Zimmerman 25:14

Do you think unconscious learning is responsible for these intuitions and gut feelings?

Joshua May 25:20

I think that's right. And it's again, another area where there's a lot of consensus actually. Really, in general, there's there's much consensus about the science lately, it's just hard to know what we should conclude from that in terms of philosophy and ethics. But the picture that I developed in the book is one that says, Yeah, unconscious learning is really an important way in which we make decisions over time, but also form our moral judgments and values, you can again see this when it breaks down in cases of patients who have brain damage or abnormal brain function. So psychopathy is a great example. It is true that people who are bona fide psychopaths, um, do have trouble with morality – they tend to behave rather immorally. And we can look at what's going on in the brain to see you know, what's breaking down. And it does seem like they do have problems with some of their gut feelings and emotions that are related to ethics, like guilt, and remorse. And those are just the kinds of things that we require for ordinary decision-making. And really, what's important is that we have those over time. So psychopathy is really something that often starts very young. And you can contrast that with people who have acquired damage to some of the similar areas of the brain later on in life. So if you have somebody who learns ordinary morality, and they can damage the brain later in life, they're not like a psychopath. What the real problem is, is when somebody from very young is having abnormal brain function in these certain areas that are characteristic of psychopathy, and then over time, they fail to learn more rules and norms and understanding and relevant moral emotions. And when you have that kind of problem over time, they just can't engage in that kind of automatic unconscious learning. It's like, I have a 10 year old daughter, and she's learning all kinds of moral norms right now. You know, when it's okay to lie, when it's not okay to lie. And you can constantly just say, I tell, you know, lying is wrong, you know, please don't lie to me. But, you know, she's also learning there are exceptions. Sometimes it's okay to lie, sometimes white lies are okay. And I'm not consciously telling her about those things. She's learning that unconsciously over time, that there are exceptions to more rules, and she's getting a kind of skill or finesse with understanding, you know, when are these rules really applicable? When Are there exceptions? How do you resolve conflicting values, and and that just takes a lot of unconscious learning over time. So I think that's really crucial and central to human learning. And we do a disservice to ourselves, if we think like a lot of philosophers do, which is the, oh it's all just conscious decision-making and just explicitly articulating rules.

Anne Zimmerman 27:48

But recognizing it all as unconscious learning, or at least that that piece matters, and that you kind of pick up values and morals along the way, can that kind of influence values and biases that are relevant to moral decision making? It's there, people would argue there's not just one morality, so if you were raised in a place where it's accepted to have something of a racial prejudice or something like that, and then does that become part of your sort of unconscious learning? And how does that then contribute to decision-making?

Joshua May 28:20

Yeah, I think that's right. It's almost like because you're learning it unconsciously over time. Often, what that involves is somewhat uncritically accepting your cultures, values or maybe your local community. And sometimes that's a problem, like you say, when you have certain kinds of racial biases. In general, though, I've kind of come to this partly through a lot of cultural anthropology, which suggests that a lot of this is actually sophisticated. And one of the things that makes our gut feelings and unconscious learning reliable is when you have a lot of trial-and-error experience with it. So like, I play guitar a little bit when I can, and I played for a long time. So I have a lot of trial-and-error experience with that. So when I'm trying to, you know, figure out a new piece of music and copy it, I have some pretty good gut feelings about you know which chords I should hold or which frets I should press, but somebody who doesn't have experience with that is not gonna have good gut feelings and intuitions about that. So trial-and-error experience is great, you don't just get it from yourself, you can get from your culture, too. So cultures over, you know, many, many years over millennia, have had experience with different kinds of moral dilemmas. And sometimes that's really useful, you know, just absorbing your culture's norms is one of the ways in which humans learned and succeeded over over the many generations. It's true, that also means that if they've got a problem, if they've got biases, and they've got bad ways of resolving conflicts, you're going to absorb that. The good news, I think, is that we do have the capacity to continue learning beyond childhood and beyond adolescence, and we can continue to revise our moral values and norms. And that's really what we do through a lot of dialogue, and especially in a democratic society. And we got to resolve these conflicts later on. But it's actually a good thing. I think that we do learn a lot of these norms over time through our culture, because a lot of times they do basically it's a shortcut, you don't have to go through all those problems yourself, you don't have to go through all the problems of potentially lying in the wrong context, you can just, you know, watch a show or watch movies that show you those kinds of conflicts and how they're resolved. And so it's a great normal way to learn. But it's true, we've got to ferret out some of those biases and problems. I think that's just the ongoing project of, of moral learning and moral progress in society. And I think really, neuroscience again, is giving us cause for optimism that there is a lot of flexibility there. It's true that there's a lot of automatic unconscious processes. But it's also true that we can, we can reflect and we can reason, and we can engage in dialogue with one another. And people really do revise their moral values. In light of those things. It's hard to teach an old dog new tricks, but it's possible.

Anne Zimmerman 30:44

Yeah, it's interesting to see how people learn and absorb. So I want to switch gears now to technology. There are all sorts in the bioethics literature of documents to do with enhancement and treatment and using implants, surgical devices, all sorts of technology and artificial intelligence. And I just wonder about that. If we go back to talking about responsibility and agency, if a surgery or implant or something medical or technological impacts behavior, and someone does something dangerous or commits a crime, what do you think that says about their responsibility? Do you think they should be held responsible? It's different from what we talked about earlier about something naturally occurring like a cyst or tumor or a disease like Parkinson's? But you know, if someone seeks out a doctor to implant something, and has that result, where do you think the responsibility lies, then?

Joshua May 31:40

Yeah, and I think there's a lot of hope for some of these implants being, you know, in our near future, I try in the book to avoid thinking too far in the future in speculating about, you know, science-fiction type scenarios. But you know, we're not that far off from implants. I mean, Elon Musk has this Neuralink company, and they have developed devices that are implants. And right now, they are just for therapeutic purposes, they're using it mostly to try to allow people who say, you know, have paraplegia and they can maybe control you know, remotely a computer to try to communicate and move arms and all that, robotic arms. But they hope, and this is clear, in their statement and their mission, that ultimately this is going to be a consumer product, that they want people to be able to have brain implants that will be kind of like your watch that tracks, you know, your heart rate and your sleeping and all that. Now, I don't know how likely that is to be a reality. I'm pretty doubtful. I think that you know, that's still surgery, people are unlikely to actually do this in a way that they would just wear a watch. So I don't know how likely we're going to get into that kind of future, at least on a large scale. But it's true. There are a lot of patients who are getting brain implants. There are a lot of patients with movement disorders who are getting deep-brain stimulation, it is surgery, but it's often to treat very serious conditions. It's expanding to psychiatric disorders as well. There's there's a lot of research on on major depression, treatment resistant depression – can we get a brain stimulator put in to help people who otherwise might be actually suicidal, and there surgery makes sense, it may be something it's worthwhile to do. Now, it does affect people's minds, it can affect their personalities, it can change, you know, what their values and decision-making might be. Whether that means that they're not responsible, I think it's kind of the same situation we had before. Same as when we have a brain tumor, or we're talking about somebody who's had surgery and a piece of their brain removed. I hate to say it, but I think it just depends. So the fact that they've got the implant doesn't really say one way or another, whether they're responsible, it just depends on how it's affecting their mental state. I think there's a great analogy with intoxication here. It's true that sometimes, you know, ingesting a drug will make it so that you're not as responsible for some behavior. But it just depends. Often we do hold people less responsible if they're intoxicated, only because we think, Well, it's because it's affecting the relevant mental states making them less in control. It's impairing their judgment. But even then, it's not a simple, straightforward connection. We don't say, well, somebody was drunk, so they're not responsible. We still hold people accountable for drunk driving. And that's

because you can think of responsibility much more broadly – are they responsible for getting themselves into that mental state. So if people are freely choosing to get brain implants, and they know that these brain implants are going to affect their decision making capacity or change their moral values, then I think there's still a role for responsibility there, you know, they're still choosing to enter into a different state of mind. Now, if they aren't aware of what the side effects are, or what the potential effects are, then maybe they aren't gonna be as responsible. But I think it just depends. It's not the fact they've got a brain implant, we have to look at what is it doing to their mental state? And were they aware that there was going to do that to their mental state, and then we can still have notions of responsibility apply.

Anne Zimmerman 34:45

Yeah, in the case of the law, where you really could see an intoxication defense, if it negates the mens rea for a specific crime, that makes sense. But when we talk about responsibility more broadly, and somebody's not, it's not something you'd go to prison for, or that has a legal ramification. I wonder how, if people can still try to pass off responsibility just for sort of moral versus immoral actions. There's an example in another book of somebody with a Parkinson's implant, who suddenly has many affairs and buys many expensive cars and things like that. And do you think it properly alters moral responsibility? It's sort of a tricky scenario.

35:26

Yeah,

Joshua May 35:26

Yeah, I mean, you'd have to know more details about the case. But I think it's the same kind of situation where we can't necessarily just say, well, there was an implant, and so there's is or isn't responsibility. It would depend on how it's affecting them. If it is leading to risky behavior, and all of that, which is a known potential side effect of deep-brain stimulation for Parkinson's, partly because, you know, it's neurons that produce dopamine that are that are dying. And so you have to, you know, stimulate those areas that are involved with dopamine reward. And some people become manic, actually, through brain stimulation, you can try to adjust it so that maybe they don't experience that as much, but still can get the movements back, but you sometimes got to crank it up enough that they're actually going to have some of these episodes of mania, where they might engage in some risky behavior, gambling, sexual affairs. I think that, you know, there we're starting to affect some aspects of responsibility and freewill, the ability to make choices and control your behavior. But if they are doing this, because now they you know, they care more about, you know, exciting things or care more about gambling or care more about sexual activity, I think that that is part of who they are, it's part of the complex agency that they are now. And then we have to evaluate it that way, especially if they got into it freely, knowing that these are potential side effects, which patients are informed, you know, that there can be these kinds of effects, and then they're sort of choosing to go down that path of potentially having a change in their their moral values. And I think this is really just something

that we all have to deal with, throughout life, you know, we all have transformative experiences, we all, you know, sometimes we travel or, you know, some people take psychedelics, or, you know, they could even just go to boot camp or, or they have some, they go through a divorce, and they have really radically new outlook and perspective on life. And they care about different things. Sometimes it's for the better, sometimes it's for the worse, and we still hold people accountable for for the behaviors that result from this kind of new personality that they have. So the theme of the book is to try to just compare some of these cases that seem unusual, that are maybe from the medical literature, to a lot of ordinary life, which I think helps us to see that maybe matters aren't so different. When it comes to ordinary life, we do again, just case by case, and we have to still hold people responsible for their actions, even if it came from, you know, experience to change their values. Still, that's them now. And I think we kind of have to say the same thing. Even if it's a medical case.

Anne Zimmerman 37:47

When you say "That's them, now,," I think that's a very interesting phrase, as if something has become the you know, they've taken it in and it becomes part of their personality. There are other regular medicines that cause big behavioral change. For example, some say statins can cause great irritability, where someone might kind of fly off the handle and behave, you know, act in anger when they wouldn't. In the drug setting, do you think that you were just very much the same, a drug and an implant, or even your life experiences, they can just make you change? And it is who you are at any moment? Are we accepting that, that the drug changes you and now you're the new person you and the drug combine?

Joshua May 38:23

I think that's right, although the big difference is sometimes with drugs or other kinds of interventions that can be much more temporary or completely reversible. So some implants even like deep brain stimulation, you can turn it off, you could also remove it. And you can also just turn it off. And some of that's reversible. And sometimes, you can't really reverse things, some people get deep-brain stimulation, and then they have marital problems, and they get a divorce. And you can turn the DBS off. But that doesn't mean that you know, now they're back together. So it says it's partially reversible. So it just depends, I think, on how you know, permanent the change is. But I think even when the changes are temporary, if we just think about normal human life, we still sort of just, you know, switch with the changes. Throughout the day, I may get hangry and irritable, and Genie still hold me accountable for that if I fly off the handle, and you know, I'm, you know, rude to my daughter or something like that. I shouldn't have been, and you know, we can explain it, we say, Well, part of the reason why is because you know, I was just really hungry and irritable. And we might actually mitigate blame to some degree. And I think we could do the same thing elsewhere. If somebody was just temporarily on, say, an antidepressant that really changed their character and personality, then you would push to hold them accountable to some degree in that context. But it does give you a better understanding, especially if they went off of it soon after, then you might say, Well, I have more understanding of what was going on, you still hold them accountable if they were a jerk to you

during that time, but it just it gives you a broad understanding of what's happening. And you wouldn't then assume that they're going to be like that continuing onward. And so it might affect you know, your choice about relationship. And if they were temporarily, you know, they are irritable, no big deal. I can deal with that. But if it's more permanent, you might decide to end that relationship. But those are choices about relationships, not really about responsibility. I think our ordinary practices do tend to just shift with changes in people's personalities and circumstances. And if they're starting to be irritable, and being a jerk, you hold them accountable for that, even if it's really temporary.

Anne Zimmerman 40:10

And when we look at things like drugs, implants, technology, AI, whatever, if they do the opposite, if we are using them to make someone calmer to make someone who is angry and has committed crimes of anger become calm. What do you think of using that in exchange or as a condition for parole? What do you think of, you know, changing a person, and then that really changes their ability to achieve freedom.

Joshua May 40:37

I think that's really tricky in the context of the law, I've gotten really interested in this issue, but largely just thinking about in terms of individual choice, that what if people could make these choices themselves, you know, maybe they just personally think you know, I'm having trouble in my relationships, and I just want to be a better person, and they could try to read a book, you know, a self help book, that's not really helping, maybe they might try to alter their brain directly, maybe take some drugs, maybe take psychedelics, maybe do some non-invasive brain stimulation. I actually am really very much in favor of that if it's safe. And if it's actually effective, I think a lot of these things won't be very effective for that, they just won't be targeted enough. I've toyed with the idea that maybe psychedelics might be a little bit more effective in that regard. But I think it's still early days to try to figure out whether that could be used for moral enhancement. Now, whether you can then you know, sort of compel prisoners to do this, that that gets much more complicated. But I think similarly, there's nothing in principle wrong with people freely choosing to do this. We choose to try to improve ourselves morally, in lots of ways. We choose to have transformative experiences, we choose to open our minds. Maybe if prisoners could freely be part of this and say, Look, this can be a condition of parole, and they can, through informed consent and no coercion, freely choose to do that as an option. I could see that being ethical. Whether it's practical and likely, I think is another question. And I think it's probably very unlikely that we're gonna get direct brain manipulation that's going to somehow drastically improve people. I know a lot of people are hoping this, we could do this with psychopathy, because there's no currently no treatment for psychopathy. And yet these people reoffend regularly. There's a lot of hope there. But I don't really think that we're very close to finding very targeted treatments for improving specific moral behaviors.

Anne Zimmerman 42:19

And on a different note, what do you think of artificial intelligence itself as a moral decision maker? Do you think it's possible that AI can make moral decisions?

Joshua May 42:31

That's a really interesting question. I have thought about this a little bit. I think the more that we understand the human brain, we can see that this could just be implemented in different kinds of systems. Now, whether AI could be competent moral decision-maker, perhaps surprisingly, I would say it depends if we're talking about uncontroversial moral issues. Here's my hot take, I think probably ChatGPT is pretty competent, as it is about giving the right answers about uncontroversial moral issues. I mean, it's trained on, you know, a large database of people talking about these issues. And there have been some studies on this already, people trying to figure out, you know, what kind of ethics do these AI systems have, because they are learning from from corpus of discussion from ordinary humans. And it looks like they reflect pretty much the general population quite well, when it comes to uncontroversial moral issues. So when you should lie, you know, when you shouldn't lie, they're pretty competent about that. But of course, that's not very useful to us. You know, what would be nice, is if we could have an artificial intelligence that can help us resolve moral disagreements about abortion, about euthanasia about issues in bioethics and beyond, I'm extremely doubtful that, you know, AI is going to help us with that, partly because, you know, we can't help ourselves with that sometimes. But it depends, I think greatly on our moral theory. And this means we just have to do some more philosophy. But take this example, suppose that utilitarianism is true. Utilitarianism really makes morality pretty simple, because it's mostly an empirical matter at that point, right? We know what the basic moral value is, it's to maximize happiness. So choose the option that's always going to bring about the most happiness. Well, AI could actually really help us with that there are a lot of philosophers in the Effective Altruism movement, trying to make these calculations and figure out what will promote the most happiness, you know, donating to this charity, you know, eating less meat or, you know, going for a different kind of diet in terms of climate change. And these are largely empirical questions, I suppose that artificial intelligence could help us answer those. And I'm not personally utilitarian. And I don't know that we can just assume that that theory is true. But I think that the right moral theory is probably going to be a little bit more complicated than that. And it'll be difficult to say that we just for artificial intelligence to help us resolve controversial moral issues. I think it's always going to be unknown what's going on under the hood with AI. And for that reason, we're not going to fully trust what's going on. We need to hash out you know, with people, what are your moral values? How are you resolving conflicts? What are your moral principles? And if we don't really know what's going on to the hood exactly, and how it's generating its moral intuitions, I think most of us will be reasonably doubtful, at least when it comes to controversial moral issues.

Anne Zimmerman 45:04

Even thinking about it, it really helps you figure out what a moral decision really is. Because when you start thinking what external data has to go into that decision, I think you see it

differently. And part of what I wanted to get at is it sort of shows that there are things that are not going to be neuroscientific that feed into what a moral decision is, that it's something that involves society and values. So as one last question, maybe to wrap up and go back to some of the original concepts: In the big picture, do you think it's possible that neuroscience is really just going to undermine some of these long held ideas that social norms parenting community, and all these things shape moral development? Is neuroscience going to try to identify morality as something from within the brain? And how will that affect responsibility?

Joshua May 45:53

I think that is the tendency. And that's a trajectory. And it's something that I think it really is important to push back against. And that's something that the book is trying to do is to try to bring this more unified, humanistic picture where we do take seriously the neuroscience, I don't want it to be an either or I don't want to throw out the baby with the bathwater. Neuroscience is really important. It's part of who we are – neurobiology does drive our behavior. But it will be a mistake, I think, if we do just focus on that, to the exclusion of society, the environment and those sorts of things. I think they can be integrated. And that we need to actually shift the conversation in that direction, so that we are talking about all aspects of human life and not just the isolated brain or medicalizing a problem. And if we do that, I think we'll have a more nuanced discussion about these issues. And it will be one that's informed by neuroscience. But it's not just going to be this simplistic kind of conclusion that well, you know, this is what's going on in the brain. So we just need to address it at that level. If you think about addiction, for example, I think that we do need to understand what's going on in the brain. But yeah, we can't ignore issues like poverty, loneliness, and isolation. Of course, those are brain issues, too, though, right of loneliness is, is something going on in the brain there. So they're not independent. And I think maybe this kind of Duck Rabbit analogy can help. It's really just the same thing looked at from different angles. And we will just do we will make a mistake if we just focus on the duck or the rabbit. But we kind of have to look at both of them at the same time, and integrate them in some overall picture. And not that the this book has done that. But it's trying to encourage us, I think, to think more in that direction, and have the conversation move towards one that's not just going to focus on the brain, even though we need to integrate that into an understanding of addiction, of mental illness, of brain implants. All of these things, we do have to pay close attention to the neuroscience. But we've got to draw some nuanced conclusions about that.

Anne Zimmerman 47:47

You've given us lots to think about. Thank you for joining us.

Joshua May 47:51

Thanks so much. It's great to talk about these things.

Anne Zimmerman 47:53

This has been the Voices in Bioethics podcast with Josh May a professor of philosophy at University of Alabama, discussing his book Neuroethics: Agency in the Age of Brain Science. Thank you for joining us.

This podcast is published in partnership
with [Columbia University Libraries](#).



Columbia University Libraries Podcasts