

BRAIN SCIENCE PODCAST

with Ginger Campbell, MD

Episode #53:

Did My Neurons Make Me Do It?

Aired January 17, 2009

The *Brain Science Podcast* is part of sciencepodcasters.org, the website where you can find high quality science podcasts from a wide variety of fields.

[music]

Welcome to Year 3 of the *Brain Science Podcast*, the show for everyone who has a brain. I'm your host Dr. Ginger Campbell. On the *Brain Science Podcast*, we discuss some of the latest books in neuroscience and I interview scientists, philosophers, and other leading thinkers. Our goal is to explore how neuroscience is unraveling the mystery of how our brains make us who we are. Please visit our website at brainsciencepodcast.com and feel free to send me e-mail at docartemis@gmail.com.

[music]

This is episode 53 of the *Brain Science Podcast*, our first episode for 2009. Today's episode is based on the book *Did My Neurons Make Me Do It?: Philosophical and Neurobiological Perspectives on Moral Responsibility and Free Will* by Nancey Murphy and Warren S. Brown. This book was published in 2007 by Oxford University Press.

Before I discuss the key ideas in this book, I want to talk a little bit about why I chose this particular book. The goal of the *Brain Science Podcast* is to explore how discoveries in neuroscience are unraveling the mysteries of how our brains make us who we are. We shouldn't be surprised that these discoveries may challenge long-standing assumptions about what it means to be human. Both science and philosophy are about asking questions, but they approach the process from different directions. Philosophy of mind interests me because it

explores the questions at the intersection between neuroscience and philosophy. For example, questions about free will and personal responsibility have a long history. Understanding that history is important if we are going to incorporate the discoveries of neuroscience into our lives in a meaningful way. There are many influential writers who have concluded that our sense of our selves as free agents is just an illusion created by our brains. But these same writers then say we have to live as if the illusion were true, not only because it is practically impossible to do otherwise, but because we can't imagine a functional society without personal responsibility.

This book's title, *Did My Neurons Make Me Do It?* captures the key question: who is really in control? Is our sense of our selves as something more than our brain an illusion? The authors of *Did My Neurons Make Me Do It?* argue convincingly that the answer is no. In today's podcast I'm going to try to give you an overview of the key ideas in this book. My goal is to keep the episode under an hour, so I'm not going to be able to go into a lot of the detail that is in *Did My Neurons Make Me Do It?* For one thing, some of the philosophical arguments will be of varying interest depending upon your background and whether you have a background in philosophy.

There are a lot of important references besides this book that I will include in the show notes at brainsciencepodcast.com, so I encourage you that if at the end of this episode you want to learn more about this subject, that you check out the show notes. My goal today is similar to that of the authors of this book. It's to show how a physicalist account of the mind can be reconciled with our intuitive sense of ourselves as free agents. Those of you who are long-term listeners may be surprised to learn that the solution lies in using some of the ideas we've actually explored in the past, ideas such as emergence and dynamic systems.

Also, if you listen to my other podcast *Books and Ideas*, you may recall that back in 2007 there was an episode, episode 12, which was a review of a book called *The Myth of Free Will*. In that episode I talked about the things about that book that I agreed with and the fact that there were some things about the book that I did not agree with, the main one being the idea that personal responsibility is negated because of discoveries in neuroscience. While I accepted some of the key ideas of Cris Evatt's book *The Myth of Free Will*, I've always deep

down rejected the idea that free will, in the sense of personal responsibility, that this is a myth. But until I read *Did My Neurons Make Me Do It?*, I really didn't know how to counteract the arguments of thinkers that argued that there was no other conclusion than concluding that free will is a myth, that you had to either be a dualist or you had to accept the fact that free will was not possible. So today, what I'm going to try to do is to hit the highlights of the argument made in the book *Did My Neurons Make Me Do It?* that to me is a very convincing solution to how we can still believe in free will and accept the fact that our brain is a big component of what makes our mind.

The authors of *Did My Neurons Make Me Do It?*, Nancey Murphy and Warren Brown, are both from the Fuller Theological Seminary. Murphy is a philosopher who focuses on anti-reductionism in both philosophy and science, while Brown is an electrophysiologist whose focus is anti-Cartesianism in the cognitive neurosciences. I think it's relevant to note that these are people who do believe in God. However, they are still arguing for a physicalist account of the mind. In the preface of their book, they say that their starting assumption is that neurobiological reductionism has to be false. Some of the important sources for this book I'm going to mention at the beginning, because this book represents the synthesis of a lot of other work. Murphy and Brown quote extensively from *Dynamics in Action* by Alicia Juarrero, and I've actually read most of this book and I can highly recommend it for those of you who want more reading. They also rely a lot on the work of Alwyn Scott, who is a specialist in nonlinear mathematics, Donald MacKay who wrote the book *Behind the Eye*. Beside Juarrero, the most important work that they reference is Terrence Deacon's work on emergence and symbolic language.

So there's going to be some definitions here in the beginning of this episode that are going to be necessary in order to proceed with the main ideas. The first one is, what is the definition of physicalism? And I'm going to quote their definition from page 1: "the denial that anything needs to be added to living human body to constitute a human being." From a philosophical definition point-of-view, physicalism and materialism are the same thing. They are both positions in opposition to dualism, which is the idea that there's something else besides the physical part of us that makes us us, whether you call it a soul or some kind of non-physical mind- that's dualism. Most philosophers of the mind are also physicalists. I think that David

Chalmers is the most famous exception to this rule. He's an explicit dualist. Dualism may be motivated by religion, but not all religious people have to be dualists. For example, the Fuller Theological Seminary is actually a liberal, I think, multi-denominational seminary. The authors of this book claim that most graduates of liberal seminaries are physicalists.

However, popularizers of recent development of neuroscience tend to be both physicalists and what the authors call ardent reductionalists. On page 2 they say, "The most radical reductionists deny the very existence of beliefs, intentions, and so forth." [And this is the kind of position that raises fears about the implications for our notions of rationality, free will, and moral accountability. The goal of this book is to show that such fears are groundless.][Is the previous segment part of the quote?] "Our thesis is that while human reasonableness and responsibility may be explained partially by the cognitive neurosciences, they cannot be explained away." The purpose of this book is not to argue for the physicalist account of the person against the dualist. Instead, what they are setting out to do is to explore whether a non-reductive physicalism is a coherent position. They are challenging the assumption that a physicalist explanation is inevitably a reductionalist explanation.

What I'm going to show, I hope, through this episode is that the reductionist point-of-view is the key problem that leads to the conclusion that we can't have free will. They're arguing, basically, that if you get rid of reductionism and keep physicalism, you can show that what we call free will actually emerges from our neurobiology. Or to quote from them, what they want to show is that the neurobiological equipment makes rationality, responsibility, and free will possible. Reductionism, in case that's not clear, is the idea that the ultimate explanation for things is in the parts. And the implication of that is that the whole is no more than the sum of its parts. And so then you end up with the idea that if the neuron is the fundamental component of the brain, then the neuron made me do it. We've discussed in the past how many aspects of what goes on in our brain occurs at a level that's below our conscious awareness. If you take a reductionist approach, then you could reach the conclusion that that sort of thing means that we are not in control. That is what we are going to challenge today.

In this book, the authors say that they're going to rely on three sources of insight- science, philosophy, and ordinary experience. Their attitude is that these three areas are what they call

inescapably linked. They say that science can direct philosophy away from implausible solutions, while philosophy can help analyze scientific findings. Common sense is the thing that argues against the reductionist explanation. I mean, most of us know intuitively that in the case of living beings, the whole is more than the sum of its parts.

I mentioned that the project for this book started almost ten years before publication. The reason for this was that they found that integrating the science and the philosophy was harder than they expected. This seemed to be partly because of some very deeply engrained world-view issues that they feel complicate problems in philosophy of mind. Two things that they think are engrained world-view issues that they need to attack are causal reductionism and also the view of the mental as inner. Now today I'm going to be concentrating mostly on the issue of causal reductionism- what that is and why it is an inadequate explanation. Both in this book and in the book my Juarrero, there's a great deal of discussion of the philosophical history of these ideas and why they are so deeply engrained and influence people who have no knowledge of where these ideas come from.

I mentioned before that physicalism and materialism are basically used interchangeably in the philosophical literature. They are both considered positions of monism as opposed to dualism. So why do they use physicalism instead of the word materialism? They chose to use physicalist instead of the word materialist just because they wanted to avoid some of the connotations or associations of materialism with atheism. On page 7 they say, "A materialist/physicalist account of the person does not entail a materialist/physicalist world-view. In particular, it does not entail atheism." So basically they're saying that accepting a physicalist explanation of the human being including the human mind does not entail reaching conclusions about the bigger world picture, such as the existence of God. This is an important point because many attacks on the idea of the existence of free will are really attacks against religious dogma. They are attacks against supernatural explanations, they are attacks against the existence of something non-physical that gives us the ability to make choices between right and wrong, and so forth. Instead what they argue in this book is our ability to make choices about right and wrong and make moral decisions is built into our biology and doesn't require another source. They stick with physicalists because they want to both be consistent with current philosophical usage since it's basically a book of philosophy, and also to, like I said, avoid the

connotations of the materialistic/atheistic world-view.

[music]

The position that they are setting out to defend is what they call non-reductive physicalism. It's roughly equivalent to what has been called emergent materialism or emergent monism. They really pretty much avoid the term *emergence* because they're trying to avoid confusion, and also they want to focus on what they think is the key idea, which is arguing against causal reductionism. On page 8 they say, "Our conclusion will be that causal reductionism is the important threat to human self-understanding, and we shall argue that to counteract it, one needs to develop a concept of downward or top-down causation, which in turn needs to be understood in the context of dynamic systems." Okay. Causal reductionism means the idea that all the causes of what happens are explained at the bottom level. If that was true then we really wouldn't have free will. And instead they are going to be arguing for what they call top-down causation.

When I first started reading this book, my first reaction to the term top-down causation was that I thought that maybe they were talking about something like God. And if that was how you reacted, stick with me because you'll see very shortly that that's not what they're talking about. Basically they are not trying to base their account of physicalism on the defeat of Descartes, but they say their account comes from the emerging recognition that the natural world is what they call a hierarchy of levels of complexity. They set out to consider developments in science from physics up through studies of primates, and I'm not going to be able to get into the details of this, but one thing that they say is that they believe that we have to look at the precursors in the animal world if we're going to understand how human thought and behavior works. This is very much in harmony with things that we have talked about in past episodes of the *Brain Science Podcast* and that we are going to continue to examine.

Now I want to continue our brief overview of *Did My Neurons Make Me Do It?* The first problem is that many people still function with a Newtonian worldview. What does this mean? Well it means that they tend to see everything as being very linear, for one thing (and I'm going to explain what that means in a few minutes), and mechanical. The famous expression

is the "clock-like universe"- it's a mechanical worldview. And at the same time, a lot of people still have what is a Cartesian view of the mental and what this means is, Descartes was known for his concept of the mind as being something totally non-physical that somehow interacted with the body to make everything happen.

The problem that comes out, according to these authors and I think that they're probably right, is that a lot of scientists just take- they don't believe in a non-physical mind but they just replace that with brain and keep on looking at everything the same except for with the expression *brain*. And this leads to what they call brain-body dualism. Why does this matter? Because both of these viewpoints tend to lead to what they call causal reductionism- again, the idea that the parts determine the behavior of the whole unilaterally. That is the idea that the parts are controlling things and that the whole doesn't have any influence on the parts.

Now I give you an example of what I mean by that. If you look at a clock, okay, the way the parts are arranged determines how the clock operates. A clock doesn't have any control on the parts. So that's a simple example. The goal of this book is to defeat causal reductionism. And you may say, "Well that's just all philosophy," but I hope that it will become clear as I go along that if we want to see how our intuitive sense of having free will and the ability to make choice- if we want this to make sense and be able to incorporate that intuition with the things we're discovering about how the brain works, we need a different model.

Now how would one defeat causal reductionism? I'm going to go over that several times today because it's a key idea. But first of all, you have to be able to show that systems can act as their own causes. That sounds impossible, but I will show you what that means. Also, you have to be able to show that top-down causation functions by a means of constraining the behavior of the parts. They say on page 10, "This would be common sense if not for philosophers having convinced us that all causal work must be done at the level of subatomic physics."

The key thing that has to happen is that we have to have a switch in perspective from things to systems, where systems include processes. And these systems are dynamic systems. The word dynamic implies nonlinearity. The simplest way to understand it is to know that as soon as you start putting feedback into a system, it makes it nonlinear. All living systems are using

feedback and so they are dynamic systems. To make this shift in thinking, we have to recognize where the Cartesian distorts our thinking, like when we replace talk of the mind or the soul with the brain. I've read several books recently where philosophers are attacking things like using the term brain in sentences in which you really should be talking about the whole person, not the brain. And I'll talk about that a little bit more in just a minute.

The other piece of the Cartesian worldview that we have to let go of is the idea of the knower as passive- like knowing is something you can do in your head just quietly sitting there. And this neglects the role of action on the life of the mind and it also neglects the social aspects of mental life. And these are two things that are very important in changing over to a more modern view of the mind. They are arguing for shifting to an understanding that sees the mental as pertaining to a higher-level dynamical system- that is, the brain and the body involved in interaction with the world, both physical and social. They say that the mind manifests as "informed engagement in action-feedback-evaluation-action loops and the environment." [are the hyphens in the previous quote in the proper place? 0:23:43] It's a little bit of an awkward expression, but what I thought of when I read this was when I was interviewing Dr. Seth Grant and he made the observation that decision-making first appeared even before our brains or nervous systems- remember he talked about the yeasts moving toward food or away from noxious stimuli yeasts as being the first example of decision-making. And that was before nervous systems or brains. Their position is that if we equate the mind with the brain alone, that this is an, what they call, egregious form of reductionism.

[music]

I want to take a moment to thank [Audible.com](https://www.audible.com) for their ongoing sponsorship of the *Brain Science Podcast*. Many of the books that are covered on the *Brain Science Podcast* are available from Audible.com. New Audible users get one free audiobook download. Just go to audiblepodcast.com/brainscience.

[music]

Before we can consider their attack on reductionism and their defense of downward

causation, we really need to think a little bit more about the problem of considering the brain and the mind as being the same thing, because the rest of the argument won't make sense if all you think of when you think of mind is just the brain. If you think of the brain as being all there is, then you really get into problems of having elements of what goes on that appear to be epi-phenomenal.

Early on in the book, I think in the first chapter, they argue that if you think of the mind as being the same as the brain, then you end up with what they consider to be sort of a materialistic dualism. They actually call it Cartesian materialism. I think they took that phrase from Daniel Dennett. The first thing is that they argue that brain events do not equal mental events. Instead, their definition is: "a mental event must be understood as a contextualized brain event." They are challenging what they call brain-body dualism. Now the reason that they argue that brain events do not by themselves equal mental events is- and this is where stuff that we've talked about in the past comes in- is because "the mind is embodied not embrained." I think this is a quote from Antonio Damasio that is worth repeating: "The mind is embodied, not embrained."

Let's consider some of the evidence that shows how important the body is to the mind. For example, the role of emotion. We've talked in the past about the importance of the parts of the brain that monitor the body's internal state. Antonio Damasio is the one that has shown that when parts of the frontal lobe that are involved with getting emotional input are damaged, a person can't even make decisions. So we need the body's input. That's why separating the brain out by itself is a dangerous example of reductionism. It's the whole person that has experiences, not the brain.

There's one other piece of this that sort of contributes to our tendency to think of it this way, and it's that second Cartesian assumption of the mental life as being something inner and solitary. One neurobiologist, Leslie Brothers, has argued that you can't even really have a mental life. He argues that a mental life is only possible as a result of socialization, language, and conversation. If you've ever read Descartes you know that he, in his writing, he had this concept of the Cartesian ego that only knew he- was only sure that he existed and had to try to prove the existence of others.

Let me give you an example of a recent finding that argues against this whole idea. Remember when we were talking about mirror neurons? Mirror neurons actually seem to allow us to have a pretty good idea about what's going on in the minds of other people. So this challenges the assumption that we can only know our own minds directly. Most normal people do not have any trouble telling, for example, the basic emotions that another person is experiencing. So this is an example of how we can know the minds of others directly without them telling us what's happening. The consequences of sticking with these Cartesian assumption of our inner life as being something totally solitary rather than having important social elements can be seen in psychology, where this assumption that the world is experienced by an inner agent leads therapists to ignore the social embeddedness of people, as if their situation was not, you know, a part of what was going on.

Murphy and Brown argue that instead of a separate inner agent, which they call the Cartesian self, we ought to be considering only the person as a whole as an agent. We don't have a separate agent inside of our heads. The reason I spent time on this is because this subtle difference is important to understanding a dynamic systems approach. If you are trying to find this mysterious inner agent and somehow give it control, then you will probably have to conclude that free will is a myth. But if you can change your viewpoint to thinking of the person as a whole, as the one who is the agent who is making choices, decisions, and has moral responsibility- then you can re-examine the question.

Another key idea that becomes important in trying to understand this from a dynamic systems point-of-view is part of this involves rejecting the idea of this inner self as a passive inner self. Instead, we need to see ourselves and even animals as active agents in the world. Remember when we talked about Gibson's affordances, which is perceiving the world in terms of the actions you can take toward objects such as things you can sit on, things you can pick up, and so forth. Basically, this is an argument for rethinking the whole field of cognitive psychology from the point-of-view of embodiment.

This is again, remember, something we talked about with several people last year, including Arthur Glenberg. Raymond Gibbs is one of the writers who has really emphasized that human

cognition is fundamentally shaped by bodily experiences. And this includes the importance of our memories of bodily experiences. So like I said, many of our guests in the past have talked about the importance of embodiment. Another writer that I hope to interview in the future is Andy Clark. He has introduced the idea that our thinking is dependent on what he calls *external scaffolding*, and this includes language. One of the things that we also know based on research is that neural processing is not based on symbolic knowledge. Remember we interviewed Rolf Pfeifer, who has been building robots that are doing a better job of emulating what living things do by not trying to be programming them with symbolic logic but programming them in a way where they basically react to the environment. In other words, so that these robots are embodied.

There's a quote on page 38 that I think sums up this key idea: "It is the person as a whole who perceives and acts with the help of neural machinery, just as hands and eyes help us grasp and see. It's important that we constantly remind ourselves that mental phenomena pertain to the entire person- brain and body- in social relations and active in the physical world." That's what they mean when they say "mental events are contextualized brain states."

Here's another quote that I think is worth reading: "A mental state is a brain-body event relevant or directed toward a social or environmental context- past, present, or imagined in the future. This is in contrast to all those theories that put mental events exclusively in the brain." Again, the reason for spending time on this point is that we have to have a bigger view of mental events and of who it is that's making the decisions if we're going to begin to approach this from a systems - dynamic systems- point-of-view. The main theme of this book is to argue against causal reductionism and for top-down causation. But that top-down causation does not refer to anything supernatural. Basically what we're saying is we are arguing against causal reductionism, which means the whole is only the sum of its parts, and we are arguing for top-down causation, which is simply a way of saying the whole is more than the sum of its parts. The fear that a physical explanation for the mind robs us of free will and personal responsibility is based on the assumption that a physicalist explanation must be reductionist. That's the reason why the authors attack reductionism and propose top-down causation as a model that actually explains how our reason and moral responsibility arise from our biology. This means that we must show that reductionism is inadequate and how

top-down causation is possible.

[music]

After arguing that the mental life is more than just what our brain does and that any consideration of our behavior needs to be considerations of our behavior as a whole, then Murphy and Brown set out to argue against reductionism and for top-down causation. I hope it's clear that if you could prove top-down causation, then you have rescued free will, so that's the key goal. There are two parts to the argument against reductionism. First, they have to show that a physicalist account is not necessarily a reductionist account, and they have to argue that reductionism is an inadequate explanation.

To reach the first part of just basically showing that a physicalist account is not inevitably reductionist, this requires looking at what's happened in science in the last 20 or 30 years and showing how ubiquitous dynamic systems have become in explaining what's going on. The conclusion that reductionism is inadequate emerges from the realization that a non-reductionist dynamic systems approach has greater explanatory power, which means that it can explain features of the world that seem paradoxical from the reductionist perspective.

The example highlighted in *Did My Neurons Make Me Do It?* comes from Steven Johnson's book *Emergence: The Connected Lives of Ants, Brains, Cities, and Software*. Murphy and Brown observe that the behavior of an ant colony can be viewed from a reductionist bottom-up point-of-view or from a dynamic systems point-of-view. In the bottom-up approach, the behavior of an individual ant is explained by using only its hard-wired rules of behavior, while a dynamic system approach sees each ant as a small dynamic system whose behavior is partly determined by the colony as a whole via environmental signals such as pheromones. It's not that the reductionist model is wrong or that it can be disproven, but there are elements of ant colony behavior that can't be explained with reductionism. For example, it is known that ant colonies of the sort described by Johnson have a lifespan of about 15 years, and that the colony becomes more stable over time despite the constant turnover of most of the individual ants. This behavior is actually typically for an adaptive dynamic system, but it seems mysterious from a reductionist point-of-view.

Making the paradigm shift to dynamic systems doesn't mean that we ignore the power of the reductionist approach. It just means that we acknowledge its limitations. An analogy might be looking at the relationship between quantum mechanics and Newtonian physics. Quantum mechanics can explain things that Newtonian physics can't, but we still use Newtonian mechanics to solve most of our day-to-day problems. Quantum mechanics can incorporate and explain everything that Newtonian physics does, plus it can explain things that Newtonian physics can't explain. Similarly, a dynamic systems approach can incorporate reductionism as a powerful tool. So just as a physicist needs to know when to apply quantum mechanics, those studying living systems need to be able to recognize when reductionism is inadequate. Murphy and Brown argue that understanding much of the behavior of living systems requires a shift from reductionism to dynamic systems. Making this shift allows us to see that while free will appears impossible using reductionist thinking, it actually emerges quite naturally in a dynamic systems model.

From here on I'm going to concentrate on exploring how top-down causation emerges from a dynamic systems approach. I'm going to start with a brief discussion of Terrence Deacon's work on emergence because emergence is the feature of dynamic systems that explains how reason, moral responsibility, and free will are possible. Emergent phenomenon are those that have novel properties that aren't exhibited by their constituents and that have regularities that can't be deduced from laws pertaining to their constituents. The key idea is that this is not emergence of behavior that violates known physics, but it is the emergent phenomenon that demonstrate a degree of autonomy from its constituents, and the emergent phenomenon is the thing that has the ability to demonstrate top-down influence over the properties and dynamics of its constituents.

This might make more sense if we consider what Deacon has proposed as the three categories of emergence. The first-order emergence is the type that can be described without considering history. An example of this would be the turbulence of large bodies of liquid. You can use higher-level descriptions to describe the liquid that don't apply to the individual molecules. So notice that emergence doesn't necessarily require anything to be alive. Now at the next level, which is the second-order level- this is when there is a temporal development or, if you heard

Dr. Wilczek's interview you know he talked about symmetry-breaking. So there's change across time. This is where both chaotic and other kinds of self-organized systems come in. Once you get to the second order you can't describe what's happening without taking history into account. This is when you start to have recursive feedback. You have feedback at first-order, but this time you have recursive feedback, which basically means that whatever happens is going to take the place of prior states irreversibly.

Finally, third-order emergence [0:43:15] involve some kind of information or memory that gets fed back into the lower levels and constrains their future states. This is where living organisms come in. According to Deacon, even the simplest forms of life are third-order emergent, which means that they can't be understood apart from both their history and function. Now to me that just means the whole is more than the sum of its parts. According to Deacon, the third-order, or evolutionary emergence, contains the second-order, self-organizing emergence as a limiting case, which in turn contains the first-order emergence as a limiting case. And this is a sort of a key idea. However you look at it from a systems point-of-view it's always nested one thing inside of the other.

Let's take language as an example. We have an alphabet where we're only allowed to use a certain number of letters which can be made into a larger number of words, which can be made into an infinite number of sentences. But the sentences determine what words are allowed, and the words determine what letters are allowed. Of course this isn't a dynamic system- it just demonstrates the idea of nesting and also how the top level constrains the lower level. So we're restricted to 26 letters, yet at the higher level we can have an infinite number of sentences.

[music]

So let's say that we accept that describing living systems requires a systems approach with concepts of emergence and different levels of hierarchy that are interacting with each other such that we can't explain what's happening at the higher levels just by telling you what the parts are at the bottom. Then the next question is then, well how do these systems cause anything or how does downward causation happen?

The first principle is that the system constrains the behavior of its constituents. The simplest way of thinking about this is, say any system in your body- your digestive, for example- how it operates is constrained within the operating parameters of the body as a whole. That is what that means. One of the key ideas of this whole approach is that while the system constrains the behavior of its components, the system as a whole gains freedom and gains repertoire. Remember the example where I showed how a limited alphabet can be used to create an unlimited number of sentences. Consider all the things that our brains can do that our neurons can't. The complex abilities that we associate with being human are a function of our entire nervous system, not its components. This basic concept is essential to understanding the emergence of autonomous systems.

Recall that when we were talking about the example of the ant colony I claimed that a key reason for going to the use of dynamic systems over a reductionist point-of-view would be when the explanatory ability of dynamic systems was superior to the reductionist approach. This is certainly the case when trying to understand the emergence of autonomy. I want to mention another point that Murphy and Brown made with regards to the example of the ant colony. If you look at the number of neurons in an ant colony, it's about the same as the number of neurons in a human brain. So if all that mattered was local interactions between neurons, an ant colony would be as intelligent as a human. But of course we know that the neurons in our brain are massively interconnected. The point is that this interconnection is what allows for the emergence of many levels of downward causation and many more degrees of freedom.

So how do we get from the concept of dynamic systems, which doesn't necessarily imply anything intelligent, how do we get to intelligent action? This is where we come closer to neuroscience per se, because we can start to look at what it is about brains or nervous systems as they evolve that allows us to arrive at the level of people with their ability to reason and even have moral responsibility.

But starting at a lower level, let's consider the difference between a wasp and a rat. They discuss this wasp in here that basically what it does is kill a cricket and brings it to its burrow

and it lays its eggs on the cricket and then flies away and then basically leaves the cricket to feed its offspring. When the wasp gets to its burrow it leaves the cricket behind, goes in, is getting the burrow ready, and if while it's in there you move the cricket, when it comes back it just moves the cricket back to the opening and goes back in and starts over again, because it doesn't have any memory. So memory is obviously an important component.

Even so, one of the principles of this book is that the same processes that operate in animals operate in people, but they are in nested hierarchies that become more and more complex. The authors say that they think it would be foolhardy to try to understand human cognition without tracing its development from the most rudimentary forms of interaction between organisms and the environment. And an example of that would be, as we talked about before, even the simple decisions of yeasts. The fundamental thesis here is that when we're looking at animal behavior, there is a continuous progression from the simplest level- we do see even then goal-direction and evaluation or feedback- but it's a continuous progression with the possible exception of what's known as the symbolic threshold, which is considered critical.

The whole issue of the development of the ability to use symbols- that is language- is one that they devote a whole chapter to. The first thing that we need to add, as I've showed in the example of the wasp- was memory. Rats can remember where food is, in contrast to the wasp's inability to remember that it's already prepared its burrow. Then the next thing that has to be developed is something called representation- the ability to have some representation of the environment that can be either non-symbolic or symbolic. If you have representation then you can have the ability to do what's known as an offline simulation. It's been shown that this is possible in animals that do not have symbolic representation. It used to be assumed that that was something that required language or symbolism, but that's been disproven by the fact that there are animals that can figure out how to solve problems, and it's been shown by how they do it that they must have done some simulation in their mind before they tried a solution. They weren't doing trial and error- they figured it out and then did it on the first try.

The last thing that comes along is the ability to evaluate one's activities, and that probably does require symbolism. That's the thing that people can do that as far as we know other

animals can't do. It appears that the frontal lobe fits into a nested hierarchy of a complex system. We can look at the frontal lobe as being part of a nested hierarchy within a complex dynamic system. And there are different models for this. One that they mention in the book is called the dynamic core model, from Edelman. The idea is that as the brain develops these increased abilities, then it becomes possible to have intelligent action.

And what about language? They talk about this in great detail, especially basing it on Deacon's work. I think the differences between humans and primates demonstrates the principles of complex systems and downward causation. In people, the motor part of the frontal lobes connects the midbrain, brainstem, and spinal cord in order to control vocalization, but in primates the control is at the level of the midbrain. And what this means is that for primates, vocalization remains involuntary because they don't have higher level parts of the brain connected to the parts that control vocalization. If you are interested more about the role of language, besides the writings of Terrence Deacon you want to look at the work of Lakoff and Johnson and also even Wiggstein, who considered the social element of language as being really important.

[music]

At this point, I feel that I have barely touched the main ideas of this book *Did My Neurons Make Me Do It?* But I hope that I have given you at least a feel for what the arguments are based on. And I want to go back over just what I think are the main ideas of this book. First, a physicalist description of mind does not inevitably threaten free will. A non-reductionist physicalist account is consistent with systems theory. Top-down causation is ubiquitous in dynamic systems. Top-down causation does not require any new laws or breaking the laws that we know. Reductionism implies that the whole is only the sum of its parts, while emergence says that the whole is more than the sum of its parts. A dynamic systems approach gives us a new paradigm with greater explanatory power than the old reductionist worldview. In the end, the answer to the question "Did my neurons make me do it?" is no. My sense of myself as a relatively autonomous person with an ability to make choices is not an illusion. It is a result of my being a whole person with a body and a brain embedded in an environment that includes me, my culture, and the people around me.

While I have shown how dynamic systems and top-down causation can explain the emergence of increasingly complex behavior, I have not dealt directly with moral responsibility and free will. In the later chapters of *Did My Neurons Make Me Do It?*, Murphy and Brown examine the question of exactly what kinds of capabilities are required for a person to have these capacities. They see moral responsibility and free will as being very closely related. Since most of us agree that only humans can have moral responsibility, their discussion naturally focuses on the capacities that are uniquely human with a special focus on the emergence of language and culture.

Obviously we must be able to remember and evaluate our actions. Murphy and Brown say moral responsibility requires the ability to evaluate one's actions in the light of what is right and wrong. But they also argue that there is a social component. Children have to be taught what's right and wrong. And even though humans don't usually see primates as morally responsible, those who study primate behavior know that primates do demonstrate cultural norms and punish violators. Obviously language takes this all to another level. The concept of free will adds the component of autonomy. Here again, top-down causation is essential, along with the ability to evaluate one's choices. If we appreciate that reason, moral responsibility, and free will are naturally emerging properties of human biology, then we can also acknowledge that some people are impaired in these areas for various reasons. But I hope I have shown you why "my neurons made me do it" is never an adequate explanation for human behavior.

In closing, if my brief consideration of *Did My Neurons Make Me Do It?* has left you unconvinced that free will not only exists but is actually a natural result of our neurobiology, then my advice to you would be to read *Did My Neurons Make Me Do It?* yourself. I also recommend Stuart Kauffman's book *Reinventing the Sacred* for a more detailed analysis of the failure of reductionism and the importance of emergence.

In future podcasts, we are going to be considering more of the philosophical implications of neuroscience, because for those of us that are not doing neuroscience as a profession, the main reason that we want to know about how our brains work is so that we can live healthier,

more meaningful lives. And part of that is to examine the implications and meanings of what we are learning about how our brains work. The more we learn about how the lower level systems work and how many things are happening outside of our conscious awareness, the more we can appreciate the importance of our making conscious top-level decisions, especially when there's things about our habits that we want to change. If we want to learn new skills, we have to face the fact that it's sometimes difficult. Sometimes our old habits will be difficult to change, but it is possible.

[music]

I hope this has been a thought-provoking episode and I look forward to hearing your feedback so that I can have some ideas of where to carry this forward in the future. Next month's episode of the *Brain Science Podcast* is an interview with neuroplasticity pioneer Dr. Michael Merzenich of Posit Science. So we will be returning to the practical applications of neuroscience, and Dr. Merzenich will tell us about how the findings of neuroplasticity can help people of all ages.

Before I close, I want to remind you to visit the website at brainsciencepodcast.com. There you will find detailed show notes. I'm going to try to put as many references from this book as I can up on the website. Also you will find links for episode transcripts and information about how to subscribe if you aren't already subscribes.

I love getting e-mail from listeners and I'd like to apologize to those of you who may have written me e-mails in the last several weeks. As I have been trying to get this episode together I have not been answering e-mails, so I hope to get caught up with that in the next few weeks. You can send me e-mail at docartemis@gmail.com. I'd also like to encourage more of you to participate in the discussion forum at brainscienceforum.com. That's the best place to put up your feedback, especially if you want to exchange ideas with other listeners. I read all the posts on the forum even though I don't respond to most of them.

I want to thank those of you that have been sending contributions to help support the *Brain Science Podcast* and I've been getting a really good response to the new mailing address that I

put up on the website, but I want to mention something that I found out from a listener from Chile. I found out that my bank will not accept checks that are written on foreign banks. And the listener from Chile said that if she tried to get her money put into a money order, that they were going to take out a pretty good chunk, which may mean that PayPal is actually the way to go if you're outside the United States. However, if you're outside the United States and you don't want to use PayPal, you may want to see if you can find a way to use Western Union.

So as I announced in the last episode, *Brain Science Podcast* is going to be coming out once a month in 2009. That should give you a chance to get caught up on old episodes and also if you need something else to listen to, you can check out my other podcast, *Books and Ideas*. I just did a second interview with Dr. Frank Wilczek, the Nobel Prize winning physicist, and I'm very proud of that interview. So I hope you will check that out at booksandideas.com. Thanks again for listening. I look forward to talking to you again next month.

[music]

The *Brain Science Podcast* is copyright 2009, Virginia Campbell, MD. You can copy this podcast to share it with others, but for any other uses or derivatives, please contact me at docartemis@gmail.com.

-transcribed by Jenine John