THE TRANSMITTER | BRAIN INSPIRED

Xiao-Jing Wang outlines the future of theoretical neuroscience

Wang discusses why he decided the time was right for a new theoretical neuroscience textbook and how bifurcation is a key missing concept in neuroscience explanations.

2 JULY 2025 | by PAUL MIDDLEBROOKS

This transcript has been lightly edited for clarity; it may contain errors due to the transcription process.

[music]

Xiao-Jing Wang
It becomes actually really a puzzle if you assume that different areas are made of the same stuff, and they all talk to each other in a dense network. How do you get differentiation of function really?

You can think about bifurcation as a mathematical machinery to create novelty.

Now, let me make one more sweeping statement.

Paul Middlebrooks
Oh, make four more. Come on.

[music]

Paul Middlebrooks
This is "Brain Inspired", powered by *The Transmitter*. Xiao-Jing Wang is the director of the Swartz Center for Theoretical Neuroscience at NYU, New York University. He sent me some bio notes. I don't usually do biographical stuff, but he sent me some bio notes, and they're interesting. Here you go.

Xiao-Jing was born and grew up in China, spent eight years in Belgium studying theoretical physics, areas like nonlinear dynamical systems and deterministic chaos. As he says it, he arrived from Brussels to California as a post-doc, and then one day, switched from French to English language, from European to American culture, and from physics to neuroscience.

I know Xiao-Jing as a legend in non-human primate neurophysiology and modeling theoretical neuroscience, and having paved the way for the rest of us to study brain activity related to cognitive functions like working memory and decision making. He has just released his new textbook called *Theoretical Neuroscience: Understanding Cognition,* which covers things that textbooks cover.

In this case, he covers the history and the current research on modeling cognitive functions, from the very simple to the very cognitive. The book is also somewhat philosophical arguing that we need to update our approach in neuroscience to explaining how brains function, essentially to go beyond Marr's levels, the famous David Marr's levels, which we discuss in the episode, and instead, to enter into what Xiao-Jing refers to as a cross-level mechanistic explanatory pursuit, which again, we discuss that as well.

I just learned he even sites my own PhD research, studying metacognition in non-human primates, and also my post-doctoral research studying response inhibition. Right there, right in the citations, so you know it's a great and worthy book. You can learn more about Xiao-Jing and the book in the show notes. It was fun having one of my heroes on the podcast, and I hope you enjoy our discussion.

[transition]

Xiao-Jing I have the book, *Theoretical Neuroscience: Understanding Cognition*. That's a slightly ambitious subtitle. There's a lot in this book. Were you asked to write this book, or did you decide it's time to write this book?

Xiao-Jing Wang
Hi, Paul. It's great to be on your podcast. I decided to write this book actually for some time now. I guess probably, it's a good time for a new textbook in the field. As you know, neuroscience has developed tremendously over the last, maybe, two decades or so in terms of two developments. You have new technological matters, and as a result, you get a lot of big data of all kinds.

There's a recognition that perhaps we need theory and modeling that go hand-in-hand with experimentation to help accelerate discoveries. At the same time, conditional science as a field has matured. Maybe this is a good time for a new textbook.

Paul Middlebrooks
If you don't mind me saying, at least from my matriculation in computation neuroscience, you already were a legend in the field when I began, and so, I suppose you're still a legend in the field. This is a single-author book. Most textbooks are multiple authors, so that stands out in that respect.

I want to understand, you just alluded to how things have progressed, and now, we're in a regime of big data. Back when I was beginning, we were recording single neurons. You would lower a single electrode and get one or two, or maybe three neurons for a recording session while your animal was performing a task. That's sort of what I knew you from as well, studying working memory.

You were really early on in the computational neuroscience of working memory and decision making. I would just like to ask you to reflect a little bit about what's different today, besides just that it's bigger, and more. What's different today than from earlier on in your career?

Xiao-Jing Wang
Well, as you said, back then, we're talking about maybe 30 years ago, 25 years ago, most of the recordings from behaving animals, writing an interesting task, were limited to one cell at a time. Perhaps, I'm sure you agree, you are from the field, that it's probably fair to say that the studies of cognitive functions like working memory and decision making roughly started around the turn of the century.

Before that, I would say the most efforts in our field were about sensory coding, sensory information processing, or movement, like central part and generators in insects, for example. What's happening in between in a more flexible delivery process, at the single cell level, really, started around before or after the turn of the century.

You think about selective attention by the work of Bob Desimone, for example. Action selection by Geoff Charles, perceptual decision making by Bill Newsome, Mike Shadlen, and economical decision making by people like Tao Li and Paul Glimcher. They already happened roughly around that time.

Paul Middlebrooks
That's true.

Xiao-Jing Wang
That frankly, in my mind, has not really been covered in a systematical way, in a pathological way in the textbook, so that's what I try to do in this book.

Paul Middlebrooks
I'm not sure what percentage of textbooks have quoted Dr. Seuss, but the last chapter in your book has a Dr. Seuss quote. "With your head full of brains, and your shoes full of feet, you're too smart to go down in a not-so-good street." Man, I'm not a good Dr. Seuss reader, but you having said that, you have this table in the last chapter that lists everything that you were just talking about in a sort of-- like a hierarchical order of all the tasks that have been used, many of the tasks that have been used to study what you're alluding to here with the cognition. It's a very thorough and modern, obviously, book. You felt like they just hadn't been put together, our modern understanding of cognitive type tasks, and theoretical notions?

Xiao-Jing Wang
Yes. For computational neuroscience, that, I think, is something really important. You just mentioned tasks. Actually, being trained as a physicist, for me, it's also really a big change. When I started in neuroscience, even when I tried to build a model for working memory, I didn't really know much about behavior psychology, so I didn't really pay too much attention on behavior performance.

Gradually, I think that's true for many people from physics or mathematics. You say, "Oh, this is a neural network model," which is some kind of physical dynamical system. I'm interested in scale-free dynamics. I'm interested in oscillations. That's a dynamical phenomenon, without thinking too much about tasks.

That is, for me, a really very interesting, personally, change and gradually how you appreciate the tradition in psychology, and how people design tasks to really get the specific questions about how the brain works.

I do think, as you know, computational science is very cross disciplinary. Each time I teach a course, there are always some students from physics or mathematics, who didn't know anything about the brain or, of course, also, a few experimentalists trained in biology with a relatively weak math background. I thought it's good to write, inspire in a way, in teaching how important it is really to think about behavior, about--

Paul Middlebrooks
Yes. You highlight that a lot in the book. Do these students come in from physics these days, and have the same appreciation that you did, where ignorance of the behavior or not a focus on the behavior and the psychological aspects of these things, and you try to hammer that into them? [chuckles]

Xiao-Jing Wang
It depends. These days, maybe more students are interested in our field because of AI, because of machine learning. I would say it's probably still the case. It's still common, at least. People from quantitative fields may not have had exposure to this richness of behavior and brain function.

Paul Middlebrooks
People that are interested in AI might be more interested in benchmarks, which is different than behavior. I'm not sure if that's a nuanced difference. What do you think about that?

Xiao-Jing Wang
Yes. Sure. That is about performance. Right?

Paul Middlebrooks
Performance. That's a very general term but yes. Okay.

Xiao-Jing Wang
Performance of what. I guess different people are interested in different things. The other thing I found interesting, perhaps we can discuss in this dialogue, namely, in some sense, we're a bit schizophrenic when we talk about the brain. I'll tell you what I mean.

Paul Middlebrooks
Please.

Xiao-Jing Wang
Of course, we all know that different mental processes, different functions, somehow depend on different parts of the brain. We know that for vision, we know that for audition, we know that for motor behavior. Very often, when we talk about "neuron computation," generically, for example, when we talk about the relationship between the brain and AI, we kind of throw that away. We don't really emphasize functional specialization.

Paul Middlebrooks
Modularity?

Xiao-Jing Wang
Like modularity. For example, you can maybe do a survey for the general public, or even for AI researchers. How much do they know about the differences between the ventral stream of the visual system that were the original inspiration of the convolutional neural networks, and the dorsal stream. What are the differences and why? How to explain the differences? I bet very few people really, really think hard on that.

That's what I mean by-- [chuckles] we know that they're different, but still, perhaps, it's good to really discuss, diving into more depths, how different parts of the brain really can subserve different functions. That's the question of modularity. Does that make sense?

Paul Middlebrooks
Go ahead.

Xiao-Jing Wang
Does that make sense? Yes?

Paul Middlebrooks
Yes. I immediately want to ask you now about working memory. I think of you as being famous for working memory, and then revisiting a lot of your historical work, you're famous for a lot of things. One of the things you talk about is that there are lots of circuits, lots of areas that have working memory types of-- what you think of as working memory types of activity within the circuitry throughout the brain. This speaks to the modularity that you were just discussing. Is that a good example, do you think, to illustrate the modularity?

Xiao-Jing Wang
Yes. Definitely. As you know, in neuroscience, usually, we'll get a hint from damaged patients, patients with damaged brain. Then, later on, you could do lesions, careful lesion studies. Those studies point to the role of, say, the prefrontal cortex in working memory. That led people like Fuster and Patricia Goldman-Rakic to put the electrode into the prefrontal cortex in behaving animals, performing a working memory task.

It's very nice because all you do in delayed response task is not to allow animal to respond right away to a stimulus. You have to hold in mind in working memory, maybe do something about it during a delay period before you can perform a memory-guided response.

Paul Middlebrooks
Right. We should maybe define working memory as a temporary memory, but also information processing within that temporary hold.

Xiao-Jing Wang
Right. The key is that it's not enslaved by environment. It's something internal. When you see some working memory signals, they are not directly driven by external input. It's during a time when you can hold something in your mind. That's really interesting in contrast to, say, primary sensory neurons responding to stimulation.

Paul Middlebrooks
Our entire brain is highly recurrent. These types of signals are basically everywhere. There are some areas where it's more pronounced, I suppose. Speak to that for a moment.

Xiao-Jing Wang
Well, there are multiple parts in addressing this puzzle. I would say it's still really a puzzle. Actually, it's becoming, I would say, one of the really key central questions today. I'll explain to you why, unpack a little bit. These days, in contrast to what we said just earlier, that you could record from one cell at a time.

Today, with the advances of tools like Neuronpixels, people record from thousands of single neurons in multiple brain regions at the same time, from animals doing some task. People then analyze those data in different ways. For example, you can decode what's encoded in neuron activity in different parts of the brain, and they found that you can decode, say, working memory signals in many brain regions.

That's very interesting, so suggesting that working memory is distributed. It also, I would say, creates some confusion. We can ask a bunch of questions. We don't really know the answers to those questions. Why is that when you see some neuron signal in a area, you can ask how does that happen and why?

For example, as you said, areas really interact with each other. If you just look at the connectomic measurements in the primary cortex, for example, about 67% of all possible connections are there. What I mean by this is that if you have n areas. For each area, you have (n - 1), possible long-distance connections. Altogether, you have n * (n - 1) possible connections. Out of that, about 67% of all connections are there. The proportion is even higher in mice. People say 95%, 97%

Paul Middlebrooks
I did not know that. That's interesting. Wait. That means it's more modular at "higher levels" in the taxonomic tree, in non-human primates and human primates? That would make it more modular. Right? It would make the mouse brain more homogenous?

Xiao-Jing Wang
Maybe. Yes. In principle, there are other things, but just look at this connectomic measurements, you'd say. You can go from anywhere to anywhere else by one or two synapses. Is it possible, for example, just take an extreme example, maybe in the end, working memory depends on only a few areas. By virtue of interactions, some other areas have some kind of working memory signals as a result of receivers, as a receivers of projections from the core, for example.

That's a possibility in principle. Then, you can, of course, discuss functionally. Is that a bug? Is that useful? I think the brain would define ways to use signals. It depends on behavioral demands. Those are the questions that people are starting to study, motivated by new data.

Paul Middlebrooks
Right. That is an interesting-- that redundancy. Von Neumann wrote about this as a feature of the brain, redundancy, and how basically if you send a bunch of copies, then you reduce the noise. That's a very short way of saying it. What you were just saying reminded me of how power-efficient brains are, and that it might not be such a cost, actually, if you have that extra stuff outside the core.

Xiao-Jing Wang
Right, but if you have memory signals in V1, how do you use that?

Paul Middlebrooks
Sure. I was going to ask you. Where is working memory then? [chuckles]

Xiao-Jing Wang
By the way, actually, I'm not so sure. Redundancy, certainly, is desirable for several reasons, but I also want to say that what I just told you could be understood a bit in terms of this old notion of equipotentiality by Lashley. Remember, we can-- big and bigger parts--

Paul Middlebrooks
Bigger and bigger parts of the brain. Right. Yes.

Xiao-Jing Wang
Oh, it's more or less doing the same thing. I actually don't believe that. That's actually the puzzle. We tried to study and try to propose some solutions. You take two premises. One is that you have these long-range collections. By the way, that's also very different, in my mind, fundamental difference between neuron circuits and the physical systems. In physical systems, interactions are local.

Paul Middlebrooks
What do you mean physical systems?

Xiao-Jing Wang
Well, like molecules. Interactions are by collision or chemical reactions. They are local interactions. That's very different in the brain circuits where you have these long-range connections that allow you to go to different places. This is number one, number one premise. The second premise is the notion of canonical circuits.

Paul Middlebrooks
Mountcastle. Canonical microcircuitry in the cortex.

Xiao-Jing Wang
Exactly. David Huber, Torsten Wiesel, and later, Kevin Martin, Rodney Douglas. It's a very elegant principle. Basically, the cortex is made of the same stuff. You just have more and more repeats of the same local canonical circuit.

Paul Middlebrooks
If we figure out what one of those little micro circuits are doing, then we have solved the cortex. [chuckles] That's the goal.

Xiao-Jing Wang
Exactly. Also, across species, from rodents to monkeys and humans. That's a very powerful, elegant idea, and people love it. Then, if you take those two premises, it's become even more no trivial. How do you explain modularity? Functional modularity, by the way, people use the term modularity in different ways, people who use graph theory, for example. When you think about the graph, apply to neural networks or even the brain connector.

Paul Middlebrooks
The physical structure, you mean. Right? The physical structure?

Xiao-Jing Wang
Yes. Network science. When you see a graph, actually implicitly, you think about nodes being all the same. What differ is input and outputs.

Paul Middlebrooks
Well, also, there are connectivity differences, small world network, et cetera, but--

Xiao-Jing Wang
Yes. That's input and outputs. Each node has different inputs, different outputs, but the nodes are all the same. In that sense, it's a bit like canonical local circuit idea. Of course, they use the word modules, modularity in different ways. I want to just to be clear.

What I mean by modularity, I mean functional modularity, like working memory is a module dedicated to working memory. It becomes actually really a puzzle. If you assume that different areas are made of the same stuff and they all talk to each other in a dance network, how do you get differentiation of function, really?

Paul Middlebrooks
I know that you have multiple answers to this. It actually doesn't take much. Well, perhaps, you could enlighten us [chuckles] how that would work then.

Xiao-Jing Wang
We don't know the answer. We do try hard over the last years to build this model.

Paul Middlebrooks
In principle, I think that we do know the answer. The answer is that not all cortical columns are the same because they express different proteins, and even these nuanced differences then give rise to different function within the cortical column.

Then, the modularity itself, the connectivity itself, one cortical column is receiving inputs from these 14 different areas, and another cortical column is receiving inputs from 12 of those, and one of these, and so, there is differentiation just with the input-output structure, just as you were saying. I thought that's what you're going to [chuckles] talk about.

Xiao-Jing Wang
Great. That certainly is, some other people say, just maybe the canonical circuit idea is appealing, but not sufficient. Maybe it's actually not true that different parts of the cortex are the same. Of course, we see all kinds of differences. These days, with the new transcriptomic data, connectomic data, all kinds of data point to that direction, so you say, "Oh, it's just not true. There are heterogeneities."

I don't know if you know this. For example, in primates, of course, in the cortex, there are excitatory neurons, pyramidal cells, and inhibitor neurons. In V1, about 15% of neurons are inhibitory, 85% of neurons are excitatory. In the PFC, actually, there are twice as many inhibiting neurons.

Paul Middlebrooks
There is prefrontal cortex, which, if you think of the brain as a feedforward input-output system, the prefrontal cortex is at the top of that hierarchy in the cortex.

Xiao-Jing Wang
Yes. The prefrontal cortex or PFC is the part of the cortex right above your eyes, in the front, and that has been mysterious for a long time. Only in recent decades we start to realize how important it is for cognitive functions and executive control of behavior. Sometimes, it's called a CEO of the brain. I actually want to contrast. Sometimes, PFC with primary sensory area like V1 and just to say the difference is quite marked.

Paul Middlebrooks
Well, that's one big difference is just the number of inhibitory neurons, as you just said, is 15% to-- how high is it in prefrontal cortex?

Xiao-Jing Wang
30%.

Paul Middlebrooks
30%? You hear all the time that excitation-inhibition balance is key to neural functioning, so what does that difference mean functionally between those areas?

Xiao-Jing Wang
I guess people still try to figure it out. We've done some research to try to address this question. In fact, it's something, if you don't mind, I can dive into a bit of a detail. I think it's probably interesting for your audience for several reasons.

We talk about working memory. When we try to build a model for working memory, we were worried about how you can ignore distractors. When you try to hold something in your mind, it's really hard to ignore intruding signals that are not relevant. That can be either external stimulation, it can also be some internal thoughts that you should ignore.

This was 20 years ago. We published the paper in 2004. By chance, at that time, I was talking to an anatomist from Hungary called Tamás Freund. We used to think about inhibiting neurons targeting excitatory neurons, controlling excitatory neurons. He was telling me that they just discovered a group of inhibiting neurons that avoid pyramidal cells. They don't target E-cells. That was a big surprise to him. That was in hippocampus, and then a bit more about some bits of anatomical evidence, and it turns out that those guys that avoid the pyramidal cells target another class of inhibitor neurons.

Paul Middlebrooks
A different class than their own?

Xiao-Jing Wang
A second class of inhibitor neurons, and those second-class of inhibitor neurons target dendrites of pyramidal cells. They're actually controlling input flow to pyramidal cells. If this second type is very active, it would just block gate out input flow to pyramidal cells.

Paul Middlebrooks
Is that for sparsity? Is that for an efficient coding kind of scheme? You mentioned hippocampus and you need sparsity and parts of the hippocampus, but--

Xiao-Jing Wang
We first found that kind of cells that avoid pyramidal cells in hippocampus, but I thought that may be a way for gating, so that gives some names. The one from the discover can be labeled by some biomarker. This states VIP neurons, or Calretinin, as a marker, and that's called that in the neuron targeting the neurons. Okay?

Paul Middlebrooks
Okay. [chuckles]

Xiao-Jing Wang
Then the second type can be labeled by somatostatin and they target dendrites. Again, if they're active, they can just block input to pyramidal cells, but if, for some reason, the first type, the VIP neurons, are active, they would suppress SOM in the neurons, thereby opening the gate.

Paul Middlebrooks
Yes. This is a concept like disinhibition, which is, I still, it's a knot-knot. It's hard to even conceptualize that happens in the basal ganglia circuitry a lot, but-

Xiao-Jing Wang
That's true.

Paul Middlebrooks
-disinhibiting something. For some reason, for my human brain, I have to pause and think about it for a second.

Xiao-Jing Wang
Okay. Right. This one is more about dendrites, about inputs on the pyramidal cells. We built a model with actually three types of inhibiting neurons. Those two, and then a third one, in turn, actually, controls the spiking output of pyramidal cells.

You have the ones that are labeled by SOM that control the inputs, and the other one, PV neurons that control the spiking outputs of pyramidal cells. This motif with three kinds of inhibitor neurons, I don't know how many people really know it, it was first proposed theoretically in the modeling paper and--

Paul Middlebrooks
I'm sorry, which-- so I can link to it in the show notes. This was 2003, you said?

Xiao-Jing Wang
2004.

Paul Middlebrooks
2004? Okay.

Xiao-Jing Wang
Yes. Then, nowadays, with genetic tools, people have established this motif basically universally across the whole cortex, but then, just back to your point, maybe different parts of the cortex have different needs for output control and input control.

Paul Middlebrooks
Oh, okay.

Xiao-Jing Wang
Right?

Paul Middlebrooks
Yes.

Xiao-Jing Wang
Think about the primary visual cortex. Maybe there are a few sources of inputs, but in PFC, there's this huge convergence of inputs on the pyramidal cells in the PFC. Maybe you need more input-controlling inhibitor neurons in PFC than in early sensory areas. That's exactly what you find. The proportion of input-controlling versus output-controlling inhibitor neurons is very different from area to area.

Paul Middlebrooks
It goes from inhibitory neurons, sorry to backtrack a little bit, go from 15% in primary visual cortex, which is the early visual cortical areas, up to 30% in prefrontal cortex. Does it increase linearly along that gradient?

One thing that we know is that there are longer and longer timescales as you go from early sensory areas to prefrontal cortical areas, which tracks, I suppose, with the percentage of inhibitory neurons, but is there a jump in there, or does it just increase along that axis?

Xiao-Jing Wang
Yes. That's a great question. When people say there are a lot of heterogeneities, it's interesting to quantify. Is it just random heterogeneity that kind of high-dimensional distribution, or there's alternatively some systematical change along some low-dimensional axis?

We and others like John Murray have used different kinds of data to address that question, including the inhibitor neuron proportion, and all kinds of things, even transcriptomic data. We found that the answer is the latter. In fact, there are systematical changes of biological properties along certain low-dimensional axes, which we now call microscopical gradient.

By the way, there are two slightly dimmed things here. One is the proportion of inhibitor neurons relative to excitatory neurons. The other one is among the inhibitor neurons, what's the proportion of neurons that control inputs, and what's the proportion of neurons that control the outputs?

Anyway, they all show pretty systematical microscopical gradients, which is nice because you quantify them. You have numbers, so then, you can build a model, a large-scale model, which is, I think, still very new, but promising. Basically, you say, I don't invent connectivity. We just use connectomic measurements to build a multiregional, large-scale, say, macaque cortex.

Then, you introduce gradients of synaptic exudation or inhibition that ensure that you have some graded differences in different cortical areas. Mathematically, it means that you should use the same equations for each local area. That's the canonical part, and then the heterogeneity part is to say, well, even if you use the same equation, the values of parameters may change. It's a disciplined way as it is to try, to build a biologically-constrained model.

Paul Middlebrooks
I'm sorry to take us as an aside, but as you're speaking, I realized you were so early in-- when I think of neuroscience now, I think, "Oh, it's all computational neuroscience." When you began, it was not. There was very little computational neuroscience. You were one of the original people who ushered in theoretical ideas, mathematical models into neuroscience.

Can you reflect just a moment, because do I have it right that now it seems like it's all computational? Do you feel justified? Was it like did you come into neuroscience and think, "Oh, these biological stamp collectors. I'm going to fix this." What was that like back then? Did you feel alone [chuckles] in your endeavors?

Xiao-Jing Wang
Thank you. I guess when I teach a course, I sometimes say that there are always pioneers in any field. I'm thinking about people like Hodgkin and Huxley, who built the Hodgkin-Huxley model, and then applied mathematicians like John Rinzel, Will **[unintelligible 00:39:01]**, Barton **[unintelligible 00:39:04]**, Nancy Kopell, but as a field, perhaps, it really started when data collection came to a point, and then people like David Marr coming from more functional perspective push the idea of using mathematical models to try to understand how the brain works.

Paul Middlebrooks
Yes. I won't ask you about David Marr, again, in a moment, but yes. You said it happened when data collection began?

Xiao-Jing Wang
Well, okay, maybe I'll frame it slightly differently. I guess in my mind, at least, that when the systems neuroscience is developed to some point, in our case, you had Mountcastle, Hubel and Wiesel, et cetera, and then you start to use math. Again, combined maybe with mathematical psychology, you use signal detection theory. You use a point process, so you start to use mathematical tools to analyze.

Paul Middlebrooks
Yes. Very simple. Well, I won't say simple, but very small models? You'd call these small models.

Xiao-Jing Wang
Yes. Then, you use information theory to quantify coding, and then there are people who like Eve Marder try to look at rhythms, underlying motor behavior, start to use dynamical systems theory. I would say those are examples, early examples.

Then, again, think about Hodgkin-Huxley, Roe(ph), and Rinzel, it's more neuron physiology, which is always very quantitative. I came from the background of dynamical systems theory and statistical physics, but just by chance, in a way, I got into the prefrontal cortex very early on. [chuckles]

Paul Middlebrooks
What do you mean by chance? [chuckles]

Xiao-Jing Wang
It's a bit accidental. When I switched to neuroscience, I was in an experimental lab of Walter Freeman at UC Berkeley. I don't know if you know that his father was the infamous person who performed frontal lobotomy.

Paul Middlebrooks
Oh, no. Okay. I did not know that. Infamous. [chuckles]

Xiao-Jing Wang
Infamous. He actually performed thousands of cases, he would drive--

Paul Middlebrooks
Through the ice pick frontal lobotomy? Yes.

Xiao-Jing Wang
He was the one who would drive his-- all the car from state to state and offer it as a treatment. He was the one who performed frontal lobotomy on Rosemary Kennedy, a sister of John Kennedy.

Paul Middlebrooks
He performed that. Walter Freeman was his son?

Xiao-Jing Wang
Yes.

Paul Middlebrooks
Gosh. Do you know how Walter Freeman thought about all that?

Xiao-Jing Wang
I don't remember if we had really that conversation. [chuckles] Anyway, I heard about PFC, and maybe it's important to understand in psychiatry. Although--

Paul Middlebrooks
That's hilarious. You came to PFC because we were lopping off PFC and someone's dad was getting rid of PFC, the thing that's not important, and you thought [chuckles] maybe that's important.

Xiao-Jing Wang
Well, it's true. It's terrible. Rosemary Kennedy, she had a procedure when she was like 23, and that failed miserably, so she could not function afterwards.

Paul Middlebrooks
Oh, yes. I don't know much about her, but I've heard that it was bad, particularly for her.

Xiao-Jing Wang
Yes. She lived in the institution for more than five decades, was a terrible example. I just showed we didn't know anything about what PFC is good for, and that actually, in Pittsburgh.

I actually started my first faculty position in Pittsburgh. I got to know David Lewis, and he told me about PFC in a more scientific way. That's actually how I learned that PFC is important in psychiatry because apparently, all major psychiatric disorders somehow involve abnormal PFC. It was him. He introduced me to Patricia Goldman-Rakic. That's how I got into--

Paul Middlebrooks
I see. Just to pause there. Patricia Goldman-Rakic, this is the way I was taught about working memory also, recorded neural activity in the prefrontal cortex of monkeys performing a working memory task, and the old story was that you would have these single neurons that when the monkey is holding something in its working memory, would have this persistent activity.

You'd have a single neuron that would start becoming active when there was no stimulus, and the monkey is trying to hold something in mind, and it would be active until the monkey was queued to respond to indicate the answer of its working memory content, and that was the story of working memory back then.

Xiao-Jing Wang
Yes. Again, there were others like Foerster and Watanabe in Japan, but Goldman-Rakic, I think, was unique because she really tried to use multiple approaches to get to the mechanistic understanding across levels that she-- not only in her lab.

They did single-unit recording in the working memory task as you described, but also, she did brain imaging. She did in vitro slice and she trained as an anatomist, also did a lot of atomic analysis. I guess that's her really unique contribution, really trying to use multiple approaches to try to understand the mechanisms.

Paul Middlebrooks
Okay. Well, you've just queued me now to go back to David Marr, so let's talk about cross-level mechanisms, which is one of the unique things about your textbook is that it's not just a textbook, it's a calling for ways to approach the brain.

We've talked about David Marr a lot on this podcast, but just to recap, and you mentioned David Marr earlier, his approach became popular in the neurosciences, thinking about how to understand brains, how to explain brain activity in brains. There are three famous Marr's levels. You have your computational level, computational/functional level, which is like what is the task? What is the object of the cognitive activity?

You have the algorithmic/representational level. What are the steps that the brain has to do to get to accomplish that task? Then you have the implementation level, which is just how do the neurons act, how do the neurons carry out the function, to carry out these algorithms, to carry out the task. One of the things that you argue very early on in the first chapter of the book is that we should go beyond Marr's levels now. Why is that?

Xiao-Jing Wang
I guess historically, maybe let me talk a bit about the logical aspect of David Marr first. Historically, back then, in the early '80s when he proposed the three levels. By the way, he's actually initially proposed three levels together with Tommy Poggio.

Paul Middlebrooks
Poggio. I'm going to have Tommy on at some point. Yes.

Xiao-Jing Wang
Yes. Please, ask him. I had a conversation with him, so it's interesting. I guess there are two motivations. One was that back then, neuron biologists didn't really think too much about behavior, frankly. A lot of people tried to understand how single neurons work, how snap transmission work. For that, you can do a lot of in vitro slice studies. Of course, there's no behavior in the slides.

Paul Middlebrooks
That's true. I guess, you could call that kind of stamp collecting but--

Xiao-Jing Wang
Stamp collecting. It's just not about behavior. Then, even when you record in vivo back then, very often, you record from anesthetized animals.

Paul Middlebrooks
You're not asking anything about behavior, you're trying to figure out how the stuff works, which is a very valid and necessary thing.

Xiao-Jing Wang
Absolutely. I guess, David Marr and his friends feel like as a complementary approach, maybe you can study behavior without worrying too much about implementation, so they say, "Okay. Vision is really hard. How do I understand the stereo vision?" Let's first define stereo vision quantitatively. That's the function part. Then, let's try to propose some mathematical algorithm. That's the software part.

Paul Middlebrooks
Yes, how to accomplish it despite whatever the stuff is made of underlying it, whether it's brain material, computer material.

Xiao-Jing Wang
Yes. That's the third part. Hardware. Okay. There's a function, software, and hardware. Actually, by the way, the second motivation was a bit sociological because apparently, from Tommy, back then, people think that all you need to know is molecular biology.

Paul Middlebrooks
Yes. [chuckles] From Tommy, that was his perspective.

Xiao-Jing Wang
There was a bit on that. David Marr said, "We can study behavioral on its own merits." That makes sense. I think over time, we have to read his original writing, actually. Over time, maybe by some people, maybe naively. Somehow, David Marr's three levels as understood or perceived in a bit naively as a unidirectional hierarchy. The most important is function, then software. If you are interested, at the end, you could worry about implementation.

Paul Middlebrooks
That is the common way, but are we misreading that? Is that just folklore now?

Xiao-Jing Wang
I don't think-- Remember that when you described it, you actually used the word just for hardware.

Paul Middlebrooks
Did I?

[laughter]

I think culturally, I've been inundated with that. Yes.

Xiao-Jing Wang
Well, for one thing, I actually like to make an observation. If, say, someone really just cares about function, let's say AI, it's good to notice the workhorse of today's AI systems, it's convolutional neural networks, deep neural networks.

Paul Middlebrooks
I think it's transformers now but--

Xiao-Jing Wang
Yes. Sure.

Paul Middlebrooks
The last generation was convolutional neural networks.

Xiao-Jing Wang
Deep neural networks to start with. Right?

Paul Middlebrooks
Deep nets. Yes.

Xiao-Jing Wang
That's a software thing, but it was initially inspired by what we learned about the hardware of the visual system.

Paul Middlebrooks
That is so underappreciated, at least in the AI world.

Xiao-Jing Wang
It's hardware that we learned from the visual system, that was an inspiration for people like Yann LeCun to develop this kind of architecture.

Paul Middlebrooks
And Fukushima.

Xiao-Jing Wang
Yes. I can give you examples. I would advocate that maybe we are not done yet by learning from neuroscience, and especially, say, the prefrontal cortex or dorsal stream of the system. That's the other thing that I actually, I'm not so sure it's so productive. That is people believe that you can implement the same software with all kinds of different hardwares.

Paul Middlebrooks
The functionalist viewpoint. Yes.

Xiao-Jing Wang
Yes. That, in part, is because of the joint analogy between the brain and the computer. You can say, "Oh, a Turing machine can be implemented by the old-fashioned vacuum tubes or by chips." I'm not so sure if that really is true, in fact.

Paul Middlebrooks
What? That a Turing machine-- Isn't it true in principle?

Xiao-Jing Wang
That is true. No. No. I mean about the brain functions. I don't know to what extent that's really true. I also don't know how fruitful, how productive that view is in brain research, or maybe that part of the joint analogy with the computer is not so interesting, so useful.

Paul Middlebrooks
Did you always think that way, or because you're an early computational neuroscientist, and your bread and butter is thinking of the brain as a computer, in some sense?

Xiao-Jing Wang
More as a dynamical system, actually.

Paul Middlebrooks
Well, that's troubling to know that. [chuckles]

Xiao-Jing Wang
I can tell you a bit more how is that different from a computer, in my view, at least. Anyway, if you do care about how the brain works, I do think that we are in need of a neural framework. That's what I try to describe. Although, actually, briefly, because the lack of space, happy to expand that view in the future.

Namely, we do have enough data, and I think it's fruitful to try to go across levels. Maybe 30 years ago, in David Marr's time, it was not possible. Now, I think it is possible. We have the obligation, really, try to benefit from big data in transcriptome, in connectome, to understand function.

Paul Middlebrooks
That's the dream. Despite whether you go top-down from function to implementation, which is the classic sense in which we writ-large interpret David Marr, or if you go bottom-up. The whole thing with David Marr was that you keep these levels distinct, and there's no need to go across levels. You're arguing that now is the time to do this cross-level mechanistic understanding?

Xiao-Jing Wang
I don't know if you say the word there's no need, but I think it's reasonable for him to say you can study one level on its own merit.

Paul Middlebrooks
Okay. That's another way to put it.

Xiao-Jing Wang
Yes. That's, I think, valuable and important. Indeed, again, maybe over time, people misunderstood, or naively took it in a too simplistic way. It's like a one-directional hierarchy, which I think is not so necessarily fruitful now. It depends on your goal. I'm just saying, "Hey, it is now really doable."

Actually, I'll just give you an example. Ignoring destructors in working memory is a functional thing. I cannot say that, in fact, this disinhibiting motif of three kinds of interneurons is not just implementation, but a specific circuit mechanism that potentially can explain gating. Then, you can go across levels in that sense. At least it's doable, depending on your goal. I think very often it's worth trying this way.

Let me give you another example, if you don't mind, for my own experience. That is about building a model for working memory, so you may, I don't know, wonder how do I get into decision-making, and that's, in a way, also, is by accident of how that happened.

Paul Middlebrooks
Not lobotomies. You didn't get there by lobotomies. [chuckles]

Xiao-Jing Wang
It's basically our struggle. This is my struggle. The idea of working memory signals representation is that it's maintained internally by recurrent interactions. It's not driven by external stimulation, like neurons in V1, for example. Like you and me, I talk to you, you talk to me. If our voice is loud enough, not whispering too loudly, we can keep going, like what we're doing now, without external input.

Suppose that we are neurons, that's the idea of reverberation. If we were neurons, through this recurrent excitation, we can maintain some activity that can maintain working memory, freed from the environment, so to speak. That's a key that we try to test quantitatively.

When we build a model, you have to crank up the strength of excitation in a model. We either didn't get precise activity, so we could not have a working memory circuit, or everything blows up when the excitation is too powerful. You have this runaway excitation. You say, "Oh, well, maybe you should also crank up the inhibition." You balance. Right?

When that happens, now, time and dynamics come in. Back then, people assumed that excitation is very fast. If you take into consideration this little biological detail in the model, that excitation is faster than inhibition, you still cannot fix it because it's like an engineering system.

If you are from an engineering background, if you have very strong positive feedback that's very fast, and you have strong negative feedback that's slower, you can't really stabilize your device. For months, I struggled and struggled. It didn't work. I tried all kinds of things, short-term synaptic facilitation or depression, adaptations, things like that. It didn't work.

Paul Middlebrooks
What do you mean for months? You could just tune the knobs pretty quickly, and then run it and see? [chuckles]

Xiao-Jing Wang
That's the thing that sometimes-- I talked to-- I've been fortunate to have collaborations with many experimentalists like Pat Goldman-Rakic. Sometimes, half-jokingly, to tell my collaborators that modeling takes time. [laughs] It's not like really just turning the knob and then you can--

Paul Middlebrooks
I know. Coming from an experimental background, I was always jealous of people who did models because it seems so fast. I guess it's not as fast as it seems.

Xiao-Jing Wang
Anyway, I ended up saying, "Oh, okay. Maybe you don't have this problem if your acceleration is slow, slower than inhibition. You have this slow, gradual reverberation, and then you have fast inhibition that keeps you in check all the time, quickly, efficiently. That worked. That led to the idea that slow reverberation depends on AMGA receptors. Now, it turns out that it's exactly the slow reverberation that you need to explain physiological observations related to decision making.

There's this famous experiment by Roitman and Shadlen. They found that when an animal is doing a very difficult decision task, there's this gradual ramping activity that is how neurons and neuron population accumulate information in favor of different options.

Paul Middlebrooks
Just to let the less informed listeners know, this is the famous, what's sometimes called the dots task, but it's a random percentage. You can

imagine looking at a screen. It is filled with some dots. A proportion of those dots are moving in one direction. You can vary how easy it is to tell which direction they're moving by various ways to do it. One way to do it is by how many of them are moving in that given direction. Your job is to report which direction this somewhat random collection of dots is moving. This is the random motion coherence task. Sorry to interrupt. I just want to make sure that people are on board.

Xiao-Jing Wang
Exactly. I guess the beauty of it is that you can parametrically change the task difficulty by changing the fraction of dots that move coherently in one of the two directions.

Paul Middlebrooks
It goes from super easy to you can't tell at all. You can vary it at very minor steps. The thing that has been used to explain this in neural activity is this ramping of neural activity toward a threshold, like that you're ramping your evidence, you're accumulating evidence toward a threshold, and there is neural activity where that looks as if it is doing this computation.

Xiao-Jing Wang
Exactly. While you make the task more and more difficult with less and less coherence, you see that at the single cell level, also these days, with neuropixel at the population level, this gradual ramping activity over time. By the way, apparently that's the same kind of algorithm that Alan Turing used to decipher the Enigma code in the Second War.

Paul Middlebrooks
He didn't know about NMDA. [chuckles]

Xiao-Jing Wang
He didn't know. It turns out that it's really frankly a shock that we just took the model designed for working memory from the shelf and applied it to this dots perceptual decision-making task. You can pretty much explain everything observed in that experiment. That's because you have this duality, so you have this gradual ramping by slow reverberation, which is slow transient dynamics. It's not attractor or anything, but at the same time, you also have a winner-take-all leading to a categorical choice.

I thought this is another example, I guess, how you could try to go across levels. In the end, what explains the behavior is this emergent collective population dynamics described by dynamical system theory. You can go down and ask, "What's the cellular receptor mechanism?" On one hand. On the other hand, you can really compare the model performance with monkeys' psychophysics, so it's possible now to do this.

Paul Middlebrooks
I want to ask you about bifurcation eventually. There are inhibitory neurons in primary visual cortex, not as many as in prefrontal cortex. There's high recurrence in primary visual cortex, maybe less so than in prefrontal cortex. What's qualitatively different about-- why wouldn't you have a working memory in primary visual cortex? You have the same sort of thing, but on a faster timescale. Is there a cutoff point where it is released from the sensory activity?

Xiao-Jing Wang
Yes. I guess you are asking about bifurcation. [laughs]

Paul Middlebrooks
I guess I am, yes. [laughs]

Xiao-Jing Wang
Right. By the way, I'm writing a piece for The Transmitter.

Paul Middlebrooks
Oh, you are? Okay.

Xiao-Jing Wang
The title is something like, *The Missing Half of the Dynamical System Theory.* [laughs]

Paul Middlebrooks
Oh, okay.

Xiao-Jing Wang
The missing half is bifurcation. These days, dynamic system theory is becoming really common and popular in neuroscience.

Paul Middlebrooks
That's where you came from. Did it go away for a while, and it has re-emerged? What is your viewpoint on the [chuckles] popularity of dynamical systems?

Xiao-Jing Wang
I guess it's really driven by data. Basically, again, when you have just one spike trend from a single cell at a time, you probably tend to focus on time series analysis. Now, if you do recording from thousands of neurons at the same time, what do you do?

Paul Middlebrooks
That goes out the window. Yes, what do you do?

Xiao-Jing Wang
One thing pioneered by people, like actually Gilles Laurent with small species and then Krishna Shenoy and his collaborators. Nowadays, you see neuroscience journals full of papers on trajectories in the state space, dimensionality reduction, subspace communication, manifold discovery. That's why. They are really, really important. Again, it's still early days. It's cool to see that.

Paul Middlebrooks
You think it's early days?

Xiao-Jing Wang
I think it's the early days because you're going to see more data. We're not done yet trying to understand manifold, things like that, also because frankly, it's still descriptive. It's the way we look at data. It's not mechanistic understanding.

Paul Middlebrooks
There's some mechanistic understanding that sneaks in there when you're talking about dimensionality, for example. If you have a system and you can't explain it with low dimension, that means that it has to remain in a high-dimensional regime. In some sense, is that not mechanistic?

Xiao-Jing Wang
No. It's really important knowledge. Sure, it's some understanding, I guess, about dimensionality. One example is that you mentioned sparse coding. I think there's some evidence that sparse coding is really desirable in sensory systems, whereas, again, in PFC, we, together with Stefano Fusi, proposed that you actually want to have high-dimensional representations. That's functionally desirable with the use of what we call mixed selectivity. It's certainly some understanding.

Paul Middlebrooks
I'll just say mixed selectivity refers to the idea that neurons can respond to lots of different signals, or their activity is related to-- it's not a single function, essentially. It can have activity related to lots of different functions, so it's selective to a mixed number of things.

Xiao-Jing Wang
Maybe that's really important for flexibility in that circuit. That requires a high dimensionality of representations. I agree with you that certainly is very important. By mechanistic, I guess, I mean circuit mechanism and across levels again. For example, when people say communication between areas are done by subspaces, probably, we would like to know-- that's interesting. That's a discovery by itself, but maybe we want to understand how does that happen.

Paul Middlebrooks
It's also just a slippery notion. I've used the terminology myself enough that I feel comfortable with it, but I don't really know what I'm talking about when I say subspace.

Xiao-Jing Wang
That's a very interesting topic. I can do that some other time, perhaps. Again, let me give you an example if you don't mind to go into a bit technical detail. This again has to do with PFC-dependent behavioral flexibility. Brandon Milner, a psychologist, pioneered this paradigm as a test for normal prefrontal function called Wisconsin Card Sorting Test. You are given a deck of cards, and you're supposed to sort the cards in one of the three ways. On each card, you have certain number of colored shapes.

For example, this card has three green triangles. Another card has, say, four red squares. You are sorting the cards either by color or number, or shape. If the rule currently in play is color, you sort all the red cards in one pile, green or another pile, et cetera. Without being told, the rule can change suddenly. In principle, when the rule changes, you now have to do with the same cards sorting in a different way.

Paul Middlebrooks
Learning the new rule just by observation.

Xiao-Jing Wang
Just by observation and feedback, so you're told that you did it right or wrong. People with schizophrenia, for example, or frontal lobe damage, have real difficulty performing this task, especially switching the rule. They tend to perseverate, just keep following this old rule, even though they are given negative feedback. We build a model for something like this, which in principle requires internally maintain the rule across a long stretch of time, and then switch the rule one warranty.

What we found, just cutting to the conclusion of this modeling study, is that the representation according to different rules correspond to different subspaces of neural population activity. They're more or less orthogonal, so if you follow the color rule, you are working, you're representing, you're processing information along this subspace. If you're following the shape rule, everything is in here in this subspace. Now, we designed the model to test the idea that this standard targeting inhibiting neurons are doing the gating. This is a sound somatostatin neurons.

First, we found this orthogonal subspaces overrepresentation, and then in the model, you can do a lot of things that you like. In the model one, we optogenetically manipulate inactivating.

Paul Middlebrooks
Using air quotes with optogenetics.

Xiao-Jing Wang
It is model simulation, right?

Paul Middlebrooks
Oh, yes.

[laughter]

Xiao-Jing Wang
The model. Then the two subspaces collapse, and the performance is gone.

Paul Middlebrooks
Performance is gone.

Xiao-Jing Wang
That's what I mean by how do you go from manifold description to circuit mechanism. That's a specific prediction. You can go to someone who is training mice to do this kind of task, for example, and really use real optogenetic **[inaudible 01:14:16]**

Paul Middlebrooks
No any quotes.

Xiao-Jing Wang
I hope that illustrates what I mean by going between this description and the circuit mechanism. Now, back to the missing half.

Paul Middlebrooks
Yes, bifurcation.

Xiao-Jing Wang
Shall we? [chuckles]

Paul Middlebrooks
Yes.

Xiao-Jing Wang
Bifurcation may be familiar to many people, but not for everyone, but it should be well known in fact. One example everybody should know, actually, is a single neuron. Think about how you learned about single neuron responses to a current injection described by Hodgkin–Huxley model. If your input current is weak but positive, neuron membrane potential would go up a bit and reaches steady state. That's what you see also in the experiment. In the slides, you disconnect neurons, you just watch one neuron at a time. You inject the current, weak current gives you depolarization of the membrane.

Paul Middlebrooks
Slight depolarization, yes.

Xiao-Jing Wang
In the Hodgkin–Huxley model, that correspond to a steady state. A fixed point, so to speak. Then you just gradually increase the intensity of your current injection. At some point, suddenly, you don't get steady state anymore. Instead, you start to see action potential that repeats at some frequency.

Paul Middlebrooks
Yes, injecting a steady, high enough current into the cell.

Xiao-Jing Wang
That's no longer steady state. That's mathematically described as oscillation, technically called limit cycle in the state space. It's attractor in sense that if you perturb it briefly, it will go back after perturbation to the same trajectory in the state space. It just circle like this in the state space of neuron population of the activity space of a single cell, in this case. All you did really is just gradually increase something, in this case, a current.

Paul Middlebrooks
Let's say linearly, you could do it linearly, right?

Xiao-Jing Wang
Yes, and then that linear gradual quantitative change can lead to a qualitative change of behavior. If you want to describe that rigorously, mathematically, it's called a bifurcation. I'm sure you're very knowledgeable about attractor networks applied to things like high-durational cells, play cells, and, in our case, working memory. They really can be described as emerging collective phenomena by changing something in a modest way. That's the beauty of it. That's something in this case could be the strength of recurrent excitation, back to what we talked about earlier.

All you did compared to V1 and PFC, all you do is say you start with a generic canonical local circuit and just crank up the strength of recurrent excitation, and suddenly you start to see attractive states. That was, by the way, the initial insights. Sorry, I didn't mean to interrupt you.

Paul Middlebrooks
No, that's fine.

Xiao-Jing Wang
You should first finish your sentence.

Paul Middlebrooks
No, no, I was going to ask you about the-- before you said the word emergent, I was going to ask you about the relation between bifurcation and emergence because it's a qualitative change, but you should finish your thought.

Xiao-Jing Wang
I just want to acknowledge that this type of ideas were early on proposed in neuroscience by people like John Harfield, which is, of course, a well-known name, but also Daniel Amit, who did a lot of pioneering this attractor network paradigm. That, I think, is one example where the idea of bifurcation is useful in neuroscience, and not as well-known as it should, I think.

Paul Middlebrooks
Do you think that that's missing from the current zeitgeist of the dynamical systems approach? Do you think it's not appreciated enough? How is bifurcation missing the other half?

Xiao-Jing Wang
I think--

Paul Middlebrooks
I'm sorry to interrupt, but how is it related to a phase change in state space?

Xiao-Jing Wang
It is related to phase change. In some sense, just at some general level, it's been a phase transition in physics. You have water, you increase the temperature to 100 degrees. At that point, suddenly, you see vaporization of- H2O, so it's similar to that. How do you go from quantitative changes to sudden transition? I guess it should be more widely known in neuroscience. We use this idea to understand the emergence of modularity. I call that emergent because it's not something you build in. It's not even built in by, I would say, modules in the sense of graph theory. It's really emergent.

Paul Middlebrooks
It's bottom-up.

Xiao-Jing Wang
Through dynamics. No, it's through dynamics. Connectome, by the way, is really important and really exciting to see all this new database coming out. A beautiful example is connectome-based model of the navigation system in Drosophila flies, but it's not enough to explain dynamics and function. One example I like to give in the book is to start with two neurons that connect to each other through mutual inhibition. What do you get? You get, say, half-center oscillator, one is active, the other one is not active, and then they switch like this.

That may be a motif for central pattern generator, but surprisingly, actually, under some conditions about dynamics of synaptic interactions, this mutual inhibition actually can produce perfect synchronized oscillations in the system. You need dynamics to explain really the function. I guess this bifurcation space, I think, is, in my mind, a way to say, "The idea of bifurcation is useful if we want to understand the emergence of functional modularity in a large-scale multi-regional system. Let me give you another example just to make it maybe really clear what we mean by functional modularity.

We talk about decision-making, like the random dots task. Imagine that in a difficult example, in a difficult trial, the dots are really random, the evidence is very weak, it is in favor of A, leftward motion.

Paul Middlebrooks
Barely in favor of leftward motion.

Xiao-Jing Wang
Yes, but your judgment is rightward motion. You say, "I think it's rightward." Now, what happens in your brain? You'd say the retina encodes faithfully the physical stimulation, and this retina should have more evidence in favor of leftward motion. Maybe it's the case with V1. Maybe it's the case for MT, which is a specialized visual system for motion information processing.

Paul Middlebrooks
Along that dorsal stream that you spoke of.

Xiao-Jing Wang
Then where suddenly, you see the signal about subjective decision. There are certain areas that are really coding the physical stimulation, A, A, A, and then there are some other areas that say, no, no, no, "I think it's--" Maybe it doesn't know A. It says it's B. That really is responsible for your subjective choice.

Paul Middlebrooks
Then, eventually, you get down to the motor neurons that are inactuating the action. Then, at some point in between those, it flips.

Xiao-Jing Wang
I would say that's a question of functional modularity about subjective choice, let's say. Let's define a functional modularity responsible for subjective choice. Then, repeating whatever I said earlier, you have canonical circuit, you have all these fancy interactions. You can go anywhere by one or two synapses, so how do you get this emergence of modularity dedicated to subjective choice? We think that this bifurcation in space could be a way to explain that.

Paul Middlebrooks
It makes it sound like bifurcation is a very sensitive thing. You mentioned Eve Marder earlier and her work with rhythmic gustatory patterns in things like lobsters. One of the take-home messages of her work is that there is a lot of different ways to get the same result. There's a lot of degeneracy. There's a lot of multiple realizability, which I'm trying to think about how that sits with the idea of bifurcation, which, what you're saying is, bifurcation is a super valuable thing in your system to be able to flip the state space, to change phases essentially. Her work shows that you can be in the same phase under lots of different regimes, even if you try to bifurcate it. How should we think about that?

Xiao-Jing Wang
No, that's interesting. It's a bit, I guess, more related to the question of redundancy. You want some function to be robust in spite of changes of environment, changes in the face of perturbations. Actually, some people took Eve's findings in terms of multiple realizability. It's the same function that can be realized in different ways. That turns out to be a bit too simplistic. Her own work later showed that, in fact, the crabs and lobsters live in different environments. The temperature changes because of climate change. She actually had a paper on climate change and how that---

Paul Middlebrooks
She's been on my podcast, and she's passionate about that. She observes it in the ocean, where she lives.

Xiao-Jing Wang
Yes, exactly. That's about the usefulness of redundancy, but here, if I put it in a radical way, in a way, but actually, if there's a kernel of truth, you can think about bifurcations as a mathematical machinery to create functional novelty. Functional capabilities. How do we really explain functional novelties? Of course, biological evolution is the ultimate answer, but if you want to understand, if you are given the same stuff, so to speak, canonical local circuits, how can you really explain different functional capabilities in different parts of the brain? In a way, maybe all you need is quantitative differences of some biology.

Paul Middlebrooks
Of the same stuff. More is different.

Xiao-Jing Wang
Exactly, but different, more.

Paul Middlebrooks
Slightly different.

Xiao-Jing Wang
Slightly different, exactly. That's true if we can, as a field, not just my lab, can push this idea to see if it really is an interesting idea. Maybe we can

have a general way to try to understand different kinds of functions. Let me just say one more thing about the robustness. The tricky thing, interesting about bifurcation in space, is that it's actually not requiring fine-tuning. Just as a contrast, when you go back to single neuron example, you do change your current injection density in a careful way until you see this bifurcation. If you're on one side or the other, you miss it.

In that sense, you need fine-tuning to really get to the transition point. Somehow, the experimenter has to change something by hand. Now, bifurcation in space is something that happens somewhere in the cortical tissue.

Paul Middlebrooks
Is that necessarily true? Does it happen somewhere because it could be just distributed? You think it happens in this much of cortex or in this much of cortex? Maybe it's a moot point.

Xiao-Jing Wang
No, it's not. I think it's still not well understood. By the way, right now it's still a theoretical proposal. That's it. We have to come down to specific predictions that should be testable experimentally. If you think about code access as two-dimensional spatial system, what would bifurcation transition be located? I don't know. What's the contour of the transition, if that's what you're asking?

Paul Middlebrooks
Yes.

Xiao-Jing Wang
The key thing is that somehow is localized because it has to be able to separate, say, those areas that are engaged in working memory from those that are not. I guess alternatively, you'd say everything's everywhere, so working memory is everywhere. I'm actually taking the opposing view. I say you do have a module for working memory. It's not everywhere the same. If you take that view, at least, and see how far that's compatible, that can explain data. Then there's some localization space that separates this module, that defines this module. Let me just say one more thing because you ask about fine-tuning.

Paul Middlebrooks
Yes.

Xiao-Jing Wang
That transition in space is very robust. When you change any parameter in your model. For example, I would say in the real brain system, maybe exactly where that transition occurs may be shifted. The phenomena, this transition itself, will not require any fine-tuning of parameters. Does that make sense?

Paul Middlebrooks
Yes. This goes to how you define an area and what you're saying is, the borders can shift, and you don't need to be so precise with whether it's 100 neurons or 101 neurons that are doing the task, implementing the function.

Xiao-Jing Wang
Maybe even that level of areas, I don't know. Depending on behavior demand, for example, you may want certain areas to be engaged in working memory, in some other tasks you may not. That boundary can shift according to behavior demand in principle. I'm doing pure speculation now.

Paul Middlebrooks
What you're saying is, by boundary, you mean different modules active together, like different subsets of areas active and coactive.

Xiao-Jing Wang
Exactly. That's certainly is the case in model. In that sense, we still wrap our mind around it because bifurcation space is so new that, I think, we need to do more work to really understand it. We're talking to experimentalists trying to test some specific prediction from that.

Paul Middlebrooks
When you say it's early days in dynamical systems usage to explain brains, you mean because we're halfway there, because we need to explain bifurcations and study them?

Xiao-Jing Wang
Yes. Again, I think bifurcation should be better known and maybe useful for answering certain questions. If we try to be specific, I'm really trying to stay away from sweeping statements. Again, I already made a sweeping statement. Maybe this is one way to create a novel functionality. That's pretty sweeping.

Paul Middlebrooks
That is pretty sweeping. [laughs]

Xiao-Jing Wang
Maybe it's worth also mentioning that why maybe we should stay away from bifurcations. Brai is so complicated, we should embrace different perspectives. It's important for us to have many different approaches to understand how the brain works. One approach is control theory. If you really think about the brain as a machine, from an engineering perspective, like you do motor control, you really want to stay away from bifurcation. That always implies some instability, something that's unpredictable. For certain things, we probably rightly try to stay away from instability and bifurcation.

Paul Middlebrooks
Oh, that's interesting. The cybernetic view and maybe cybernetics is coming back too, but maybe for your sensory and motor systems, you really want to be cybernetic. We really want to be non-bifurcational.

Xiao-Jing Wang
Exactly.

Paul Middlebrooks
For everything else cognitive, everything that we value as our high human cognitive abilities, maybe bifurcations then are actually beneficial. Is that another sweeping statement that you're going to make?

Xiao-Jing Wang
It's certainly one way to try to understand.

Paul Middlebrooks
That's interesting.

Xiao-Jing Wang
Now, let me make one more sweeping statement.

Paul Middlebrooks
Oh, make four more. Come on.

[laughter]

Xiao-Jing Wang
One more sweeping statement is that, how do you explain psychiatric disorders? Give me an example. In spite of enormous efforts in neuroscience and in clinical research-- by the way, there's a new book that just came up by Nicole Rust. You should get her on your podcast.

Paul Middlebrooks
It'll be released next Wednesday, her episode that I recorded recently.

Xiao-Jing Wang
You did?

Paul Middlebrooks
Yes.

Xiao-Jing Wang
Her book is rigid about how basic research in neuroscience should really better meet the challenge of mental health.

Paul Middlebrooks
She embraces dynamical systems and complexity in the book.

Xiao-Jing Wang
Did she mention bifurcation?

Paul Middlebrooks
Did she? In the book? I don't believe she did.

Xiao-Jing Wang
Here's the thing. We can ask the question. If you really compare, "normal" subjects and people afflicted by some disorder like schizophrenia, are you expecting quantitative differences in biological abnormalities, or are you expecting qualitatively totally different kind of things?

Paul Middlebrooks
Doesn't it depend on the disorder? There are gradients and lots of disorders.

Xiao-Jing Wang
Maybe. Yes. At least it's a possibility. We're in the realm of speculation here. There's a possibility that all you need is quantitative differences in biology to explain qualitative behavior. Maybe there's not enough control, then people become impulsive. Things like that. That could be also potentially at least a useful thing to think about. There's this nascent field called computational psychiatry. Maybe that's one angle at least to think about.

Paul Middlebrooks
That's another sweeping statement. I like that you cautioned against-- the way Claude Shannon cautioned against using information theory to apply it to everything. Right after he invented information theory, everyone was applying it to everything. He said, "No, no, this is a very specific thing, it shouldn't be applied to everything." You made a few sweeping statements about bifurcation, but you also caution against it being the solution to everything, essentially.

Xiao-Jing Wang
Of course. Definitely. Again, different perspectives, different angles. I think as part of the theory of dynamical systems. By the way, just to mention a detail, you will never get bifurcation if your system is linear.

Paul Middlebrooks
It has to be--

Xiao-Jing Wang
It has to be nonlinear dynamical system. It's part of the theory of nonlinear dynamic systems perspective that I just feel like I was mentioning that bifurcation needs to be more widely known and appreciated.

Paul Middlebrooks
I was just at a conference, and we were talking about human cognition and how we can scale brain data to better understand human-specific cognition. I was asking someone, "What's beyond our current brain evolution? If you expanded it further, what would that do?" I'm not asking you that. What I'm asking you is, let's say, you come from the dynamical systems background, you've seen it now flourishing and very embraced. Let's say we add bifurcation, and that explains a lot. I've been swimming in this manifold, everything's a manifold, everything's subspaces. Now I'm thinking in the same way of information theory, and we were just talking about bifurcation. It seems like everything's a manifold, and that's not right. What would be beyond dynamical systems thinking? [chuckles] I guess we don't have it yet.

Xiao-Jing Wang
It's a good question. I don't know the answer to it. I think we should be humble and be modest. The truth is that we don't know many things about the brain, about the PFC. PFC used to be called a riddle in psychology. I think it's still mysterious. By the way, I tend to think that the book that I wrote really covers very elemental cognitive building blocks. Even the list of 26 tasks I summarized in the last chapter, really are the ones that could be studied with non-human animals without language, in relatively simple setting. What's really exciting is that most of them actually cannot be studied in mice even, not just primates.

Paul Middlebrooks
Oh, okay.

Xiao-Jing Wang
That's great. The question is how far we can go in that direction. I hope that at least what we learned are going to be useful to understand more and more complicated mental processes and even fluid intelligence. What is fluid intelligence?

Paul Middlebrooks
Raven's Matrices? [chuckles]

Xiao-Jing Wang
Raven's Matrices, for example. I guess one thing that I'm becoming more and more interested in is compositionology. That's the idea that you learn some primitives, some building blocks. Then, you learn some rules, some grammar, and then, depending on what problem you try to solve, you flexibly combine different building blocks according to some syntax to create arbitrary, complicated sequence of things. I don't know if you know this, it's interesting. People who have frontal damage had difficulty to cook a meal, because when you cook a meal, you have to be creative.

Paul Middlebrooks
Compositional, yes.

Xiao-Jing Wang
Right. You can plan, and then you organize your task into subtasks, and you have a goal in your mind all the time, you go through steps, subtask, and then go back to the task or to the second sub-task, et cetera, et cetera. Very often things don't happen the way you planned, then you have to

come up with new ideas to solve new problems until you reach the goal. People who have frontal lobe damage couldn't do this task. If we could really understand how the brain does it, I think that would be amazing.

Paul Middlebrooks
You mentioned mouse earlier, and a lot of the tasks that you list in the last chapter, a lot of them can be assessed using a mouse model. I'm currently working with mouse, but I come from a non-human primate background, and the reason why I was in a non-human primate lab is because I wanted to study something as close as I could to subjective experience or consciousness. Then, I realized [chuckles] it's a fool's errand, basically, but that's what my PhD is in.

Mice have become more popular again in terms of, "Oh, mice can do these cognitive tasks," but how far do you think we can go using rodents in general to study, understand, and explain these higher cognitive functions that we want to know about? Is decision-making in a mouse the same bifurcation as it is in a human brain, for example? What's your viewpoint on using mice for these cognitive studies?

Xiao-Jing Wang
See, I try to avoid subjective mind questions.

Paul Middlebrooks
Yes, because you're bright and I'm not. [chuckles]

Xiao-Jing Wang
Let me actually say something about it. I don't expect mice to cook a meal for us, so that's out of the question. I do think that a certain mental subjective experience can be studied with mice. You asked about decision-making. The simplest decision-making is actually detection. If you show a visual stimulus, and all you need to do is to say, "I see it," or "I don't see." The trick is that you changed the contrast of your stimulus. This is psychometric function as function of the contrast. When contrast is low, you don't see it, when contrast is high, it's very easy to see, and it's nonlinear. It's like a sigmoid curve.

If you're right at the decision's threshold with the same physical stimulation, the same photons onto your retina, sometimes you see it, sometimes you don't see it. Now, that's subjective, I guess, awareness of a stimulus. People do this kind of experiment. You had Pieter Roelfsema on your show. He did a very interesting monkey experiment using exactly that paradigm. Sometimes they don't show the stimulus. Most of the time, the animal says, "I don't see it," but occasionally the animal says, "I see it," so that's a false alarm. What they found is that early sensory neurons essentially reflect physical stimulation, hit trial, miss trial, more or less the same response, no response when you don't show the stimulus.

PFC seems to reflect subjective awareness. You see the activity level in the hit trials or in the false alarm. If you agree that's a simple way to look at subjective awareness, I think that can be done with mice. Don't you think?

Paul Middlebrooks
Yes, but their prefrontal cortex-- for example, the gradient of inhibitory neurons from visual cortex up toward their frontal cortex is different. The bifurcations are going to be all different, even though it's made of the same stuff. That's the worry. That we're--

Xiao-Jing Wang
It could be different, in details, but again, there should be a module of areas, somehow responsible for subjective choice.

Paul Middlebrooks
You think if artificial intelligence has a prefrontal cortex, it'll be subjective? [laughs] This is what we'll end on. I know you're hesitant to talk on this, but before we do that, is there anything else that I haven't asked you that you want to share?

Xiao-Jing Wang
No, I think we covered quite a bit.

Paul Middlebrooks
Perhaps we'll end on this, if you're willing.

Xiao-Jing Wang
Frankly, I have not thought about the question of awareness in machines.

Paul Middlebrooks
That was a joke, but you wrote to me, "AI needs a prefrontal cortex." Why would that be?

Xiao-Jing Wang
That's more about flexibility and fluid intelligence. For example, machines are not so great at multitasking. These days, if you try to train, say, a robot to do more than one task or software, usually what you do is to just add different cost functions, you add them up, for different tasks, one for

each. That doesn't work. PFC is actually crucial for multitasking. If we learn about how PFC does it, perhaps there are some new insights that we can translate through computational modeling to build smart machines capable of doing multiple things.

I would argue, actually, thinking in a way, intelligence, solving new problems, in a way, it's a bit similar to complicated sequences of events, except that they are not motor acts, they are internal events in our mind. If we really understand how this flexible generation of sequences with recursiveness, with compositionality, I think we can go a long way in thinking about smart machines.

Paul Middlebrooks
Okay, Xiao-Jing, it is a joyous and also deep historical and modern view and recounting and perspective on theoretical neuroscience, and it's an excellent resource. It has it all. It reminds me how long I've been in this field and how little I still know, for one thing.

Xiao-Jing Wang
You're being modest.

Paul Middlebrooks
[laughs] Anyway, great to see you again. Thank you for coming on. It's been a pleasure having you on.

Xiao-Jing Wang
It was a real pleasure. Thank you.

[music]

Paul Middlebrooks
“Brain Inspired” is powered by *The Transmitter*, an online publication that aims to deliver useful information, insights, and tools to build bridges across neuroscience and advanced research. Visit thetransmitter.org to explore the latest neuroscience news and perspectives written by journalists and scientists. If you value “Brain Inspired”, support it through Patreon to access full-length episodes, join our Discord community, and even influence who I invite to the podcast. Go to braininspired.co to learn more. The music you hear is a little slow, jazzy blues performed by my friend Kyle Donovan. Thank you for your support. See you next time.

[music]
