

## Minding the Brain: Unraveling the Mystery of Consciousness

<https://mindmatters.ai/podcast/ep260>

Pat Flynn:

Hello everybody, and welcome back to the podcast. I'm your host Pat Flynn, and today we are continuing our fascinating and extended discussion around this wonderful new volume called *Minding the Brain: Models of the Mind, Information, and Empirical Science*. I have the three editors with me today, Angus Menuge, Brian Krouse, and Robert Marks. And boy, we have covered a lot of territory so far. So if you have not heard the first and second part of this episode, I am going to strongly encourage you to listen to those. We covered many of the difficult, if not insuperable, problems facing physicalist models of the mind. We began to explore alternative ranges of options. And we discussed where we believe the best empirical data leans in which direction, which is the broadly dualist direction. So what I want to do in part three is begin to explore in a little bit more detail some of these other options, dualism, and hylomorphism, and idealism.

Obviously we're not going to be able to cover them all, but I'd like to give people a good idea of what is being proposed, and some of the motivations for at least some of them. And then of course I would like to talk a little bit more about artificial intelligence. This is all the rage these days. People have lots of questions about AI, so trying to get some conceptual clarity around AI. And what sort of predictions we think we might be able to make, depending on what position we hold in philosophy of mind. So gentlemen, thanks again for doing this. It's great to be back with you.

Robert J. Marks:

Great.

Brian Krouse:

Yeah, great.

Angus Menuge:

Glad to be here.

Pat Flynn:

So if you wouldn't mind, Dr. Menuge, we'll start again with you. You obviously have identified yourself as something of a dualist, but of course dualism comes in different flavors. I'd be curious just to hear a little bit more about the position that you find most tenable. And some of the problems you think that it might be able to solve, related to contemporary philosophy of mind. And we'll just go from there.

Angus Menuge:

Yeah, I suppose I would describe myself as an Augustinian dualist, if I was asked to be specific about it. And that's because, although I find some of Descartes' arguments persuasive, I do think it's reasonable to say that the soul, or the mind is located where the body is. In the sense that it is always present wherever there is sensation throughout the body. It's just that the mind is located in a different way than physical objects are located. So only part of the body can be present in any of its parts, but it seems that the mind is wholly present anywhere that there is sensation. And I think that the primary reasons for being a dualist in that sort of sense, that we find in introspection, that there is a unity of consciousness, that all of our experiences and thoughts belong to one subject. And that we also need to

assume that unity at, and over a time, in order to account for rationality itself. In particular, the rational activities of a scientist.

So for example, if I have a hypothesis I wish to test, and so then I design an experiment. And I look at the results to see if they show that my prediction is true or false, I have presupposed in that whole activity that I'm the same person who had the hypothesis, and who designed the experiment. Who survives long enough to actually determine the results of that experiment. For that to make sense, it seems to me that we have to have a subject which persists over time, so that with the very same person who made the prediction, who later on finds out whether that prediction is true or false. And in general, I think the reason this is so overlooked, as I argued in my earlier book, *Agents Under Fire*. Is because, of course, scientists spend most of their time focusing on their objects, and don't tend to think very much about what they themselves as scientists are doing. And there seem to be presuppositions that one makes in doing science.

And among those are that you are a conscious subject who persists over time in the work of your scientific discovery. And so I see dualism really as not a foe, but a friend of science. It's there in the background to make sense of the rationality of science itself.

Pat Flynn:

Yeah, that's a really great point Dr. Menuge. And something we weren't able to fully discuss in the previous episodes, but I think it's at least worth bringing up now. Is not just the unity of the subject, which is very difficult if you're coming from a physicalist perspective to describe how we achieve such unity from such a base of disparate, unintentional entities. But also the binding issue, of how that unity is maintained over time. Especially when it seems like a lot of our material parts are being swapped in, and swapped out. Do you think that those issues are something that can be covered by all the alternatives equally well in terms of substance dualism, hylomorphism, and idealism? And is just a problem for physicalism? Or do you think that some of those positions are theoretically more advantageous than others for, call it the binding and the unity problems?

Angus Menuge:

Yeah, so what I would say is that there are some issues where I think that sort of traditional substance dualism, hylomorphism, and idealism can do an equally good job of accounting for them. And then there are some other issues where there might be a reason to prefer one option to another. So for example, when I look at hylomorphism, I think it does a really good job of explaining this boundary between those aspects of our cognition that can be easily manipulated because they have to do with the sensory motor system. And on the other hand, the apparent independence of abstract thought and free will. Going back to Aristotle, we recognized that in reason we contact universals, which we have never experienced. And therefore that reason must in some sense enjoy a degree of autonomy from our physical senses and brain. And he believed that individuals have the power of self-movement and free will, and recognize that that as well was not something that simply emerges from matter.

So I think that that position does very well there. A little bit of a challenge to the hylomorphic position though, would actually be the near-death experiences. Because in that case it's difficult to understand how people are having, what appear to be sensory qualia, which are not using their physical senses. Now that doesn't mean the hylomorphists could not come up with an explanation of it, but that's the kind of challenge that one has. And likewise with idealism, I think their strongest suit is to point out that maybe all of us are operating with sort of out-of-date physics. So their big complaint is that it's not just a physicalists, but even some dualists who tend to think of physical matter in terms of easily locatable stuff.

But when we get to the quantum realm, and you have all these complex superpositions of states that we can't determine the position and the momentum of a particle at the same time, and all these kinds of problems. We begin to lose a clear sense that we know what a material objects is. And idealists capitalize on this problem to suggest that maybe what's really going on is fundamentally transfers of information between minds, and that the old idea of a physical object is actually redundant. That's kind of a very exciting idea. I'm not a proponent of it myself, but I'm very eager for it to be given a fair hearing.

Robert J. Marks:

In fact, one of the chapters in the book by Bill Dembski, I think deals with this. If I get things right, Dembski describes informational realism as the belief that the defining characteristic of reality is the ability to exchange information.

Pat Flynn:

Yes. Yeah.

Robert J. Marks:

So that's his chapter, how informational realism dissolves the mind-body problem.

Pat Flynn:

Yeah. One thing I love about this book is how you are giving a platform for these different positions to get a strong initial hearing, so people can go in and read the different accounts. And see what ultimately convinces them, or starts to lean them in a certain direction. Real quick again about the binding problem, I just want to come back to this, is I think it's pretty obvious there's a distinction between us and the modes of us. So for example, I have my thoughts, I have my emotions, but what binds all those together? What are they ultimately grounded in through time? And the answer is me. It's me, and we need to theoretically make sense of that. And that's where I think something like hylomorphism does a really nice job with the principles of matter and form. And if people want a further explanation of that, I'll encourage them to see my interview with Dr. Jim Madden who contributed the chapter to this volume on hylomorphism for some of the more details and background.

But I agree with you, Angus, near-death experiences are sort of an initial uncomfortable fit within that worldview. And even though I'm quite partial to hylomorphism myself, I've gone back and forth with many other mere hylomorphs if you will. Saying, "Hey, this is something we need to be able to accommodate and think about." I think it can be done, but I agree that it's one of those initial data points that might more comfortably fit, in say a substance dualism or something like that. And it's always good to just be honest about the limits and range of predictions for each theory. It doesn't mean that the theory is immediately falsified, but it might cause you to have to rethink certain aspects or expand it in some ways. And that's always fruitful, isn't it?

Angus Menuge:

Yeah, and I think you're right. I mean the whole goal of the book is to give a really fair expansive hearing of these positions by their best proponents, and without any prejudice. Obviously as editors we have our own opinions, but we didn't let that get in the way of making sure that the reader just gets the best versions of these options from their own proponents.

Pat Flynn:

Yeah. And I think you've done an excellent job of that, and how needed volumes these are. So Bob and Brian, if you wouldn't mind, I'd like to hear now some of your perspectives or their positions that each of you lean toward in philosophy of mind. And if so, would like to hear some of your thoughts and motivations behind that?

Brian Krouse:

Yeah, I guess I could jump in. So, I'm a bit hesitant to pick a favorite to be honest. I think it's kind of like when my kids ask me what my favorite ice cream flavor is. And there's so much to like about the different flavors, I have a hard time picking. But I probably have a leaning towards substance dualism. Or at least I could say what substance dualism I think has going for it, is when we approach it through the philosophical problems, the category differences that we find between the mental and the physical. In very brief terms, it really seems that we need the mental to be rooted somehow in a non-physical substance. And this, it seems to me that substance dualism emphasizes and gets right. Or at least to me it seems right. And some of the objections, like the interaction problem, it seems to me like they have good answers that don't put it down.

Pat Flynn:

Well, if you wouldn't mind, I wanted to get there because that was an aspect that I wanted to cover. So we might as well go into it right now, is of course one of the major issues with dualism. And it's not just towards substance dualism, but this will be directed at hylomorphists, I'll just call them hylomorphs, as well, is this famous interaction problem. So if you wouldn't mind just steering in that direction, just briefly explaining what that is. And some of the solutions on offer for it.

Brian Krouse:

I think Angus could do a far better job at that than I could. So let me actually defer to Angus, maybe. Do you want to try to cover what the interaction problem is, and some of the responses to that?

Angus Menuge:

Yeah. So the main interaction problem is just, how can substances of two fundamentally different kinds interact without a shared medium. And so it arises because it seems that physical causation is unproblematic. If 1 billiard ball runs into another one, we can see how that happens because they share the medium of space. But if a mind is somehow not spatial, then how does it interact with the physical? Well, of course, one thing, as I indicated, it's open to the dualists to in fact say the mind is located in space but in a different way. Another thing though is to go back and ask, how strong an understanding do we really have of causal relations to begin with? David Hume pointed out that in fact all we observe is one event followed by another one, and we never actually observe a necessary connection. That's something that we assume is there, but that's a metaphysical notion, not something that we really derive from our experience.

But that said, I think that it's fair for physicalists to challenge dualists to come up with some sort of explanation of what's known as the pairing problem. And that's the, why is it that one particular mind is paired with one body rather than another. And so why is it that when I want to raise my arm, my arm goes up and not Robert Marx's arm. And why is it that when I stub my toe, I feel in pain and not Robert Marx, and so on.

Robert J. Marks:

Ouch. Okay.

Angus Menuge:

Okay. Yeah, well played. All right. So what I've tried to develop in my own thinking about this is an account of the flow of information. And I think this is promising, because information is something which we're aware of existing in both mental and physical formats. Every day we have ideas and we say we write them down. Which means that we make physical marks, which with certain conventions we can interpret as words that express the thought that we had.

Every day we read books, and after looking at those physical marks, we have ideas in our mind. It seems that memory works a bit that way, that when we remember things that are ingrained, stored in our brain. And it seems that when we decide to move our limbs, somehow our volition activates a motor control program that moves our body. So I think the way to think of this is that we need to have an intermediary that is common between the mental and physical realms. And I argue that some form of information is just the right thing to do that, and that may help us to illuminate it. So I'm not one of those philosophers whose content to sweep the mind-body problem under the rug. It can be framed as a serious problem that needs a serious answer, but I'm not convinced that the dualists has got no responses available.

Pat Flynn:

Okay. So obviously dualists have thought about the interaction problem, and there are actually many different proposed solutions out there. We've heard just one plausible model. I know obviously thinkers like Richard Swinburne have their own. So I think again, it's important that people actually look into this. And they just don't take this comment of objection as by any means decisive. This is something that very smart people have thought about, and I think come up with some very clever responses. So apologies for that diversion, but the interaction problem is so commonly brought up that I'm really glad we actually got to talk about that. So just going back to you, Brian, please continue your line of thought. Is there anything else that you want to say about what inclined you towards substance dualism?

Brian Krouse:

I'll move on to just give a couple of thoughts on how I'm wrestling with hylomorphism. So I find hylomorphism is actually a particularly tricky model to understand, because it's situated within a metaphysical framework, Aristotle's metaphysics that uses a lot of terminology. It sounds very similar to something that it's not. Uses words like form and matter, and they aren't meant to invoke the same concepts that might pop up in my mind, or the mind of someone off the street.

Pat Flynn:

Yeah. Yeah. Right. Just to give an example, when Aristotle's thinking of matter, he's really thinking of a principle of individualization, right?

Brian Krouse:

Yeah. Right.

Pat Flynn:

Not necessarily something that has anything to do with modern physics. It's not completely disassociated, but it's not the way that many modern people think of the term matter. That's a really good point, right?

Brian Krouse:

Yes. And I think I remember reading, I mean your friend Dr. Madden is, I think, just a wonderfully clear exponent of all of this stuff. But I think if I remember right, in a larger book on this subject, he talked about how, in the hylomorphist view, that this matter, which is, as I understand it. It's like pure matter is pure potentiality. It could become anything. And that's not to be equated with the sort of the matter as we think of it as just physical stuff, or particles, or impersonal things. That it could potential have, I guess would you put it nascent mental properties within it as well? I don't know if I've got that right or not, so correct me if I'm wrong, but...

Pat Flynn:

Right. Yeah. Well, if you want to think of something conceptually difficult, try to think of prime matter. And from an Aristotelian perspective, it's very difficult.

But does it relate to modern science? I will, for people who really want to go into this, because this is obviously tall grass metaphysics, which is fun to do. But one of the principle proponents of modern hylomorphism is Dr. David Oderberg. And he's got an article 2021 I think, asking his prime matter energy. So he tests the hypothesis of whether we think we can equate these two. And he doesn't come down definitively on one particular side, but it's worth seeing the plausibility of that hypothesis. To see the compatibility, or at least converges between traditional Aristotelian line of thought and modern physics.

Brian Krouse:

Interesting. Interesting. Well, okay. So as is pretty well known, I think the term "soul" in hylomorphism is used to not only describe the form of humans, but it's describes other living things have a soul too. So plants and animals, just with different capabilities or qualities. And as I understand the emphasis in hylomorphism with what's distinctively human about the human soul, is the rational capability. And when you pull on this thread, you get into the ability to comprehend abstract truths and whatnot. And this is the piece, that when they pull on this thread, that's what they lead to saying there's something immaterial, something that could survive the death of the body in a human. Which I think is very interesting, and it has a lot of overlap with some of the arguments you get from idealists and substance dualists.

But one of the things about the hylomorphists that I've puzzled over a bit, and not quite sure how to connect the dots with what the substance dualists might talk about when they're addressing the issue of qualia. Or some of these other seemingly categorically distinct properties of the mental, like the fact that it seems to imply that you have to have a subject.

Pat Flynn:

Yes.

Brian Krouse:

And then, these things also seem, as we've talked about, have a real poor fit with material. With the physical.

Pat Flynn:

Yes. Right. Yeah.

Brian Krouse:

And so I don't understand why there's not more focus on that in the hylomorphist system. But that opens up some interesting questions, because I guess the sensory aspect actually pops up in animals.

Pat Flynn:

Right. Yeah.

Brian Krouse:

So, I look at my dog, he looks like he's having some qualia in there, I don't know for sure. And so how does this fit in? I'm sure there's a good explanation for that, but that's one place in which I tend to go towards substance dualism explanations.

Pat Flynn:

Yeah. Brian, those are really good considerations, and they obviously extend beyond I think what we can cover in this podcast. But it does bring up a general point, and that is that people often talk across purposes because they haven't clearly defined terms. It's true that many hylomorphs, especially somebody like Thomas Aquinas. For him, the thing that is most relevant to our immaterial aspect is not qualia. He never thought of qualia in the way we think of today, but he would say that that is a material operation. But his understanding of matter is way more expansive than the contemporary understanding. For him, it's rationality. It's a rational aspect. Our ability to gauge, engage in formal thinking, that proves and demonstrates not that we're two substances, but we have an immaterial aspect to one substance, for example. And then if you want to talk about how you survive bodily death, that's a very complicated topic for hylomorphs.

And there's many different models out there, including notions of incomplete substances until the resurrection, and getting your body back in a whole nother topic. But I think the general point is yes, there's too much ambiguity. That sometimes you think that positions might be at odds when actually when you gain conceptual clarity, they're not really so different or as much at odds as you may initially think. At least within the broad camp of dualism. That's why before we started to record Brian, I said, I don't know how to classify hylomorphism of whether or not it's dualism. Because so much of that depends on these conceptual refinements that just often take a lot of time to work through, if that makes sense. Dr. Menuge may have some thoughts on this as well. I'd love to hear them. Right.

Brian Krouse:

It seems like there might be some different opinions within the hylomorphist camp too.

Pat Flynn:

Oh, yeah. There's not just one position in there. Yeah. Dr. Menuge, I love your thoughts on any of all that too. And same with you, Bob.

Angus Menuge:

I would certainly agree with that. As I to hylomorphists, they do seem to have several different positions. Some of them actually almost sound as if they are very close to being materialists. It's just that they have a richer notion of matter. So that doesn't mean that they're materialists in the same sense that Daniel Dennett or Richard Dawkins is. On the other hand, some of them are very close to being substance dualists, those who have sort of a modified thomistic account. Because they so much

emphasize what is distinctive, and independent in the ability to engage in abstract reasoning. And so I think that it's fair to say that there's a spectrum of positions there. And that's another thing about this book. We're trying to open people's mind to the fact that there's lots of options here that need to be thought out, and we really want to open the conversation with this book.

Robert J. Marks:

Yes. Absolutely.

Pat Flynn:

Yeah. That's great. So the reason I'm asking about your personal perspective is really just so we can explore some of these options. So Bob, I'd love to hear from you now. Coming from the computer science perspective, is there any position that you're currently leaning towards? Or you feel is worth exploring more?

Robert J. Marks:

Well, working with these philosophers, I've had to learn a lot of terms. I've had to learn a lot of terminology like hylomorphic, and substance dualism, binding, epiphenomenalism. My gosh, that's a word with seven syllables. People like me don't like words more than three or four syllables. And idealism. What I can comment on is, things which are going on in engineering now, which are attempting to measure something with consciousness for example. How is the computer world trying to establish what consciousness is? This doesn't directly address the theme that we're talking about, but is parallel. I'm aware of six different models that people are trying to do in order to model consciousness. First of all, I'm not even sure there's a good definition of consciousness. We mentioned panpsychism, which is that every material, however small, has an element of individual consciousness. So I'm sitting here with a book in front of me.

My book is a little bit conscious, and we have been the lucky recipients of lots of consciousness. So good for us. There's emergence, which we've also talked about. That's number two. And this has been investigated, I think in terms of computers for numbers of years. Most famously by a guy named Thomas Ray, who came up with this model of Tierra, who was going to try to do emergence. There was a whole field called artificial life. And their entire purpose was to show intelligence through emergence. And basically that field has been abandoned. Well, yeah, the emergence as we talked about, is this idea that if you exercise a horse enough, it'll turn into a tractor. There is integrated information theory. This is a theory which has been developed by Tononi at the University of Wisconsin. All of these, by the way, are materialistic. They're trying to explain things through materialistic models.

And he concludes the consciousness is measured by the degree of complexity of a system. And the more complex you get, the greater that it can be conscious. There's a great chapter, by the way, in the book by Selmer Bringsjord, that explores integrated information theory and the flaws of integrated information theory. So that's a great read. There's the Sims theory. And this is, I don't think something which has made it deep into philosophy. This is the idea that we're all simulations. And if we're a simulations, then we are being controlled by some higher entity by this major computer program, which allows us. Elon Musk is a big proponent of this, by the way. And I don't know if it's an urban myth or not, but he has set out teams to try to prove his theory of the Sims theory. The fact that we are simulations.

Pat Flynn:

Yeah. Well, the simulation hypothesis actually has gained a considerable amount of traction in certain circles to be sure. Right?



Robert J. Marks:

Oh, so I'm so sorry. That really just kicks the can down the road, doesn't it?

Pat Flynn:

Yes, it does. Yeah.

Robert J. Marks:

Because that means that there's greater simulations than we are in. Then the questions are they simulated?

Pat Flynn:

Well, I mean, we brought up Dr. David Chalmers. One of his latest works is reality plus. The notion of virtual worlds and the problems of philosophy. So yeah, it's in philosophy for sure. Right.

Robert J. Marks:

It's in philosophy. I'm sorry to hear that. It reminds me of turtles all the way down. We have simulation theory, who does the simulation? There's another one, and there's a chapter in the book by Bruce Gordon on quantum consciousness. And this looks at a very interesting aspect of the quantum world upon consciousness. Now, Roger Penrose, who wrote the great book, the Emperor's New Mind. He said that we are not algorithmic, and therefore, because we are not algorithmic beings that we cannot be... I'm sorry, that computers can't be creative and can't create qualia, and don't understand what they're going to do. But he looked around and he says, "Well, that's because we're not algorithmic. What in the world can be non-algorithmic?" And the only answer that he came up with is quantum collapse. Quantum collapse is not algorithmic. And so therefore, he said, "The answer must lie here."

And he worked with an anesthesiologist. And they came up with this thing called the Orch-OR theory, but there's been really no research into quantum consciousness that I'm aware of. And those are five different models of consciousness that are being looked at from, I guess a computer science, a scientific point of view. And then the last one is dualism, and the type of substance dualism that we're talking about here. And this is of course one that is embraced by Theus. And it seems to me that a lot of these things are silly. I think the Sims theory is silly. I think that Panpsychism is kind of silly. The Sims theory though, if you look at the Christian religion, are we computer programs that are generated by a higher entity? Well, according to Christianity, we are entities created by a greater God. So does that make us simulations, we're created in his image? I don't know. Maybe some parallels could be made there. So that's my input.

Pat Flynn:

Yeah. Well, that's really great. And I think that opens up a few potential paths to getting towards artificial intelligence. But I want to just hit a few of those points all along the way, especially concerning the simulation hypothesis. Now I'm with you, Bob. I think it's silly, but silly ideas are often seriously entertained by even very intelligent philosophers. So, I think it's worth maybe saying a few more things about. Now, I personally think that there's good metaphysical and epistemological arguments against this thesis, but I'd be interested to hear from any of you, particularly you, Dr. Menuge, if you've thought much about this. Of if there's any particular launch point or basis in philosophy of mind against a thesis like the one that Chalmers is proposing? Or other philosophers. Or even I guess just owners of social media companies who hop onto the simulation hypothesis.

Robert J. Marks:

By the way, Chalmers came out and he said, ChatGPT will probably be 20% sentient. So I don't know where he comes up with this idea of sentience and what model he's using. But we throw these terms around without really defining them. I'm big into definitions, and we really haven't even defined what sentient and consciousness is.

Pat Flynn:

Yeah.

Angus Menuge:

Yeah. I'm inclined to say I'm not really comfortable in saying too much about JAMA-specific hypothesis. But the worry is that because of the scientism in our culture, when we have a difficult intractable phenomenon, we tend to redefine it in terms that a particular science can measure. And as a result, we can spend a lot of time trying to capture some correlated phenomenon, but we're not really understanding the thing itself. And so I don't really see what simulation has to do with the understanding of what it's like to be conscious itself. After all, we can certainly produce now machines that will act as if they are conscious, but that doesn't mean that they are.

And I'm worried that in many fields when they talk about consciousness, actually what they're doing is simply substituting something that they can measure. So they're finding a neurological correlates, or they're finding some computation. And it might be true that there is, for example, in the IAT, that there is this correlation between a certain kind of complexity and consciousness. But that still doesn't really explain what it's like to be conscious. It changes the subject to something that we can more easily measure.

Robert J. Marks:

The thing with sims theory that I find interesting is that indeed, if we do have non-algorithmic properties, if we have properties which cannot be generated by computer programs, then whoever is our creator in this Sims world must have the ability to do a computer program for non-algorithmic things. And that, to me, seems to be a paradox. Unless they have some sort of super-duper intelligence, I don't know.

Pat Flynn:

Yeah. Well, one thing I think just as a general point is that if you're going to kind of entertain seriously this sim hypothesis, you really do need something like a computational theory of mind. And if you have reasons to think that that's inadequate, then you have reasons to sort of reject that particular perspective. And of course, a lot of what we've covered in this discussion would be exactly that, right. So let us now consider, at last, artificial intelligence. Obviously this is something that has sort of exploded on the scene in the past few years, especially with the emergence, if you will, of things like ChatGPT, and Midjourney and what have you.

And what I would like to discuss is not just whether or not we think that ChatGPT is 20% sentient, which is an interesting question, I guess we should cover that as well. But what we think, or what we should predict concerning the future of AI. And what limitations we think it might have, or will encounter from the perspectives that each of us hold. Particularly that each of you hold in philosophy of mind. Do you think we can get some fairly confident predictions along that front? Bob, I know this is an area that you obviously have a lot of relevant expertise. Maybe we start with you on this one.

Robert J. Marks:

Sure. Yeah, I'd be happy to. This is a thesis in my book, *Non-Computable You*. There's also chapters on this in the book that we're talking about, *Minding the Brain*. A wonderful chapter by Winston Ewert, which totally debunks the idea of super-duper intelligence. And there's a number of things which are brick walls that artificial intelligence will go through. One I think is understanding AI will never understand what it is doing. Computer programs, they can add the number 16 and 14, but they don't know what the number 16 and 14 are. The greatest explanation of the lack of understanding goes back to, what I'm sure you guys would know in philosophy, but the philosopher John Searle. And the Searle's Chinese room. He said that he doesn't know Chinese, can't read Chinese, doesn't understand any Chinese, so lock him in a room. And he's in this a room with a number of file cabinets, and somebody slips a note through the door that's written in Chinese.

Searle looks at the note, it's in Chinese. He starts going through a number of file cabinets in the room until he finds a match to the question slipped through the door in Chinese, which he can't read. Well, it turns out on the slip of paper that he finds in the file cabinet that matches the question, which is submitted, is the answer to the question. So he jots down the question and then he goes, and he slips it out the door after refiling the card in the file cabinet so it isn't a misplaced for the next time. Now external to the Chinese room, it sure looks whatever happens inside of there is something that understands Chinese. Although Searle doesn't understand Chinese, he doesn't speak it. Doesn't read it. Yet, it sure looks like it. And I think this is a smack down of the idea that AI will understand.

AI will never understand. Now, what happens in AI with large language models like ChatGPT and other things, is that they're doing algorithms which are much more complex than simple table lookup. But nevertheless, it's a big number crunching machine. So the modern part of Searle's Chinese room is a slip through the door of a question in Chinese. Searle puts it into this big number crunching machine. He turns the handle, out spits an answer, he takes that answer and slips it out the door. He doesn't understand what's going on. So I think this is really important. The other thing that is a brick wall that artificial intelligence will never go through is creativity.

This was wonderfully defined by Selmer Bringsjord, who has two chapters in our book. And this is the so-called Lovelace test. Now the Turing test, by the way, for those familiar with it, has been passed by ChatGPT.

I don't think there's any question about that. The Lovelace test does something different though. Most tests just look at the output of the artificial intelligence and try to determine the consciousness, or whatever the intelligence of the underlying AI by just looking at the output. The Lovelace test by Bringsjord looks under the hood. He doesn't judge the book by its cover. He looks inside the book. So he raises the hood. And the requirement for creativity, according to the Lovelace test, is that AI will be creative if it does something that is beyond the explanation, or the intent of the original programmer, or programmers. And this has not been achieved yet.

Every AI program which is generated, has been generated because of the intent, and the intellect. And the creativity of the computer program. The AI itself has never generated anything which is creative. We talked offline a little bit about a recent paper, really exciting paper. That what happens if you take a large language model, like ChatGPT, and you train it with all of Wikipedia and everything in the web. And you take this large language model, and then you use the output from it, and only the output from it, to train an other large language model.

Then you use the output from that second large language model to train a third large language model. What eventually happens? Do we have artificial general intelligence? Does it become smarter? No, it suffers from something the authors of this paper called, they call it Model Collapse. And the model actually collapses. And pretty soon, the ChatGPT a few generations down the road sounds like a

babbling idiot. So no, this is not going to happen. In fact, I mentioned this to Pat and he says, "Wow. Before I gave him the answer, he says, "Wow, this might have some implications for AGI." And I said, "Yes indeed. And it shows that AGI isn't going to work. That these large language models just become babbling idiots after a while."

Pat Flynn:

Yeah. Real quick. For people who aren't familiar with the AGI thesis, can you explain that for us, Bob?

Robert J. Marks:

Sure. AGI is artificial general intelligence, sometimes called strong intelligence. And this was hypothesized by a Google engineer named Ray Kurzweil who said, "We're going to reach the singularity." And the singularity occurs when artificial intelligence duplicates the intelligence of a human being. And then AI will go on to be a super intelligence at some point. So AGI is artificial general intelligence, where the intelligence generated by artificial intelligence is equal to, not simulates, but equal to that of a human being. That has not yet happened. And artificial intelligences, I've explained, will never understand what it's doing. It'll never be creative. And then I think as we talked about in a previous podcast, it'll never be taught to be sentient. So these are brick walls that AI will never go through.

One might ask, "Well, what about super-duper computers of the future? They're going to be able to crunch numbers." Well, there's something called the Church-Turing thesis that basically says anything we do on a supercomputer today, can be done on Turing's original 1930s computer. Now, it might take a billion or a trillion years to do, but computationally they're equivalent. But just because computers get more and more sophisticated, they will still be subject to this Church-Turing thesis. And these arguments against understanding creativity and sentience are still not going to be breached. It's just not possible.

Pat Flynn:

Right. That's really helpful, Bob. Thank you for all that. Angus and Brian, any further thoughts on anything concerning artificial intelligence? Or where you see it's headed in the future?

Angus Menuge:

Well, I'll just say that in terms of practical problem solving, the latest artificial intelligence can seem very impressive. And that's because it's producing intelligent output in this sense. It's output that it would require intelligence for human beings to generate. And so if you want to create a website very, very quickly, then ChatGPT can do it probably more quickly than human designers can.

Robert J. Marks:

And certainly, I don't want to diminish the power of ChatGPT. It is incredible, and I use it periodically to help me do stuff. It is astonishing. And by the way, it's recently been tied in, I believe, with Wolfram who does mathematica. So if you want to ask ChatGPT to do a mathematical problem, and I'll ask it to do some pretty sophisticated mathematical problems, it does it pretty well. It'll actually write out equations for you. It is astonishing. But it doesn't understand what it's doing. It's not creative. It's only regurgitating what it's been trained to do.

Angus Menuge:

Right. And so in terms of the output, it's impressive. It can do quite well on computer programming tests, for example. It can create computer programs, which is impressive. But of course, the real issue is not the behavior, it's what intelligence is. And going back to Turing's imitation game, the problem with the test that he designed is its behaviorist. You're counted as intelligent if you pass as a human as often as a human does. But that of course is just behavior. And we know that there's no direct correlation. After all, I can listen to a humble radio and be impressed by the intelligence of the voice, but that doesn't mean that the radio is intelligent. It's simply a conduit of intelligence. And with its sophisticated algorithms and its ability to troll the internet, of course, ChatGPT, and similar programs, can troll all kinds of data and put it together in impressive ways, but they're not really the origin of the creativity.

And so what we really need to pass Lady Lovelace's test is that the machine actually generates some novel information, which it understands. And of course, the metaphysical issue is that despite its great complexity, there really isn't any reason to think that these AI systems are subjects, that they're just very complex aggregates of parts. All of these switches ultimately are just on or off, and it's a very sophisticated physical system. But there is nothing which credibly could be called the subject of its states, and which could be credited with understanding, or reasoning, or any of these higher cognitive functions. So it appears intelligent, but I don't see any good reason to say that it really is.

Robert J. Marks:

Well, fundamentally, Noam Chomsky has called these large language models, or just generative AI in general, digital plagiarism. In fact, right now, there's a number of lawsuits going on. One for example is Getty Images is suing. Getty Images has a library of millions of photographs that have been used to train generative artificial intelligence. So they are suing these people, because they said they have violated copyright loss. We're also seeing computer programmers saying that these generative AI programs have stolen their intellectually protected software. So it's going to be interesting to see how the courts play out in terms of this digital plagiarism and copyright law.

Pat Flynn:

Yeah, obviously there's a huge range of issues that aren't immediately related to the questions that are interesting us on this podcast. But I mean, certainly I have many friends that are musicians and artists that are terrified of what the future is going to look like for them on this front.

Robert J. Marks:

Oh, here's something you can tell them. Remember the ChatGPT generating worse, and worse, and become blabbering idiots. In the same paper they said this would also happen with music.

Pat Flynn:

Oh, interesting.

Robert J. Marks:

If you music to train music to train music.

Pat Flynn:

So you'll need real musicians to come back in at some point, right? Yeah.

Robert J. Marks:

Exactly. In fact, the paper concludes that in order to keep fresh, that human intelligence is required to update these generative models.

Angus Menuge:

That's really interesting because that's a sign of, as it were, digital entropy. And what we know is that human beings can generate new, coherent information when they design a new piece of technology, or they write a novel, or composes a symphony. So we're a source of coherence. But that evidence suggests that there's a law of information, entropy that these systems left to themselves, if they only interact with each other, that information will degrade and ultimately become nonsense.

Pat Flynn:

This has been a really fascinating and rich conversation. I want to thank you all for everything that's been contributed here. I'd like to ask you each one final question. That is, what is your favorite aspect about this volume? What really stands out about this work to you, either in particular or as a whole? Anything else that you're excited that people discover in the volume? I'd like to hear from each of you on that. And of course, Bob, if you wouldn't mind at the end, let people know the best place to get it as well.

Robert J. Marks:

Okay, great.

Angus Menuge:

I just comment that what I think is perhaps unique about it, is that it is fully interdisciplinary. It brings together disciplines that are related to the mind-brain issue that often don't talk to each other. And secondly, that it tries to be very fair in canvassing all of the options to try and show what are the alternatives to physicalism. It's fair enough to say that physicalism has got plenty of exposure. And so what we want to do is just to open up people's minds to the possibilities. And track it all the way down to, what difference could it make for how I look at neuroscience, or computer science, or some other empirical field?

Robert J. Marks:

Yeah.

Pat Flynn:

Excellent. Brian, how about you? Yeah.

Brian Krouse:

Yeah. I'll just add to that. I agree with everything Angus said there. Add to that is that we targeted within this book to try to make... Or targeting as we thought of as a broadly academic audience. So it doesn't mean that all the chapters are easy, maybe popular level book reads. It might take a little bit of concentration. But what we wanted to do was to be able to have a neuroscientist who might be interested in these philosophical models, to be able to read the philosophy chapters. And not run into a bunch of jargon that they have to go outside the book to try to understand. Make them self-contained, but still robust, and cutting edge, and current. And vice versa with the scientific topics. So that I think is unique, and hopefully we accomplished that.

Pat Flynn:

Excellent. And Bob, how about you?

Robert J. Marks:

Yeah, well, I guess I would echo what Angus said. We have 25 different people that have contributed to this book, and usually many of us work in our own silos. As Angus said, we have people... And I made a list, and this is kind of cool. We have people with expertise in computer science, philosophy of mind, philosophy of science, biology, neurobiology, cognitive science, neuroscience, business, neurosurgery, theology, computer engineering, machine learning, and psychology. So we have contributions from all of these fields in the book. And I have learned so much outside of my silo from this. And we hope that the impact will be that other people will be able to look at this book and learn about the mind-body problem over a very broad spectrum of expertise.

Pat Flynn:

I certainly agree. And again, I want to emphasize how impressed I've been by this volume. And don't let it intimidate you, gentle listeners, even if you're new to this, everything is laid out in a remarkably accessible way. Yes, you'll be challenged, but that challenge will be very productive, I assure you. So the book, once again is called *Minding the Brain: Models of the Mind, Information, and Empirical Science*. And Bob, if you wouldn't mind, where is the best place for people to grab a copy of this?

Robert J. Marks:

Well, the book itself has a site on Discovery, and it's called [discovery.org/b/minding-the-brain](http://discovery.org/b/minding-the-brain), with hyphens between it. So that's [discovery.org/b/minding-the-brain](http://discovery.org/b/minding-the-brain), with dashes. Like everything else in the world, it's available on [amazon.com](http://amazon.com). So that's a good place to go to purchase the book, or any other places that you buy books.

Brian Krouse:

And Bob, I believe if you just go directly to [www.mindingthebrain.org](http://www.mindingthebrain.org), that will get you there too.

Robert J. Marks:

Oh, thank you. [Mindingthebrain.org](http://Mindingthebrain.org). Okay, thank you. [Mindingthebrain.org](http://Mindingthebrain.org). Remember mention the website three times, [mindingthebrain.org](http://mindingthebrain.org). I learned that from a politician. Said that you have to do that in order to get it into people's brains.

Pat Flynn:

Yes. Well, it's definitely effective. Gentlemen, I want to thank all of you for taking the time to have this discussion. It has been extremely enriching. And if people are just catching the last part here, of course, I'm going to encourage them to listen to part 1 and part 2 as well. I really hope we can have the opportunity to do this again sometime. Thank you all.

Announcer:

This has been Mind Matters News. Explore more at [mindmatters.ai](http://mindmatters.ai). That's [mindmatters.ai](http://mindmatters.ai). Mind Matters News is directed and edited by Austin Egbert. The opinions expressed on this program are solely those of the speakers. Mind Matters News is produced and copyrighted by the Walter Bradley Center for Natural and Artificial Intelligence at Discovery Institute.

