Vijay Namboodiri and Ali Mohebi on the evolving story of dopamine's role in cognitive function

Researchers discuss the classic stories of dopamine's role in learning, ongoing work linking it to a wide variety of cognitive functions, and recent research suggesting that dopamine may help us "look back" to discover the causes of events in the world.

27 September 2024 | by PAUL MIDDLEBROOKS

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

Vijay Namboodiri
The type of things that you want in an algorithmic way of thinking is to look for invariants in the data. What are the things that don't change? This is a very powerful invariant. If that's true, then, this should be the thing that everyone should be modeling because this is the core thing that any algorithm should capture.

Ali Mohebi
We might not know what dopamine does and what dopamine is, but we almost definitely know that dopamine is not pleasure. One takeaway, dopamine is not the pleasure signal, and maximizing your dopamine is not necessarily a good thing. There are specific predictions that predictions where TDRL fails to make predictions about what happens.

Paul Middlebrooks
Why would that be controversy as opposed to progress?

Vijay Namboodiri
It's a very good question. Why is that controversial?

Paul Middlebrooks
This is "Brain Inspired," powered by *The Transmitter*. Hello, good people. I'm Paul. Vijay Namboodiri runs the Nam Lab at the University of California, San Francisco, and Ali Mohebi is an assistant professor at the University of Wisconsin-Madison. Ali has been on the podcast before a few times, and he is interested in how neuromodulators like dopamine affect our cognition. It was Ali who pointed me to Vijay because of some recent work that Vijay has done reassessing how dopamine might function differently than what has become the classic story of dopamine's function as it pertains to learning.

The classic story is roughly that dopamine is related to reward prediction errors. That is, dopamine is modulated when you expect a reward and don't get it, and/or when you don't expect a reward but do get it. Vijay calls this a prospective account of dopamine function since it requires an animal to look into the future to expect a reward. Vijay has shown, however, that a retrospective account of dopamine might better explain lots of known behavioral data. This retrospective account links dopamine to how we understand causes and effects in our ongoing behaviors.

In this episode, Vijay gives us a deep history lesson about dopamine, also his newer story, and why it has caused a bit of controversy, and how all of this came to be. Coincidentally, I happened to be looking at *The Transmitter* the other day after I recorded this episode, actually. Lo and behold, there was an article titled "Reconstructing dopamine's link to reward." Vijay's featured in the article among a handful of other thoughtful researchers who share their work and ideas about this very topic. Go check that article out for more views on how the field is reconsidering how dopamine works. I link to that and all of Vijay and Ali's information as well in the show notes at braininspired.co/podcast/194.

OK, hope you enjoy our conversation. Vijay and Ali, thanks for coming on. I've had Ali on before, and Ali pointed me, Vijay, to your work and said that you'd be an interesting person to talk to, and he pointed me to a specific paper. I thought it'd be fun to have Ali on as well because he's a dopamine expert, you're a dopamine expert, I don't know anything about dopamine. I thought it'd be fun to have a conversation here. Welcome, and thanks for being on the podcast, first of all.

Vijay Namboodiri
Of course. Thank you.

Paul Middlebrooks
Having said that, I put the word out to my Patreon supporters in my Discord community last night. It was late. There wasn't much time. All I said was that I was going to have a couple of dopamine experts on, and to send me any dopamine questions you might have. The question I got reads like this, "I'd really like to hear a view on how the causal association story contrasts with the reinforcement learning story. I understand it's been quite controversial in the literature, and so it would be great to hear from experts." Then he sent me a link to your paper.

Vijay Namboodiri
I see. Wow.

Ali Mohebi
There you go. There you go, Vijay.

Vijay Namboodiri
That is a very well-informed Patreon community.

Paul Middlebrooks
They're on top of it. I do want to talk about that, but I also want to talk about dopamine, the dopamine story writ large and just, historically, and how it's evolved.

Vijay Namboodiri
Yes, absolutely.

Paul Middlebrooks
I'll start off by reading, I did a little ChatGPT research here. I'm going to read 10 functions, 10 theories about the function of dopamine, and I'm not going to describe them. Reward prediction error, incentive salience, motivational drive, learning and habit formation, attention and cognitive flexibility, motor control theory, aversive learning and punishment prediction, effort-based decision-making, dopamine as a generalized neuromodulator, and finally, social and emotional regulation. That's too many functions, isn't it?

Vijay Namboodiri
Yes, absolutely.

Paul Middlebrooks
It's telling that the first one listed was maybe the first real, oh, I don't know, well, maybe you can correct me on this, success in the dopamine literature was that temporal difference learning, reward prediction error. That's the classic story of dopamine, right?

Vijay Namboodiri
Yes, absolutely. 100 percent.

Paul Middlebrooks
What is that story? If one of you could just [crosstalk]

Ali Mohebi
May I interject? I would say that that's not where it started.

Vijay Namboodiri
That is correct.

Ali Mohebi
I think we always learn function from malfunction. With dopamine, what happens if you deplete a person, a patient, from dopamine? That's Parkinsonian. I think the motor control aspect is where it all started.

Paul Middlebrooks
I should back up and say, so the big success story in my little area, the thing that this podcast focuses on, and how neuroscience and AI can coexist and inform each other, was the temporal difference learning, reward prediction.

Ali Mohebi
Right. I think, early in the '80s, when people started doing electrophysiology, sending down electrode in monkey brains and looking at dopamine cells, they were expecting to see motor-related signals. Early Schultz studies were all trying to look for motor-related signals. They would see if a monkey's sitting there, reaching out to grab a piece of apple, there's some dopamine activity when they get there. Surprisingly, it was always after the movements. If a signal is related to movement, why is it happening after that? In time, these stories of TDRL were developed. I'll let Vijay talk more about the history of that. I just wanted to say that is not where dopamine story started.

Paul Middlebrooks
OK. Yes. No, I appreciate that. Vijay, you don't have to go into technical detail or anything, but just broad strokes.

Vijay Namboodiri
Yes. I think, actually, one more thing to add to Ali's point about the history is that even pre-dopamine-becoming TDRL, in addition to the movement stuff, there's also the idea that dopamine's related to pleasure. It's a rewarding molecule. That was actually maybe even one of the OG theories before. Actually, in popular press, lots of people that I talk to-

Ali Mohebi
It still is.

Vijay Namboodiri
-it is still that theory. The great thing about dopamine, I think, as a case study, in terms of how science progresses, is that--

Paul Middlebrooks
Yes. We'll talk about this.

Vijay Namboodiri
There's been multiple ideas of what dopamine does right from the get-go. Right from when it started, there's always been multiple views. I think that's a sign of a healthy scientific discussion. That there are different things going on. I do agree that I think, for right now, definitely, the most well-established view is a temporal difference reinforcement learning view. The discovery by Schultz was seminal, that kick-started, in some sense, this whole field of at least what dopamine function is as it relates to very rapid dopamine activity.

There's this parallel literature, as that was developing, where people were looking at long timescale disruptions of dopamine signaling, or Parkinson's disease, it's like the longest timescale one, but then also pharmacological inactivation experiments. Where that story is slightly different from the standard story of the learning story.

Paul Middlebrooks
Can they coexist? You started off by saying something to the effect of there are many different views, and then you just started talking about the timescale. Oh, maybe they're all correct. Does dopamine need to have-- Oh, no, here's what you started off by saying, it's a sign of healthy science. Yet, I'm not sure how you feel about your dopamine theory, but then people who have their own dopamine theories they need to be staunch supporters of their own theory, and everyone else is wrong.

Vijay Namboodiri
Yes. I don't think that that's the right way to look at things. No, of course. there's one line that I think in one of the reviews, it's a beautiful thing. Even for dopamine, it is too much to be doing all these things.

Paul Middlebrooks
Even for the dopamine.

Vijay Namboodiri
That's, I think, a great line in many ways, some of these theories are complementary, but in many ways, they're also not

complementary. It can't be that they're all simultaneously correct. It can't be that a single molecule does everything. Even if it's involved in everything, maybe the way that it's involved in everything is probably not exactly the same algorithmic function and stuff. At broad strokes level, I think it is a sign of a healthy debate that there are different views. The next step that we want to do is really figure out exactly what the things are that are not possible to coexist in these different views, and what are the things that are possible to co-exist.

Paul Middlebrooks
Just to contrast it with a brain area. I think it's great that it has a neutral name, dopamine. It doesn't say pleasure amine or surprise amine. Whereas in the cortex, for example, there's a brain area named motor cortex. That means it does motor activity. That's not all it does, but because we've named it that, it's really hard to think of it in any other way.

Vijay Namboodiri
Yes, absolutely.

Paul Middlebrooks
Dopamine is neutral.

Vijay Namboodiri
I think that is good. There's a funny one along those lines. We've named our field neuroscience. It's named after one cell type in the brain.

Paul Middlebrooks
Yes. Right. Do we need to switch it to brain science?

Vijay Namboodiri
Brain science, that's probably the more welcoming view of all the different cell types.

Paul Middlebrooks
Yes. Maybe we can do brain science and the rest of the world can do neuroscience [crosstalk]

Ali Mohebi
Then brain and behavior. Brain is nothing without behavior.

Paul Middlebrooks
Oh, my gosh. We have to go into society and the universe.

Vijay Namboodiri
Yes, we're going right to all the key controversies.

Paul Middlebrooks
All right. Let's focus back down on that tiny molecule.

Vijay Namboodiri
Yes, we'll go back to the temporal difference reinforcement learning stuff. Yes, to back up in the history. TD, I do believe that it was really one of the big successes of this field, because it is one where we had, at least in neuroscience up until then, there were aspects of this going on in other fields. TD was one of the poster tiles for where we have these computational theories that were developed in a different field that have clear algorithmic function and can be very effective. You find that something like that actually exists in the brain. That's exactly what every theorist's dream is.

It is, essentially, what kickstarted this whole way of thinking, of thinking about things in an algorithmic way. It is easily the most successful theory in that sense. Before we get to what the current state of the field is and whatnot, just to explain what TD is, the core idea actually predates TD. The core idea comes from psychology. It's a question of how do animals learn? What do they actually learn about? Is there some algorithmic principle whereby they learn associations in their world? Associations are probably too many things. You could learn that one cue, one word is associated with something else, so on and so forth. It's a very general set of things.

People, just because it's in animal models, it's actually very difficult to assay things without any behavioral output. You typically only study associations related to rewards or punishments, because you can actually get these very defined motor outputs that you can clearly tell that the animal actually is responding and sensing something, first of all. Then responding to something else based on the fact that they predict something else. The set of associations that we study, in general, is limited in that sense, just because of the way that we assay them in animals.

Within that field, the biggest success is the Rescorla-Wagner theory in 1972. The idea being that up until then, the core idea was that the way that animals learn and people learn, associations is when two things happen relatively close together in time, then you learn that those two things are associated. People pre-Rescorla Wagner found some really interesting examples where that's not true. Where you can have stimuli that are very close together in time with a reward or a punishment, but you still don't learn that is actually associated with it. The core insight there from Rescorla-Wagner was that actually a simple way to describe all these different phenomena is that you learn based on error.

That was the first big revolution in this field, in the question of how animals learn. The core idea of that is very simple. You just have some prediction of what will happen. You have some prediction of rewards, whether they will happen, or punishments of whether they will happen, then you actually see what happens, then you look for the difference. If there is a difference, that means that your prediction was wrong. You could use that difference to actually improve your prediction the next time you see the same thing.

Paul Middlebrooks
The way that you're saying it sounds like you're aware of all this, but it might be more accurate to say that circuits in your brain take care of this for you. You don't have to be aware of predicting it. It's just happening in the brain as an algorithm.

Vijay Namboodiri
Yes, that is 100 percent true. How much of this is actually conscious versus not is unclear. Most of it is probably unconscious. I think the way that we all talk about our favorite theories, we talk about intuitions. The best way to tap into intuitions is to think about conscious things that we at least have recent proof. All of this could be completely subconscious. The idea was that prediction error, specifically, and specifically reward prediction error as far as rewards go. Is the key quantity that allows you to learn better predictions of whether things will actually be followed by rewards. That was all done in a trial-based way, if you will.

The idea was that every time you have a trial, you get, let's say, one cue predicting associated with a reward, and then you don't think about time in any other sense. It's just, basically, on this trial you had a cue, and you had a reward, or you didn't have a reward. Then, collectively across a bunch of trials, do you learn to associate the things or not? Obviously, that is not how animals live. Animals don't magically know that right now this is a trial kind of a thing. That was obviously a limiting factor.

Temporal difference reinforcement learning actually started to solve that for the first time, where it actually started incorporating time within a trial. That was one of the key advances in TDRL when it comes to the neuroscience aspect of things. There's a whole bunch of TDRL stuff related to computer science that I'm not going to get into.

Paul Middlebrooks
I was going to say that it also drove reinforcement learning algorithms in AI. [crosstalk]

Vijay Namboodiri
Yes, absolutely. It still is.

Paul Middlebrooks
Still does.

Vijay Namboodiri
Still does. The core thing as far as neuroscience and animal learning goes is that TDRL allows you to have some way that you can define the progress of time within a trial. The idea is you could have, for instance, on a single trial, you have one cue, followed by another cue, followed by reward. Now this allows you to actually keep track of the sequence of things within a

single trial, which is actually hard to do with the Rescorla-Wagner type of theory. Because that was all just at the same moment, there are a bunch of things happening versus not happening.

It clumped the entire trial into a single thing, where the sequential effects was hard to actually model within the Rescorla-Wagner framework. The key advance in TDRL is that this allowed you to actually form these predictions with good time resolution.

Paul Middlebrooks
Correct me if I'm wrong, it's attractive because it's a fairly simple idea.

Vijay Namboodiri
Exactly. Yes.

Paul Middlebrooks
That's another reason, Occam's razor approach, it's appealing in its simplicity.

Vijay Namboodiri
Exactly. It's super elegant. The core idea of TDRL is extremely elegant, the basics of it. As far as a didactic version of TDRL goes, the one that I teach students, it's a beautiful theory because it has the core elements that you need to actually explain the phenomena that you want to explain.

Paul Middlebrooks
I didn't realize that you're teaching students the history of this stuff too, because then you have a real left turn probably midway through the semester with your own work.

Vijay Namboodiri
Actually, I don't teach students yet my own work, because I actually do just one lecture on this stuff for the core systems neuroscience class here for the grad students. On that resolution, I feel like they need to know TD and Rescorla-Wagner way more than they need to know my work. The context of that is way more important, because that is the way of thinking. For students, I'm trying to get people who have not thought about anything related to this, from all backgrounds in neuroscience, to actually start at least thinking about the problems in neuroscience in this sort of way.

Paul Middlebrooks
In an algorithmic way.

Vijay Namboodiri
In an algorithmic way. For that, I think our work comes later. It's not quite there yet.

Paul Middlebrooks
It comes on the "Brain Inspired" podcast. [crosstalk]

Ali Mohebi
Having said that, I will be teaching Vijay's work next semester. I'll do that. You don't have to teach. Yes, I will be. [crosstalk]

Paul Middlebrooks
You'll be teaching a full course then?

Ali Mohebi
Yes, it's a full course on neurobiology of learning and decision-making, so causal association and dopamine story is definitely there.

Paul Middlebrooks
Where will Vijay's work come in? Have you got the syllabus ready yet?

Ali Mohebi
Yes, I have 85 percent ready. I'll send it to you.

Paul Middlebrooks
I realize we're burying the lead, but that's OK. We're building attention. You just very nicely told that historical temporal difference learning story, and why it's attractive, how it worked, and how it ushered in this algorithmic thinking. Why are there so many other theories? Why are there so many various theories that all have some supporting evidence, etc.?

Vijay Namboodiri
It's a good question. Actually, before I get into our own stuff, I want to still stick with a historical perspective.

Paul Middlebrooks
Sure, please.

Vijay Namboodiri
While the TDRL stuff was becoming very successful. Where Schultz's seminal work, you start recording dopamine neurons or putative dopamine neurons in Schultz's case. You find that the activity of these neurons in a very fast temporal resolution actually seems to abide very nicely with this prediction error idea. That's very powerful. That stuff was going in parallel. You also had all this data on dopamine pharmacological inactivation type things where you actually inactivate dopamine on longer time scales. You look for whether dopamine is important for learning, whereas dopamine is important for other things.

In general, the story there was much more complicated. The story from the TDRL side was very simple. Dopamine is really about learning, here's this, we found the magical signal that is useful for learning. It's prediction error. That story, it's elegant and very nice and narrow. That's what you want powerful theories to be. Simple and broad in scope in terms of explanatory power.

Paul Middlebrooks
Then you do an experiment.

Vijay Namboodiri
Yes. One of the things, one of the mysteries in the field that is still not resolved, and that's not even not related and not solved, I guess, including in our own work, is the idea that there are forms of learning that people find that dopamine is not important for it, as best as we can tell like pharmacological studies. The classic example of this one is sign tracking versus goal tracking. Are you aware of this?

Paul Middlebrooks
What's sign tracking?

Vijay Namboodiri
This is all work from rodent studies. What you find is basically that if you do just a simple Pavlovian conditioning, so the idea is you have a rat, let's say there's a lever. It doesn't have to press the lever. There's, let's say, a port where you get the reward. Let's say the lever is the cue. You use the lever as the cue. The lever just comes out before the reward is available. You don't have to press the lever at all, then a reward comes later. When you take normal rats and make them do this task, you find that, broadly speaking, there are two classes of animals.

Both classes of animals learn the association between the lever and the reward. One class of animals, actually, once they learn that the lever predicts the reward, the moment the lever comes out, they'll go out and hang out by the reward port. Those are called goal-tracking animals, because they know that the rewards are about to come and they're right there by the reward to actually collect it.

Paul Middlebrooks
By sign, you mean cue. It's a synonym for cue?

Vijay Namboodiri
Yes, exactly. The alternative is the sign trackers. Those animals, actually, when the lever comes out, they go to the lever and they start messing with the lever, press the lever, bite the lever, chew the lever, all sorts of things. They don't go by the reward. When the reward actually gets delivered, they'll walk over and collect the reward. It's this weird dichotomy. Both of them are learning. They're both learning Pavlovian associations in that the cue predicts the reward. They both know that.

Paul Middlebrooks
Except the sign-tracking animals are actually thinking they're learning some sort of operant conditioning, I suppose, because they're not passive.

Vijay Namboodiri
Yes, they're not passive. It's not clear that they think that they need to press the lever.

Paul Middlebrooks
Sure, you can ask them.

Vijay Namboodiri
Yes, either way. Some of the behaviors that they produce it's not just-- Typically, in the open lever-pressing test, they just go and press the lever and they don't do anything else. Here the sign trackers are different because they actually go and chew the lever.

Paul Middlebrooks
Oh.

Vijay Namboodiri
They do even conservatory-type things as if the lever itself is a food, or the lever itself is a reward. It's a weird thing.

Paul Middlebrooks
That is weird.

Vijay Namboodiri
What you find is that dopamine is actually not needed for that type of behavior, for the sign-tracking. Dopamine's actually-- You cannot get goal-tracking, but sign sign-tracking seems to be independent.

Paul Middlebrooks
Wait, you deplete dopamine, and then all of a sudden, all you have are sign trackers. You do not deplete dopamine and then you have some sign trackers and some goal trackers?

Vijay Namboodiri
Yes, exactly. That's weird. The idea was basically that you have this whole type of behavior learning where animals actually don't seem to need dopamine. That was a parallel literature, essentially, in that what exactly is dopamine useful for wasn't actually clear.

Paul Middlebrooks
They're learning something?

Vijay Namboodiri
They're learning something. Exactly. All of this was fodder for the other theory that dopamine's actually important not for learning per se, but for ascribing value to cues [crosstalk] incentive motivation.

Paul Middlebrooks
Yes, that goes hand in hand. I was going to say with motivation and, uh, goals.

Vijay Namboodiri
With motivation, exactly.

Paul Middlebrooks
Goals.

Vijay Namboodiri
Yes. Which makes me think that it's the other way. Now am not [chuckles] Now I'm blanking.

Ali Mohebi
It's a sign tracking perhaps. I would say that it's a sign tracking. You're right. If we are to recreate the theory, that makes total sense for dopamine to be needed for sign-tracking because with incentive value idea that you are attributing value to this cue.

Vijay Namboodiri
Absolutely.

Ali Mohebi
As Kent Berridge would say that it's like a magnet. Approach behavior is an important thing in motivation research, at least in rodents or animal work. When you're motivated, you are more likely to approach a reward or a rewarding cue or something. That becomes important in drug addiction research as well. It's not the reward itself that you're attributing values to, it's the sign that is the cue that you get attracted to. That's why when you're re-experiencing an environment where you had a high at you had some drug-related incident happened, you will recreate all these--

Paul Middlebrooks
You chew on the crack pipe instead of loading?

Ali Mohebi
You may do that because there's value in that cue itself, not the high.

Paul Middlebrooks
An intrinsic value.

Ali Mohebi
Dopamine is not related, and it makes sense what Vijay mentioned earlier that dopamine is not the pleasure signal when people still-- If you go ask in the street people who say dopamine is a pleasure signal, but that we know--

Paul Middlebrooks
Oh hell, if you [crosstalk] ask any a popular neuroscience communicator, they would say that.

Vijay Namboodiri
Yes, that's true.

Ali Mohebi
They would say that cold shower also has some effects on dopamine. [chuckles]

Paul Middlebrooks
You mentioned Berridge there, is Berridge the liking?

Ali Mohebi
Liking, wanting. Yes, that is exactly the idea. Liking would be the pleasure?

Vijay Namboodiri
Yes.

Ali Mohebi
When you like something, it is the pleasure part, but then wanting is the approach, is the motivation part that's more related to dopamine.

Vijay Namboodiri
Yes. I did mix that up. I've said everything correct except for the fact that dopamine's not important for sign tracking. It's actually crucial for sign tracking, not for goal tracking.

Paul Middlebrooks
Oh, OK.

Vijay Namboodiri
Which is exactly what I was leading up to, that it is important for incentive salience.

Paul Middlebrooks
OK.

Vijay Namboodiri
I just mixed them up.

Paul Middlebrooks
What is the incentive salience, what were you going to say about that?

Vijay Namboodiri
The idea about incentive salience is basically-- It's the second hit on your ChatGPT.

Paul Middlebrooks
I think so.

Vijay Namboodiri
The idea for incentive salience exactly what Ali was saying, that basically, dopamine's not important for learning per se, it's actually important for ascribing motivational properties to cues.

Paul Middlebrooks
Oh, this is related to your work that will come to perhaps.

Vijay Namboodiri
Perhaps.

Paul Middlebrooks
I guess they're all somewhat related.

Vijay Namboodiri
That's right. They're all related. They all have related components. The idea was, basically, that the animals that actually go and chew on the lever and stuff, that seems to be dopamine independent. The animals that just go and hang out by the reward port where they know exactly what's going on, that seems to be dopamine-independent. This is the controversy in the fields at the time where you had all the Schultz stuff developing where it was more and more evidence that rapid time-scale dopamine activity is all related to learning, crucial for learning. Here is one type of learning that dopamine, in fact, you would argue in a simpler form to understand, doesn't require dopamine, or doesn't apparently require.

Paul Middlebrooks
Actually, if you're learning, you're learning the wrong thing.

Vijay Namboodiri
Yes. Now what is right versus wrong is unclear because it's not operant, because the animals are free to do whatever they want.

Paul Middlebrooks
OK, you're learning the thing that is going to decrease your survival and, therefore, decrease your evolutionary lineage. One could say, if evolution is normative, you could say that's the wrong thing. Big leap there, I know.

Vijay Namboodiri
Yes. That could be an interesting discussion, but actually, let's say we skip that discussion because there are complex thoughts.

Paul Middlebrooks
I just mean, how about, a simpler thing is you're going to get satiated less quickly.

Vijay Namboodiri
Basically, if you're hanging out by the reward?

Paul Middlebrooks
You have the dopamine, you're going to go chew on the lever. That means you're going to get the reward a second later, 500 milliseconds later.

Vijay Namboodiri
Yes. The argument, now I need to go back to the data. I think that as best as I remember, these animals were not slow at picking out the reward. They were slightly slower, but not that much slower. The complex thing that I was going to say is that the argument is that if you were to go by the lever and chew on the lever and all that, maybe that is exactly what increases your survival. You know exactly what the cues are that are important.

Paul Middlebrooks
Yes, but it's more important just to see the lever and make a beeline to the-- Anyway, I'm nitpicking.

Vijay Namboodiri
No, I agree. I'm just saying that there are different views that I've heard people say. OK, cool. That was a controversy that was happening, while the Schultz stuff was developing.

Paul Middlebrooks
A lesser-known story, by far.

Vijay Namboodiri
Yes. Lesser-known story in the, I guess, the recording systems neuroscience community, but in the dopamine community, well-known, at least in the dopamine community.

Paul Middlebrooks
That's the thing that's so-

Ali Mohebi
Specifically, sorry, in the drug addiction, because big part of NIDA is National Institute of Dopamine. A lot of funding, addiction, or dopamine research has been an important part of addiction research. I would say that incentive value has been very dominant and influential in that field, but maybe not as much in the learning and reinforcement learning and AI fields.

Paul Middlebrooks
Yes, and I was going to also say, Vijay, one of the things that you point to is work by people like Randy Gallistel. One of the things that you are deeply familiar with, as is apparent in your work, is all of this literature on learning, and different modes of learning. Which I think, as Randy Gallistel has pointed out to me, is super important and a super rich history to draw from to test these different things. That there's so many different ways you can test learning and different facets of learning that will key into whether your story is more or less correct. That's an important thing that I have missed out on. I need to revisit all of that literature, but I don't know where to begin because it's so vast.

Vijay Namboodiri
No, I agree. Randy's stuff, I would say, genuinely, is lesser-known, when it shouldn't be. Everyone should be knowing this, everyone interested in learning should know all the literature that Randy has talked about. Collaborators of Randy have worked on experimentally. That literature is absolutely crucial. It forms some of the bedrock in our own thinking of how our own work evolved a lot from Randy's set of core ideas to the listing of the problem.

Paul Middlebrooks
One of the nice things about that, I keep pointing to Randy, but that's just because I had him on the podcast, but about that kind of research is the tenor of it is, as you say, "Well, look, we can show that animals behave in this way, and it just doesn't work with your algorithm, and it doesn't work with the account that you're giving of a circuit level mechanism." Therefore, and so you have to account for those sorts of behavioral findings.

Vijay Namboodiri
Yes, absolutely. That has been a tricky one, explaining, so just to back up, in case listeners aren't familiar, and this can lead into our own work. Randy has shown experimentally, not Randy, but like Peter Balsam, others, John Gibbon, etc. Then, Randy has summarized them in a very influential paper. He's talked about this for a while now and has published a lot of papers. They showed that, on the one hand, there's this competing idea. There's the temporal contiguity idea that if cue and reward are separated from each other, you learn less about them.

There's a different literature that shows that if you space things out more, then you learn more from each experience. Just at a qualitative level. How these two things interact was not clear at all. Some of the old literature actually suggests some extremely interesting relationships between both of those aspects of learning. What they found was that, actually to back up. In TDRL, temporal contiguity, even though we say that it moved beyond temporal contiguity, and by saying that error prediction and prediction errors are a thing that drive learning. Temporal contiguity is still a factor in the learning. In that, if you separate cue and reward more, there's more temporal discounting.

You actually end up learning less. It's harder to learn. Temporal contiguity is still a factor. The interesting thing in the other literature, vast literature, as you're saying, shows that if you actually increase the cue reward delay, and I'm saying cue reward, but a lot of the stuff was actually done with punishments and stuff. I'm just going to stick with cue reward just for simplicity.

If you increase the cue reward delay, it's not clear. It's not actually apparent in the data that it always increases the time to learn or the number of trials to learn. You can have the cue reward delay be long. If you increase the inter-trial interval also in a proportional manner, you find that the number of trials it actually takes to learn is actually conserved. It's invariant.

Paul Middlebrooks
Counterintuitive.

Vijay Namboodiri
Yes, very counterintuitive and very profound. The type of things that you want in an algorithmic way of thinking is to look for invariance in the data. What are the things that don't change? This is a very powerful invariant. If that's true, then, this should be the thing that everyone should be modeling, because this is the core thing that any algorithm should capture. It turns out that a lot of the original work that summarized this was a meta-analysis in the early '80s, but from work in the '70s, and then Randy's paper popularized it in early 2000s. It's largely been ignored in the neuroscience community. Very few neuroscientists.

Paul Middlebrooks
It's terribly inconvenient.

Vijay Namboodiri
Yes, exactly. It is terribly inconvenient. Especially to the dominant model, the TDRL model, it is inconvenient. I think that's the reason why maybe it's been ignored. That, I think, is a fatal flaw. That's for sure. I can say pretty confidently that we should rectify that as a field. The difficulty is that in a standard TDRL view of how learning works, these sorts of invariants don't naturally come about. You can maybe bake that into the system, into the algorithm, but that's not like it's a natural thing. It's not a natural thing. You're just adding in that thing as a constraint. That's not really satisfactory.

People have taken cracks at it and found that there are ways to take the error-based idea and try to capture this, but not in an easy way that captures all the other things with dopamine. This is, I think, a somewhat unsolved problem on the TDRL side of things. As I said, somewhat unsolved in that, there's been some attempts, but it's largely not been shown that you can actually capture the standard dopamine things while also capturing this in the same framework.

Paul Middlebrooks
You're telling the story as if this is the history of your thinking about it, almost. This led you to the conclusion that you needed something new, something else, something to account for. What I want to know is where your idea, which is a fairly simple inversion of TDLR, how you came up with that? Maybe you could just really briefly say what the idea is, and then we can unpack it more, because I want to know how you came up with the idea, because it's so simple, man.

Vijay Namboodiri
Yes, it is a very simple idea. The core essence of it is extremely simple. To explain what the idea is. The core idea of TDRL is that the way that you learn predictions is by learning clue prediction errors. You make a prediction and then you look at what actually happened, and there's an error, and then use that error to actually update the prediction.

Paul Middlebrooks
There's a bell and then usually you get a cookie after the bell. Sometimes there's a bell and you don't get a cookie and you learn, then you have a different signal. The first time you get a cookie after the bell, you weren't expecting it. The dopamine says, "Whoa, I was not predicting that. There's a big error because you get the reward."

Vijay Namboodiri
Exactly.

Paul Middlebrooks
These signals are modulated based on your predictions and the error that is generated.

Vijay Namboodiri
Exactly. The critical thing here is that everything that you need in this algorithm is forward-looking if you will. That you're always looking for future relationships. Exactly what predictions are.

Paul Middlebrooks
You call that prospective.

Vijay Namboodiri
Prospective. Exactly. The alternative view, just to state it simply and then we can go through history and how this came about. The alternative view is that actually, there's a different way that you can learn associations, and it's simply by looking backwards. Imagine that you got the ice cream or something like the output of the cookie. Then you look back to see what might have been the thing that caused this. If you consistently find that something precedes it, then you know that those two things are associated. It turns out that you can show mathematically simple Bayes'-rule-type things showing that if you know the backward associations, you know the forward associations too. Now you can compute it very easily.

Paul Middlebrooks
Here's my guess is that you were just thinking about Bayes' rule. You thought of the prospective story as part of some conditional probability. You're like, "I could just reverse this with Bayes' rule. Am I right? Is that [crosstalk]"

Vijay Namboodiri
That is definitely was one of the steps.

Paul Middlebrooks
Ah, one of the steps, I'll take it.

Vijay Namboodiri
Yes. It's very close. Now, going to how this evolved, and actually maybe before going that, just to say why is this so intuitive in some sense. The version of the story that I give in talks to get people intuition. It's also one that I tap into consciousness for just because that's the easiest way for people to actually tap into the intuition. All of this may be subconscious. The core intuition is this. Let's imagine that one day you feel sick. You feel nauseated, stomach pain, etc. You feel like you've been food poisoned or something. You ate something that you just didn't like. What do you do?

I will say pretty much everyone looks back in their memory to think about what is the thing that they might have eaten that would have caused this. Let's say that you happened to eat at a new restaurant that day. I probably bet that you would probably not go back to that restaurant again for a little while, because you will associate eating at that restaurant with getting sick. Now from that example, there are two things that are clear. One thing is the way that you associated eating at that restaurant with getting sick is a backward, is fundamentally backward.

When you were eating at the restaurant, you weren't thinking, "Am I going to get sick?" It's just that when you got sick, you were actively looking back in memory to think about where you ate. You realize that there's a possible explanation.

Paul Middlebrooks
You just made me realize I believe my wife is slowly poisoning me, but go ahead. I couldn't resist, I'm sorry.

Vijay Namboodiri
Oh. The other thing that you realize from there is that once you've associated eating at that restaurant with illness, then you can invert that retrodiction to a prediction. You intuitively do that. The thing that makes you decide not to go there is not the retrodiction. It's a prediction that if you see the restaurant, you will probably imagine getting sick. That's why you don't want to go there. You've intuitively converted a retrodiction to a prediction. Intuitively, it seems like this process is happening, that you learn a retrodiction and then you convert that implicitly to a prediction. Question is, how does that happen? Bayes' rule is the answer.

Paul Middlebrooks
Why has this not been thought of years and years ago or proposed, or has it in some form?

Vijay Namboodiri
Yes, so this was the thing. This worked out, and I'll get into that a bit more, but once I worked out the full theory on the retrodiction to prediction conversion, I was just puzzled by exactly this thing. It seems such a simple idea. Why is it this not being thought about? The answer is after working this out, and I'll tell you there are some key differences in the way that we worked it out from the prior attempts at it. I did realize that people have thought of similar things. In fact, in one of the original things that kickstarted the whole field is Common's blocking experiments that preceded Rescorla-Wagner. Common's explanation for blocking is something similar to this. That you basically look back-- He called it backward scanning. The idea is that you look back to find associations.

Paul Middlebrooks
What is blocking? Let's just say what blocking is. Sorry. I know because it's just a technical term. [crosstalk]

Vijay Namboodiri
Absolutely. This is something that I alluded to at the start. This is actually one of the most influential results in the early days in psychology, which actually moved the field from temporal contiguity as the key thing to learning from errors. The idea is this. Imagine that you first teach an animal that one cue predicts a reward. A lot of this was done with punishment, but again, I'll stick with rewards. Cue predicts reward. You've already learned this. Then, in the second phase, what you do is while you present this cue, you also present another cue, and then give the same reward.

Now, this cue is obviously temporally contiguous with this reward. If it's just temporal contiguity that is the thing that allows you to learn, you should learn the new cue reward relationship as well. It turns out that, in general, animals don't. At least, they don't show behavioral evidence that they do. Why is that? That's called blocking, in that, the first cue actually blocks the ability of the second cue to learn.

Paul Middlebrooks
To learn a new cue.

Vijay Namboodiri
To learn a new association for the new cue. The idea there was that Rescorla-Wagner showed that, well, if this is all driven by an prediction error, then this works. Because, if this cue already predicts reward, then the second time this cue is associated with reward, even though there's something else, this has already been predicted. There's no prediction error here. You don't learn the second cue. The original explanation that the person who discovered this actually gave was not that, was actually that you're looking back for causes type thing. It is a retrospective backwards scanning.

Paul Middlebrooks
That's similar to your idea that, well, so, if you're looking back, that first cue is predictive 100 percent of the time, and then the second cue is just predictive less percentage of the time because it was introduced later.

Vijay Namboodiri
It was introduced later, exactly, yes. You don't need to attribute causal significance to that thing. There's always distractors that happen in your world. You don't need to assign causality to everything that precedes something. You want to know that it consistently precedes it. This idea that was maybe the original form of the backward retrospective view. Now, the critical thing that was missing there was this inversion from the retrospective to perspective. In that view, this was proposed as a way of learning associations, but it was not mathematically formalized as a way that you could actually take this retrodiction and then convert that to a prediction, which is finally the thing that you want to learn.

Paul Middlebrooks
Using Bayes.

Vijay Namboodiri
Using Bayes, or any other form. That was missing in that explanation. Now, the other time that this retrospective thing in the literature, was actually Randy's work. Randy, as we touched upon briefly, has this idea that all of this is just based on a temporal map, or a cognitive map, if you will, of time. If you just store the exact time of when everything happens in memory, and your memory is perfect, let's say, then you can arbitrarily go back and think back about whether things precede something, or things are followed by something. The idea is that retrospective makes as much sense as perspective in that sort of a computation.

Paul Middlebrooks
I see.

Vijay Namboodiri
You can have, and they could, in fact, in the way that he writes it, it's not just a direct Bayesian version, because it's a slightly different framework of the perspective versus retrospective. Either way, that was one other place where the retrospective had made its appearance before we published it.

Paul Middlebrooks
Did you discover this after you had your ideas?

Vijay Namboodiri
Yes. Which is funny, because Randy's work has been one of the core drivers for me, in terms of, at least, thinking about the philosophical problems of learning. Randy just has published so many papers that it's-- [laughs]

Paul Middlebrooks
I know. How could you be expected to? It takes a career just to follow someone else's career, sometimes.

Vijay Namboodiri
Exactly, yes. I just didn't know about this. There's a very specific paper where Randy talks about the retrospective thing, and I just didn't know about it.

Paul Middlebrooks
Ali, has that happened to you frequently, where you have an idea, you start working on it, and then you realize it's been done seven times before in different ways?

Ali Mohebi
My ideas are so original.

Paul Middlebrooks
Original. OK, sorry.

[laughter]

Ali Mohebi
Of course, yes. [chuckles]

Vijay Namboodiri
Here too, then, with the retrospective thing, there was not this perspective-- Conversion of a retrodiction to prediction didn't exist in this way either. That part is new, I think. The way that we propose it, I think that was the first time that that came about.

Paul Middlebrooks
Because you need that because you do need to behave moving forward.

Vijay Namboodiri
Exactly. Moving forward. Exactly. Just given that we are in the historical sense of this, this idea -- There are two threads to how I realized this. One thread is that when I first learned reinforcement learning in TDRL, I had this vague feeling that something was missing, and that was just the time component of it. It took me about 13 years to verbalize what I thought was missing. It took a long time. I knew that it was something related to time, but exactly why, was hard, and we can get into that later. That was one thread where I felt I had some dissatisfaction with the standard view of TDRL.

The retrospective versus prospective thing was almost independent of that thread in that that came about because I'd actually collected some data in my postdoc where I found some weird patterns of responses in orbit of frontal cortex neurons that project a VTA. That didn't make any sense to me. The only way that I could postdoc retrospectively rationalize the results was to think about this retrospective framework at the time?

The first time that in my published work that I mentioned retrospective is actually in the discussion of that paper. Because that paper, it's not designed to test this idea in any way, so it only makes an appearance in the discussion, because it was a loosely formed idea at the end of that paper to just roughly qualitatively explain the patterns of findings that I found. I had that idea that maybe this retrospective thing could work.

The problem with the way that I was just describing it is that all of this is the Rescorla-Wagner type equivalent. It's a trial-based view in that you're just thinking about, does the cue precede the reward? We're just looking at whether on a given trial cue proceeds to the award versus not. The core advance of TD is to go from that view to a time-based view, where you have time differences. That actually requires additional work to show that the same retrospection, a retrodiction to prediction work would actually hold in the standard way that people do TDRL.

Paul Middlebrooks
Because the Bayes' equation has nothing to do with time.

Vijay Namboodiri
With time, exactly. It's just a conditional probability thing. Conditional probability alone does not get you to these long-run sums of value-type things that you define in TDRL, which we can get into as well.

Paul Middlebrooks
It turns out time is fairly important in life.

Vijay Namboodiri
Yes, exactly. It's tricky to think about how time plays a role in this stuff, like in learning. Now that we've maybe built up the intuition for these things, I can start to now poke holes in the standard way that people assume things work in TDRL. It's when I realized these things that I was able to verbalize what I thought were the problems. The question is, how can you then try to figure out a solution for this? That's where I made the connection between the retrospective thing and then the problem with the time stuff.

Paul Middlebrooks
Oh, nice.

Vijay Namboodiri
The view is this. Standard TDRL, all the idea is set in the core didactic way, it's a very simple idea. You just have a cue and a reward here, and you make a prediction here, and you make a prediction that value is zero or something at the next moment, let's say. Then you break up-time into these small components, time bins. Then this time bin, you're predicting that nothing

will happen, but then you actually get reward. Then when you get reward, you actually are surprised. You get a prediction there, you ascribe that prediction to that came before it. Then the standard didactic view, you actually ascribe it to the thing that immediately preceded it. That's, let's say, a time step, so you assign value to that time step. Then the next time, the state before that, gets value from the next state and so on and so forth. Then eventually, you go assign value to the cue, backtracks to the cue.

Paul Middlebrooks
That reward prediction error is a dopamine hit, just to bring it back to dopamine.

Vijay Namboodiri
Exactly. That's the standard TDRL view. Now if you think about how this works in reality, let's-- The thing that you're assuming happens here, is that you keep track of time from the cue. Now, in the simplest way that's obviously a simplification that you keep track of time perfectly, bin by bin. That, obviously, everyone knows is a simplification, so no surprises there. Still, the critical assumption is that you are keeping track of time from the cue. Now, this raises a problem. Think of how many cues there are in our environment. There's so many. Infinite. There's no way that that's not infinite. Are you keeping track of time from every single cue that you experience? Are you doing this in parallel? You know how much time it's been since that cue and then this much time since that cue, and all of those things are happening in parallel. Every cue, you need to have a separate clock for how long it's been since that cue has happened.

Paul Middlebrooks
You should look up Marc Howard's work. He might say that you can do that using a Laplacian transform, but no need to revisit that right now [crosstalk]

Vijay Namboodiri
Yes, very familiar with Marc's work. Marc's stuff has relationships to the learning as well, which is a different angle to this whole thing. I don't think that he solves this particular problem in the standard TD way. We're still talking standard TD, that each cue, you're assuming that you keep track of time from there. Then when you ascribe value for that reward or prediction error, you're ascribing value to that cue in terms of how long it's been since that cue has happened, so for each cue, you assign it differently.

The other critical assumption is that all of TDRL depends on states as the input to this algorithm. Basically, the idea is that TDRL is an algorithm that operates on some inputs and gives you some output. The output is the value, the input is the state of the world. In standard computer science stuff, this all makes very simple sense. States are something that you know in priori in an engineering setting, you know exactly where you're trying to learn. It's very easy to define all this.

Now, when you're trying to take that architecture and ascribing it to animal behavior, you have to make a bunch of assumptions. Basically, the state inputs that you give to this algorithmic black box, if you will, it's not a black box, obviously we know every component of it. The inputs that go into it are what exactly? What are the state inputs? Essentially, what you need is a state input that tells you that for every time moment or every time step, something tells you that this is the current state of the world.

Because TDRL fundamentally is about temporal differences. What that means is there's one time step and another time step, and you're calculating the difference in predictions between these time steps. To do that, you need to know what is the thing that allows you to even make a prediction at this moment? That is the thing.

Paul Middlebrooks
You have to have predefined states, essentially.

Vijay Namboodiri
Exactly. Now, when you think about that thing with this problem, standard cue happened, let's say, trace conditioning, so nothing is happening, there's a lot of delay. A cue happened, there's a lot of time delay where nothing else is happening. No external input is being given to the animal. Then a reward comes. Then you need to have a state defined for every moment in time. How do you define a state for every moment in time? The state that you define for every moment in time is basically that you define the time since the cue. That's the only thing that defines the state of that moment. Even though that moment

it's just like a delay period, any other delay period. There's nothing else externally that tells you what this moment of time is. Now you have to keep--

Paul Middlebrooks
It's like you have to arbitrarily keep track of nothingness.

Vijay Namboodiri
Keep track of nothingness from every possible previous thing.

Paul Middlebrooks
Going back to that for just a second, I thought about this just right after I said, it's infinite. It is infinite in principle. However there are salience differences in cues. It's like not all objects are as shiny as others. Anyway.

Vijay Namboodiri
Yes. Absolutely.

Paul Middlebrooks
You can narrow it. If you add an attention component to it or bottom up, let's say, and ten or top down, that really narrows the search space.

Vijay Namboodiri
The search space down. There are still problems with W, but each of these can be recent food, and then there are some problems that can be some solutions, and then you keep on going. This could be a whole day.

Ali Mohebi
May I just add that the salience also has to be learned.

Paul Middlebrooks
Yes, unless it's like a hawk coming down if you're a mouse.

Vijay Namboodiri
Yes, exactly. Like a loud thud of a door. You don't necessarily have to learn it. I won't get down that argument just yet.

Paul Middlebrooks
Sorry. It was really a side thought. [crosstalk]

Vijay Namboodiri
No, it's absolutely a very good point. These are some of the good points that you need to start to consider as you go down this path. It says that each of them will have its own set of issues. Just sticking to the main thread. Essentially, you are trying to keep track of time from everything. Now, the important thing is, for you to ascribe value for this reward to the state that preceded it, the state that preceded it should be a repeatably identifiable thing. At least in the standard view.

If it's a repeatably identifiable thing, then that means that whatever the neural state is in the brain at the time that just preceded that reward, on the next time that the cue got presented at the same delay, you should get the same state, brain state in the brain. That seems just hard. To get exactly the same type of things repeated reliably trial after trial and keeping track of time that precisely, that seems hard. Now, people will say that you can get TD-type things to actually work without it. That's the whole other set of discussions, too. I'm talking about the standard didactic view.

Paul Middlebrooks
Even in a natural living condition, even if you have the exact same cue, the exact same stimulus, the context is never ever going to be the same. There would never be a perfectly repeatable state.

Vijay Namboodiri
Exactly. Remember, you also have to keep track of time since every cue, because you're trying to-- Obviously, the time since every cue is not going to be repeatably the same.

Paul Middlebrooks
It's hard.

Vijay Namboodiri
It's hard. It's a very hard problem. If you're trying to learn this this way, then it is difficult. It turns out that if you are trying to do this in a prospective way, that when you get any cue, you're trying to learn what is following that cue, you necessarily have to do this thing, of keeping track of time step and the passage of time step by time step. Because what you need to do is, whenever that thing is happening, you need to give more weight to the cue in that you've ascribed more predictive power to the cue.

Then when it's not happening, you need to downweight the predictions to the cu. Because you just don't know when future things are going to happen, you necessarily have to do this time step by time step for everything going to infinity, essentially.

Paul Middlebrooks
Your insight is that if something happens that is awesome or that you want to remember or that you want to repeat, then it makes more sense to start looking back in time to see what was paired with that with high correlation.

Vijay Namboodiri
Exactly.

Paul Middlebrooks
However, then you need to look at an infinite number of things in the past. It is 000001.

Vijay Namboodiri
The problem doesn't entirely go away. There are different aspects of it. In some sense, the core insight of our work is that there's maybe two steps to learning. The first step is to know that something is possibly related to something else.

Paul Middlebrooks
It's like a threshold.

Vijay Namboodiri
Exactly. It crosses some threshold. You're just simply trying to make connections. Does this connection exist? If you believe that that connection exists, then you can go in and then try to understand the properties of that connection. Like what is a temporal delay there? What is the associated probability of reward? That stuff. The critical thing to first know is just that is whether there's a connection. If you're trying to do that, then you do have to keep track of the different cues in memory. Now, here's where the retrospective view actually allows you to solve the salience problem in some sense.

You do need to keep all these cues in memory, but you could-- Like when you're doing these backward sweeps, this is not part of our algorithm, but you could, in principle, add this. When you're doing these backward sweeps, you could do backward sweeps with varying levels of thresholds for the salience if you wanted. You only look at the most salient things as a thing that you could ask for it to. If you don't find anything that's related that way, then you bring down the salience a little bit and then you look back again. You look for a search for things that might be proceeding, but in a salience way, in a salience-dependent way.

Paul Middlebrooks
Still sounds hard.

Vijay Namboodiri
Still sounds hard. The point is that -- I think something of that sort must happen. Not necessarily the retrospective, but something off the way of filtering the salience type things and storing things in memory must happen. The core advantage of going this way is that one, we're using the fact that we have memory and we're storing things in memory. Then that allows you to actually look back for associations. The key critical thing that this gets you in the base framework is that now all the associative components or the associative learning, where you're learning associations, it's all only triggered when events come.

Rather than updating your associations time step by time step, every time step, which is a hard thing for every possible thing, and every possible outcome too. We know that we simplify and talk about just reward learning, we know that animals have different predictions for different types of rewards. We already also know that animals also learn cue to cue relationships. In reality, this is much more complex. For all those associations, you should do time step by time step updates going in the forward direction.

Paul Middlebrooks
I don't think even Ali can do that. I'm not sure.

Vijay Namboodiri
Ali might be the only person.

Paul Middlebrooks
He's talented. I know that.

Vijay Namboodiri
The only animal on Earth that can [crosstalk]

Paul Middlebrooks
He is an animal. I'll give him that.

Vijay Namboodiri
The advantage going in the backward direction is that you can do this in an event-based way. Now you don't need to do this every time step by time step. You just need to do it when something very meaningful happens, you just update backwards and take time stamps of, are things proceeding this in a reliable way.

Paul Middlebrooks
That's at least a few orders of magnitude easier.

Vijay Namboodiri
Exactly. That was the core insights. This doesn't solve the core problem of how do you learn the time delays and stuff, because I've ignored that part here. Here the core aspect is you're just still looking for things whether they precede meaningful things. You, of course, account for the time delays in the memory of what you store.

This algorithm doesn't tell you, "This is the number in terms of what is the delay between the cue reward association." That we pushed aside and said, "That's the second step to learning." You actually have to learn that, but it's much easier to learn that if you already know that this particular cue is very important and this particular reward is very important, and that association is actually where you're trying to measure the time to live between.

Paul Middlebrooks
This is where something like a replay might come in very handy. If you're just replaying offline over and over and over, that's an auto learning system because then you can match. If you're literally replaying from the event, it's a little more feasible, maybe.

Vijay Namboodiri
Exactly. Replay might be a way that you could get this retrospective type thing to work.

Paul Middlebrooks
That's what I meant. It's like reverse replay, I guess, is what you--

Vijay Namboodiri
Exactly. Anyway, the core idea is basically just this. The core advantage is that, if you do this in an event triggered way, when you know something meaningful has happened, you look back, then the number of computations that you have to do might be fewer, at least in terms of the associated computation. You still have to keep track of some other things, and that you need to do time step by time step, but those things are not associative. Those things specifically are the overall rates of different

events. You still need to keep track of overall rates. How often do cues happen in my life, this particular cue. How often does reward happened in my life, this particular reward.

Paul Middlebrooks
Then you also have to do the inverse so that you can do the prospective.

Vijay Namboodiri
Exactly. That prospective now, the advantage is you don't need to compute the prospective every time step. You can just compute the prospective when you need it.

Paul Middlebrooks
When you find the thing that you think is most causally related to the event.

Vijay Namboodiri
Exactly.

Paul Middlebrooks
Please correct me, because I'm sure I'm saying lots of things incorrectly.

Vijay Namboodiri
No, it's all correct. That's the core insight. That's how this view came in. Now, none of this is dopamine apart. This is all independent of dopamine, if you will.

Paul Middlebrooks
This is algorithms.

Vijay Namboodiri
This is algorithms. Even at the algorithmic level, we're not talking about which things you should learn from, etc. This is just simply, if you wanted to learn all pairwise relationships between every possible thing in the world, which is probably hard, then you could do it this way.

Paul Middlebrooks
Not only could you do it this way, this is a better way to do it, the prospective.

Vijay Namboodiri
Exactly. This is the better way to do it, just because it keeps track of events. Now to give credit to the other side, that's not to say that there's no possible way potentially, whereby you could do it in a prospective way. It's just that if you do it in the prospective way, you will necessarily have to do this time-binding stuff and compute that.

Paul Middlebrooks
Let's pause here then. Maybe I'll ask Ali first if you are aware of or versed in the "controversy" of this. Did this, in your view make a big-- Was there a big fuss with Vijay's papers, the recent--

Ali Mohebi
No, no. The first time I heard it, I just purchased it.

Paul Middlebrooks
I meant not from you, but from the field. You pointed me to it, cheerleading it. I know there was not a controversy from you, but are you aware of this controversy in the literature, or? There's never anything with dopamine that there's not controversy in the literature.

Ali Mohebi
Yes.

Paul Middlebrooks
This in particular is a very, specifically-- It's saying is, "Hey, you've all had it backwards."

[crosstalk]

Vijay Namboodiri
Maybe I wanted to just add one sentence before we get to that. One thing that I didn't describe in the algorithmic thing is just what exactly dopamine I think is doing in that way down.

Paul Middlebrooks
We're going to come to it, but let's go ahead and do that.

Vijay Namboodiri
Very quickly, just before the controversy, because then--

Paul Middlebrooks
Oh, that is the controversy there because--

Vijay Namboodiri
Yes, exactly. In addition to this retrospective thing, you also need to do one more thing. This retrospective thing, at this level, I just described something where you learn every possible association. Maybe you don't want to do every possible thing just intuitively, we just described. You only want to do this for meaningful outcomes. The core idea with the dopamine was that there's this additional step to this retrospective thing that filters out and tells you what are the meaningful things that you need to--

Paul Middlebrooks
The dopamine tells you what's meaningful.

Vijay Namboodiri
Exactly. That's the additional step. The controversy is more related to that part and less so related to the retrospective perspective.

Paul Middlebrooks
A reward prediction-- An error is inherently meaningful. What is the controversy? Why is it controversial?

Vijay Namboodiri
Ah. There's a lot of similarities between this idea of what is meaningful and what is--

Paul Middlebrooks
I'll just interject and also say that we're using human language for these terms.

Vijay Namboodiri
Of course, yes.

Paul Middlebrooks
Definitions are slippery, and it gets ridiculous.

Vijay Namboodiri
Exactly. We're also simplifying it to a big extent. Before addressing the controversy, the quick thing to say is that it just turns out that -- Just exactly ask you to get to it, this meaningfulness thing and prediction are a thing just sound somewhat similar. They have a lot. That's exactly the reason why we decided to look into this, which is that it turns out that when you mathematically formalize this, there's a lot of similarities between this meaningfulness type thing and RP.

The idea was, if you've not done experiments that try to look at where the differences are coming, then you wouldn't have known which one is actually-- Then our argument was that maybe all of this evidence for RP might also be consistent with this other thing of the meaningfulness.

Paul Middlebrooks
You should do experiments that test that, perhaps.

Vijay Namboodiri
Maybe, right?

[crosstalk]

Vijay Namboodiri
That's about so.

Ali Mohebi
Vijay, if I may ask a not a softball question. What defines the cue in your interpretation? Because we are under barrage of sensory information. I'm an animal walking around, I'm getting continuous visual input and auditory. How do you keep track of those events? I'm continuously seeing things, a tree, I don't know. In an experimental setup, it's easy to define that, to have four set of cues, and then try to see which one is meaningful. In the real world, how would you?

Vijay Namboodiri
This is a hard problem This is a hard problem that I think basically all these theories just show away and let's say some other smart region takes care of it. My version of that is to say that a cue is something that the higher order sensory regions that define what the objects are, sensory objects. Let's say IT, area IT for visual stuffs defines what cues. Visual cues are those things that are identified by those nerves. Like objects, sensory objects, if you will.

There's learning associated with that, and there's filtering associated with that, and there's complex operations like you take sensory input, and then Ali turning this way versus Ali turning that way is still Ali. That's a hard problem to solve. That problem is typically studied by sensory neuroscientists, and they have come up with reasonable solutions for that. We have shown that there are neurons in the brain that can do that.

Ali Mohebi
The correct answer is that the most important part of the brain, which is the downstream structure or upstream.

Vijay Namboodiri
No upstream.

Ali Mohebi
Yeah ... always. [crosstalk]

Vijay Namboodiri
Upstream structure.

Ali Mohebi
Upstream will take care of it.

Vijay Namboodiri:: Sensory structures will take care of it. If you want the full solution to this whole problem, that's basically saying, how does the brain work? It's essentially impossible to describe it. I think the way that we start to formalize this is we at least start to now get into all the things that we're assuming.

Ali Mohebi
I'm asking because I'm very interested in that because I think that would relate to attention as well.

Vijay Namboodiri
A 100 percent, yes.

Ali Mohebi
How the dopamine system is now involved in attention-

Vijay Namboodiri
Yes, exactly.

Ali Mohebi
-in gating of incoming information. Which I think, we will not get into that, Paul, but one of the most important things about dopamine is its role in attention.

[laughter]

You actually missed that in the 10 commandments.

Vijay Namboodiri
Even in ChatGPT. [crosstalk]

Ali Mohebi
You have to be responsible ChatGPT user.

Paul Middlebrooks
I also intentionally didn't include retrospective. It'll be on the list next year.

Ali Mohebi
I'm going on a tangent, but I think this is another important thing about like, yes.

Paul Middlebrooks
Let me just interject and say one of the things that I have appreciated about your work also is that the way that you've approached it, at least-- The way that you approach it in the literature, which is super helpful, I think, in terms of thinking about how to tackle a problem in general, is that you list out some of the assumptions of TD learning and then point out how they're wrong. That's a powerful way to build your own argument to say, "Here are the holes," and, "Here's how we can fill those holes."

Vijay Namboodiri
I wouldn't use the word wrong per se. I would just say hard, impossible. It's our argument.

Paul Middlebrooks
Incomplete.

Vijay Namboodiri
Or incomplete. Exactly.

Paul Middlebrooks
Good save there, by the way.

Vijay Namboodiri
Anyway, so we get back to the controversy stuff. Ali, you want to take the controversy bit?

Paul Middlebrooks
I was just asking in the beginning, is the controversy due to the fact that so many people's careers or their reputations are at stake and they're just feeling not hurt by it, but defensive, perhaps? Because a lot of controversy begins by powerful people feeling defensive.

Vijay Namboodiri
I know some of these powerful people. I will say that I don't think that that's it alone. Maybe there's a component of it, but I wouldn't say that that's really the main driver. I think the main driver is that extraordinary claims require extraordinary evidence, and the idea that

Paul Middlebrooks
Makes sense. In papers in science.

Vijay Namboodiri
Yes, papers in science, exactly. The dopamine story and the TD story, was 25 years old at the time that we had our paper published, 1997 to 2022. Something that has lived on for 25 years with thousands and thousands of papers supporting the idea--

Paul Middlebrooks
It's bound to be wrong. Makes a lot of sense.

[laughter]

Vijay Namboodiri
No. It's not going to be over-turned with one paper.

Ali Mohebi
Great. I think what was beautiful about this work, it's not my work, so I can just say nice thing about it, was that it looked at, like you mentioned, Paul, looked at predictions of TDRL, and some of them were not explained by actual data that we will get into it. Then came up with an idea, this retrospective learning, that could explain both TDRL predictions and places where it could fail. I think that was the beauty of it. That is, I think, the main controversy, maybe.

Paul Middlebrooks
How is that controversial?

Ali Mohebi
There are specific predictions that Vijay may go through some of them. There's, I don't know, 13 in that paper, 14. It's a huge number. That predictions where TDRL fails to make predictions about what happens with the real animal.

Paul Middlebrooks
Why would that be controversy as opposed to progress?

Vijay Namboodiri
It's a very good question. Why is that controversial? Here we get into the heart of TDRL and what TDRL is. TDRL, the way that I describe it, I said there's a box, and that's where the box in which the algorithm lives. There are some inputs that go into it, and there's some output. Now, which of these is TDRL? TDRL is really the box where the computations are happening. The things that go into it are not TDRL per se.

Paul Middlebrooks
Oh, OK. Are you saying you're doing TDRL?

Vijay Namboodiri
Exactly. The argument is, if you're doing TDRL, that's not a single thing, because the input that goes into the algorithm can be many different things.

Paul Middlebrooks
That's your argument.

Vijay Namboodiri
No, no. It's the controversy argument. It's the alternative argument. It's the argument from the TDRL folks as to why this should not make us talk about TDRL.

Paul Middlebrooks
Because it doesn't matter what's going into the box. It's what the box is doing.

Vijay Namboodiri
Yes. Because of that, the problem that comes down to it. If I'm supporting TDRL, this is the counterargument that I have. You have convincingly demonstrated that this TDRL box with a specific set of inputs is wrong, and those predictions are wrong. How do that this TDRL box with a different set of inputs is wrong? This is the problem.

It becomes almost philosophical, and it's a good thing to get into. The argument is this, basically, that now, when it comes to-- These things that are the inputs are states, it's a formal term that I use. The way that you define states is inherently ambiguous when it comes to time delays, because there's no objective thing in the world that tells you exactly what state you're in. You can define that in many different ways. Now, people have defined them concretely in many different ways.

We did look at the concrete things that people did define it as, and showed that none of those concrete things actually fit with the data. I think that, from our perspective, we argue, at least the published concrete things we've looked at, and those don't fit the data. Now, the counter argument is, but that's not to say that you've not done a good job of looking for things that are possibly the different state inputs. You could have had different state inputs where you could have looked for this.

Then you could have also assumed different parameters of the TDRL algorithm and how sensitive they are to the TDRL. That, obviously, is another aspect. Those are the defined free parameters. Within the box, there are free parameters. There are many different free parameters within the box. Then there's the undefined things that are outside the box, which are the inputs, which technically, those are infinite dimensional because they could be anything, basically.

Paul Middlebrooks
Yes, but my guess is you would say that these are valid arguments.

Vijay Namboodiri
Yes. I think that's a valid argument. To actually argue against TDRL as a family, you would need to show that any possible inputs with this algorithm would not fit. My question to that is, how do you ever show that?

Paul Middlebrooks
You need to keep track of all prior possible inputs as you move forward in time without knowing what the reward is. Every possible input cue, sorry, I'm trying to bring it back to you. Retrospective.

Vijay Namboodiri
Exactly. There may be different ways to keep it. You can store things in memory in many different ways. Which way of storing things in memory should you be considering. This is where I think that the TDRL field has a-- There's a little bit of a philosophical problem.

That relates to falsifiability as a concept in science and to the extent to which that is a core concept that we still stick by for scientific progress, for theories. In my view, TDRL is a framework, it's not a theory. If you were to ask folks that really defend it, I have done it, to give me a list of predictions, here are exactly the predictions where if you were to find these, I am telling-- That's it. There's no way that TDRL can rescue it. There's no clear answer to that question.

How can you prove this theory wrong? Give me a way whereby you're 100 percent certain that if this was true, this theory is wrong. I don't know the answer yet. Until you know that, you can't fully falsify a framework. That's a valid point. To me, then the flip side to that is, if you cannot falsify a theory, how much does that aid in our understanding anyways? Don't you want to have a theory that is very concrete, where you can falsify it. If you can't do that, then that's a framework, and it's useful. I'm not saying that it's useless. It's extremely useful as a way to think through things once you get the results and to come up with explanations for it. It's all postdoc explanations.

Paul Middlebrooks
It's not the way that science should be done as Popper envisioned it, or as Popper argued. [crosstalk]

Vijay Namboodiri
At least definitely as, it's definitely not the Popperian view. This is where we get into the philosophy of science debate. A lot of the controversy in this field actually comes down to this particular problem. That's a lot of it.

Paul Middlebrooks
What's actually pretty beautiful about that is that you arrived at that through your algorithmic approach.

Vijay Namboodiri
Yes. It came to it just from looking at the assumptions. Looking at the assumptions and which ones can be defended versus

not. That's when you realize that some of the assumptions actually are so flexible that you have so many degrees of freedom there. Then how exactly do you test?

Paul Middlebrooks
See, I think, however, that this same argument, look at the assumptions can be applied to almost any topic, in at least neuroscience and biology. Earlier when you said that you just had an intuition that something was missing about TDLR, and then it took you 13 years to be able to vocalize what that intuition is. I've had an intuition that there's something very wrong with almost all of neuroscience. That means it's going to take me 100 years to even get close to vocalizing it.

Vijay Namboodiri
That's the tough part.

Paul Middlebrooks
I think that yours is a more tractable intuition. I'm envious of that.

Vijay Namboodiri
This also goes back to some of the Randy stuff, maybe with memory. Randy's obviously been arguing that there's a lot that's wrong with neuroscience and all of how we think about memory is wrong. To actually be fair to folks on the opposing side, where on the TDLR side, there's one more aspect of this that is controversial. The second aspect of this that is controversial is that though the retrospective to prospective conversion, that is fully within the view of the standard TDRL perspective view, that it's all long run sums of things, events and discounted sums. It's all the same view.

That part is less controversial. I think the controversial bit is that the way that we handle the definition of what is meaningful, is you can't write it out as a solution to a problem that you define. This is a problem that you're trying to solve, and this is exactly the solution. It's more of an intuitive type approach to say, "There's a set of core intuitions that whatever meaningfulness is should abide by. Here's a way to mathematically formalize that."

Paul Middlebrooks
You have to operationally define meaning then. Because if not, then we are back into the unfalsifiability problem.

Vijay Namboodiri
Exactly. We have taken a very concrete way of defining it. The problem with that is that is not-- There are elements of this that have a lot of intuitive and there's an intuitive derivation for it. There's not a problem statement where we can say, "If you want to maximize this objective criterion, this is exactly the way that you would define it."

Paul Middlebrooks
I see.

Vijay Namboodiri
That's a valid concern.

Paul Middlebrooks
There's not a theoretical formal solution.

Vijay Namboodiri
Normative solution.

Paul Middlebrooks
Normative, OK.

Vijay Namboodiri
Yes. Normative solution. That's a fully valid concern. In that sense, I think of this as the first step towards going in that direction. We're working on this. Who knows if we will end up coming up with something that is a fully normative sort of thing that-- My suspicion is that if we work through that, the eventual solution that we come up with is not going to be exactly like the anchor prediction. It would have a lot of the intuitive features of it.

That's why a lot of the things that we test are those intuitive aspects of things. Not exactly how much up or down. We're not looking for dopamine is 20 percent up from or 10 percent up from this than what it should be. We're looking at, dopamine should be higher versus lower. It should be positive versus negative. That's the level of predictions that we're testing at. There's a completely different story from this retrospective view that is beyond the Jiang *et al* paper. We're finding other things that seem to be coming directly from the retrospective view, where those happen to be true.

Those are experiments that we did explicitly to falsify our model. It turns out that it wasn't falsified. Those things exist. At least so far in our attempts we have looked at some really wacky predictions of this retrospective view. Those wacky predictions seem to generally hold up, so far.

Paul Middlebrooks
Have not been falsified. When you tweet about something like that, do you often use that raised hands emoji? Like, "That's right, but we're still not wrong."

Vijay Namboodiri
No, it's a tricky one It's a tricky one, because this is not a common approach in neuroscience.

Paul Middlebrooks
What's not trying to falsify your own--

Vijay Namboodiri
Trying to falsify your own.

Paul Middlebrooks
It's supposed to be.

Vijay Namboodiri
It's supposed to be, but unfortunately, it's not. It's a tricky one, because when we say that we end up finding results that were consistent with our prediction, wacky they may be, the way that, at least some members of the audience, of the neuroscientific audience, take it as, you just looked to prove your theory, and you're claiming this as proof for your theory.

Paul Middlebrooks
There's no way to convince them otherwise.

Vijay Namboodiri
Yes, how do you convince them otherwise? I can't give you the full thread of the thinking that I had.

Paul Middlebrooks
You need to prospectively write out all your future thoughts so you can--

Vijay Namboodiri
Exactly. What I'll say is that we haven't discussed that. Some of the things that we have found are very inspired by Randy's stuff, but not exactly Randy's stuff. There are key differences from Randy's stuff too, but very inspired by them.

Some of the things that we find, if my job is to try to find evidence supporting this theory, let me just say that will not be the thing that I'll go and try to test. Because my core intuitions are that there's no way that those some of those predictions could be true. I was starting out. If I wanted to just find some evidence consistent with the framework, it's way easier to look at things that, this thing could fit and also many other things could fit. That at least I could say, "That thing is consistent with our stuff."

It's not uniquely consistent with our stuff, but at least it's consistent with our stuff. If I wanted to just build out evidence, that would be the approach that I would do in a purely strategic way. If you want to test these things, I would be hard-pressed to try to come up with a way to look at the experiments that we've done with not an intention of, at least an expectation, that would falsify our results.

Ali Mohebi
May I say that-- Oh, sorry. The fact that we are just having this conversation, I'm biased, but would speak to the maturity of the dopamine learning field because the majority of neuroscience, we don't even have these frameworks. Here, I think we have a framework and now we can start questioning some of the assumptions. I think we've come a long way. Just say that.

Vijay Namboodiri
100 percent.

Paul Middlebrooks
It's, in general, a pretty friendly community, the dopamine-- Not without controversy.

Vijay Namboodiri
There's definitely controversy. I think people talk. People talk respectfully to each other. That's all you can hope for.

Paul Middlebrooks
Thank you so much for spending so much time with me and for the careful elucidation of the history. I didn't really know that we were going to get such a history lesson. I learned something. I'm not sure I know what dopamine does, but I think this is super valuable for people who do think they know what dopamine does. It's not going to change popular science outlets who will still say, "Oh, I got a dopamine hit with that piece of cheesecake because it made me happy," the happiness. At least, hopefully, it'll reach some people and they'll realize that, "Man, the story is hard." There are lots of different ways of approaching it. I really appreciate the simplicity and the elegance of the solution that you have come up with, Vijay.

Vijay Namboodiri
Thank you.

Paul Middlebrooks
I appreciate it.

Vijay Namboodiri
There's one thing that maybe I add, if you have time still. Just very quickly. One thing that I wanted to say is there's both a positive message and a negative message, I guess, in this trajectory. In that all of the debate about dopamine that we've been having, I have a big controversy in the field, are not about all the details that Ali was talking about, all the variability across different regions, the axonal regulation of dopamine release, et cetera. It's about the simple thing, this one-dimensional signal, what does it represent. We're throwing away a bunch of stuff. Even that simple one-dimensional stuff, we haven't yet settled as a field. To me, the negative side of that is that shows you that neuroscience as a field still has a lot of maturing to do.

Paul Middlebrooks
It's young.

Vijay Namboodiri
It's very young. The fact that maybe one of the most investigated questions where we're talking about just, we're literally a one-dimensional signal is still not settled, actually says that other things probably will also need to have these moments. Now, the positive side of this is that, this, I think, is the sign of a field that is actually starting to get to the direction where you're starting to have these debates. It's a thing that, in general, we need across all of neuroscience, across all the different topics in neuroscience. The dopamine field is leading in that sense.

Paul Middlebrooks
The video is going to end with celebrating our incompetence, I like it, in neuroscience. That's not what you said.

Vijay Namboodiri
The main point that I'm making is that it's a very exciting time to be a neuroscientist in that even though we've collected a lot of data, I think as a conceptual field there's still so much to be done.

Paul Middlebrooks
I agree.

Ali Mohebi
My last line would be that we might not know what dopamine does and what dopamine is, but we almost definitely know that dopamine is not pleasure. One takeaway, dopamine is not the pleasure signal, and maximizing your dopamine is not necessarily a good thing.

Vijay Namboodiri
I do second that.

Paul Middlebrooks
Thanks, guys for coming on again. Good luck to you both.

Vijay Namboodiri
Yes, thank you so much, Paul. That was great.

Ali Mohebi
Thanks for having me. It was lovely. 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 why invite to the podcast. Go to braininspired.co to learn more. You're hearing music by The New Year. Find them at thenewyear.net. Thank you for your support. See you next time.

[music by New Year]
