Rationally Speaking #212: Ed Boyden on “How to invent game-changing technologies”

Julia Galef: Welcome to Rationally Speaking, the podcast where we explore the borderlands between reason and nonsense. I'm your host, Julia Galef, and I'm here with today's guest, Ed Boyden.

Ed is a neuroscientist at MIT who has pioneered two technologies that have been/and are about to be really transformative for what we're able to do in neuroscience: Optogenetics, which was called arguably the most important technical advance in neuroscience in the past 40 years, and then more recently, expansion microscopy.

Ed has won a bunch of awards including the prestigious Brain Prize. Perhaps my favorite award that he's won was mentioned in an article abut him that I happened to read — it’s a certificate in his office for “Mr. Most Likely To be Late because he is teaching students how to build a microscope.”

Ed, welcome to the show.

Ed Boyden: It's great to be here, Julia.

Julia Galef: So, I have a lot that I want to cover with you. First, I want to talk about why optogenetics and expansion microscopy are so revolutionary. And then I want to jump meta - which will not be a surprise to any of my listeners - and talk a little bit about:

What's your process for making major innovations like that? What heuristics are you following for scientific research, and just for thinking in general, that are so effective as to produce two major breakthroughs in basically a decade? You're only in your 30s. And then also, how did you decide that these were the right problems to work on?

So, it's a lot to cover. Let's jump in. I guess first, I want to ask you: Would you say that you could name a single goal or problem that your whole career is aimed at solving? Or is that too hard?

Ed Boyden: Well, for me it's been an obsession since I was a kid to discover scientific principles that might help shed light on some of the deep questions that have been the domain of philosophy. You know, why are we here? What's the nature of happiness? What's the meaning of life? When I was a child, I was pretty obsessed with these kinds of questions. And I really wanted to work on these problems, but from a scientific standpoint.

So I started college pretty young. I worked in a group, Paul Braterman's group, which was working on the origins of life. You create life in the laboratory. And although our efforts did not yield life - that's a really hard problem - it really taught me how to think about problems at the border of philosophy and science.

And I went on to MIT where I worked in quantum computing, another area where it's sort of at the boundary of physics and science and philosophy, and the mysterious questions of what's the nature of reality. And finally, I went on to study the brain.

So for me, it's been one long arc of how do we understand the world and existence in scientific terms? Can we aim for a certain kind of enlightenment, if you will?

Julia Galef: So this “single question” basically reminds me of the single question that the giant supercomputer in "The Hitchhiker's Guide to the Galaxy" was designed to solve, which is, "The answer to life, the universe and everything."

Ed Boyden: Mm-hmm. What my experiences with the origins of life and quantum computing taught me was, you also have to pick the right problem to work on.

And those are both really hard problems. You could argue, we still don't have quantum computers that are universally applicable. But, it was just sheer dumb luck that all my experiences in chemistry and in physics prepared me to tackle the brain, which in some ways is not as difficult a problem. I regard it as problem where we want to solve the brain in terms of chemistry and physics.

So, it's a problem where we kind of know where we want to go. We just have to get there, and to get there requires new technology.

Julia Galef: Well, how would you characterize where we want to go? Do you just mean, understanding the brain, from the smallest pieces up?

Ed Boyden: Well, my dream is that we could make a biophysically realistic computer model. So a piece of software, let's say, that could simulate a thought or an emotion that is also — ideally, there's no guarantee that it will be, but ideally — human understandable. So, we would really understand what a thought is, what an emotion is.

And so my dream is that this would help us become more enlightened as a species, because we would know why we do what we do, and we would know why we feel the way we feel, because we would be able to peer inside and see the mechanisms of that.

I also think that this means that we need certain technologies. We need to see what's going on inside the brain. We have to be able to make maps of the brain, and we have to be able to control what's going on in the brain. And so as a byproduct of that quest, my hope is that we can develop all sorts of treatments for diseases which are ... almost all brain diseases, currently, they're intractable.

Julia Galef: Yeah, so maybe now's a good point to ask, what was the course that led you to discovering optogenetics? Was there a specific question that you were pursuing? And what is it?

Ed Boyden: Well, I think that all the great questions in biology at some level are kind of clear. We want to see what's going on inside the brain and control what's going on inside the brain. When I was a first year graduate student, I had joined Richard Chen's group, and also Jennifer Raymond's group at Stanford. And in the Chen lab, there was another student, Karl Deisseroth, who was finishing medical school at the time, and we just started brainstorming:

How would you control a brain circuit to make it do what you wanted it to do? Could you fix a brain disorder by canceling out pathological activity?

And so we just started going through all the laws of physics. Could you use magnetism? Could you use mechanical perturbation?

And we decided that light would be the best way to do it, if you could, because light, of course, is as fast as anything ever gets, and you could aim it very precisely. And so the trick then became, how can you make a brain cell sense light? Because of course, most brain cells don't.

I stumbled across this paper from a group in Japan that had been published the year before, and they had found that some of these natural molecules that convert light into electrical signals, they just happen to have the right kinds of properties that might lend them to express in brain cells. So we started collecting these molecules from colleagues, and we were off to the races.

Julia Galef: So what is possible with optogenetics now, that wasn't possible before?

Ed Boyden: Well, we've now given the molecules out to literally thousands of research groups all over the world. People are using the molecules to activate sets of brain cells.

So for example, if you can drive a set of brain cells, you could figure out what kinds of behaviors or pathologies or processes can those

brain cells trigger? And if you can delete those brain cells, you could ask the question, “What processes, or brain functions, or computations, or diseases, are those brain cells needed for?”

So just as a couple of examples: people have found that by activating certain brain cells deep, deep in the brains of mice, you can actually trigger them to become aggressive, or even violent, to whatever is next to them. Even a rubber glove, frankly.

By triggering certain brain cells, you can figure out if they can cancel out an epileptic seizure.

And by turning off brain cells, you can figure out what parts of the brain might be involved with storing a memory. If you can delete a part of the brain just for a few seconds, and memory recall is suppressed, maybe you could start to hunt down where memories are actually encoded in the brain.

Julia Galef: I have to admit, it surprises me to hear that by messing with a single cell in a mouse's brain, you could cause it to become aggressive. That doesn't fit my crude, naïve model of how brains and emotions work. I wouldn't have thought, "This is the cell for aggression." Does that-

Ed Boyden: Just to be precise, each of these studies was manipulating sets of thousands, maybe millions of cells.

Julia Galef: Ah, ok.

Ed Boyden: They would put the molecule that encodes, the molecule that makes the neurons light sensitive into a population of brain cells. You can do that using all sorts of different tricks from fields like gene therapy, for example, because these molecules are genetically encoded. And then you can shine light on a large chunk of the brain, and only the small subset of the brain cells that manufacture this little light activated solar panel, if you will, only they will be sensitive to light.

Julia Galef: Okay, got it. So basically, at a high level, the significance of optogenetics is just that it gives us a much more fine grained tool for connecting activity in the brain just at the cellular level, the level of single neurons, to what we see brains doing at a higher level, the emotions or the behaviors that are generated by the brain. And we couldn't do that before, because we didn't have that high, extremely precise ability to observe what was happening in the brain.

Ed Boyden: Yeah. Eventually, our goal is to be able to control a large population of brain cells at a single cell resolution. And we’ve already begun

that process. What some of my colleagues like to call “playing the brain like a piano.” Could you enter a distinct code into each brain cell?

But even just activating a whole set of cells can be very useful. I'll give you an example. My colleague, Li-Huei Tsai at MIT, directs the Picower Institute at MIT, and she used some of our optogenetic tools to drive a certain fairly rare population of brain cells. These are called parvalbumin-positive interneurons. And these are basically little cells that shut down other cells. But they're wired in a weird pattern that lets them resonate. What that means is, if you drive them in a certain frequency, they will prefer to be driven at that frequency.

So her group had been studying these, and found out that they resonate at a certain frequency of 40 Hz, 40 times a second. And so their group went on to find, using our technologies, that if you drive these interneurons, these parvalbumin-positive cells at 40 times a second in Alzheimer's model mice - these are mice genetically engineered to have some of the genes that cause, or mutations that cause Alzheimer's in humans. So they're not perfect models of human Alzheimer's, of course, but you gotta start somewhere.

... Anyway, if you drive these cells, interestingly, the brain's immune system turns on. And the amyloid plaques and some of the other molecular hallmarks of Alzheimer's in these model mice, they go down.

And we went on to — in our collaboration, Emery Brown, another professor at MIT contributed — the collaborative team went on to find that you could simulate that pattern of brain activity through, effectively, movies. Blinking lights that you would see through your eyes. No optogenetics, no other technology than basically just watching a movie through your eyes.

And amazingly, these mice also got better. So Li-Huei and I finally co-founded a company, Cognito Therapeutics, where the goal is simply to build movies to treat Alzheimer's.

Julia Galef: Whoa.

Ed Boyden: And so we've begun human trials. Many mouse data sets do not translate to humans, so we have to wait until we see how the trials go, of course. But we're very excited about this possibility of a non-invasive way of treating an otherwise untreatable, in terms of stopping the progression of the disease, the condition.

Julia Galef: That is extremely exciting. And I think if it were me who had invented optogenetics, I probably just would have run with that and been like... all right, well, that's my big innovation for my career. That's what I'm gonna be doing from now on.

But instead, you went on and invented expansion microscopy. Why was that the next step for you? And also, what is it?

Ed Boyden: Yeah. Well you know, optogenetics was very powerful for perturbing brain circuits and figuring out what they could trigger. But how do you know what to perturb? Ideally we would have a map of the brain.

It's not a great metaphor, but I'll use it anyway, because I think it can frame the problem a little bit: Suppose I want to understand and reprogram a computer. So ideally, I would have a map of the computer. I would know what the chips do, and how they're wired, and how the wires are configured inside the chips. That's where the expansion microscopy that I'll tell you about in a second comes in.

Ideally, we can reprogram the computer with a keyboard, right? That's analogous to the optogenetics. And then thirdly, and this is something that a lot of people in my group are working on now, ideally, we can watch the brain in action. We can eavesdrop on the high speed computations as they occur, like having a monitor or a probe to watch the computer in action.

So expansion microscopy, we hope, will help us rapidly get the maps of the brain.

Now the problem here is that the brain is a really multi-scale system. And what do I mean by that? Well, brain cells in the human brain are enormous. A single brain cell in your brain might extend for centimeters in spatial extent. They're by far amongst the biggest cells in the entire body.

Yet, the connections between brain cells these are called synapses that exchange chemical transmitters - those are nano scale. And they're also full of molecules, like transmitters, interceptors and so forth, that are themselves nano scale, and often organized with nano scale precision.

So how the heck can you see a large scale system like a brain circuit, and many of these cells wired up in a specific configuration, without losing sight of the molecules and the connections?

And so the time that we started this project, we were really trying to figure out, how could you image a large scale system with molecular

precision? And at the time, there really wasn't a great way of doing that.

With electron microscopy, you can see with very fine precision, but it's hard to analyze what the molecules really are. And then, with so called super-resolution techniques, some of these - their inventors won the Nobel prize in chemistry a couple of years ago - you can see molecular information with fine detail, but it's hard to scale up to large 3D structures.

And so we were thinking - why don't we just do the opposite of everybody? If they're all zooming in , why don't we blow it up?

And so two really great grad students, Fei Chen And Paul Tillberg, we were all trying to figure out, how would you blow up a brain to make it a hundred times bigger or a thousand times bigger or a million times bigger? And then you could image it with really cheap optics. We even have some collaborators who are trying to figure out how to modify inexpensive webcams, so that you could actually do potentially molecular imaging with dirt cheap cameras. Because we're gonna take these biological systems and physically blow them up.

And the basic chemistry is intriguing. We take these specimens of preserved brain tissue - this doesn't work on living brains, obviously - and we infused them with a very even chemical mesh, that's a lot like the stuff in baby diapers; a swellable polymer. And then we add water.

And if we chemically treat the tissue to make it very soft, adding the water will cause the baby diaper material to swell. And if we've treated the tissue just right, the swelling of the baby diaper material will bring along all the biomolecules as well, physically blowing up the brain tissue, until you could image the finest connections even with inexpensive optics.

Julia Galef: Wow. That's a lot to process. Does it just happen naturally that in that expansion process, the basic shapes and spatial relations of the brain structures are preserved as they blow up? Or is it more chaotic? I could imagine some parts expanding slightly more, some slightly less, and you'd just get a messier structure when it's done.

Ed Boyden: We designed the polymer, and also the softening process, to preserve as much biological information as possible. It's not perfectly isotropic. We get a few percent distortion over length scales of tens to hundreds of microns. But that's still good enough to preserve the vast, vast majority of the kinds of biological information that we need.

The design is important for two reasons. One is that we synthesize the baby diaper-like polymer inside the tissue very densely and very evenly. So the spacing between these polymer chains is about the size of a biomolecule itself.

And then secondly, we soften the tissue up with a process that tries to really saturate the mechanical homogenization of the tissue. So if one tissue's tough, and one tissue's soft, and another tissue's in-between... in all three cases, we're trying to make the chemical process of softening - which involves detergents and heat and sometimes enzymes - we really want to soften all the tissues until they're equally soft. And that allows us then to separate the molecules from each other in a fairly even way, although as I mentioned earlier, it's not perfectly even.

Julia Galef: Right. Now I really wish you had ... The term expansion microscopy is a good one, but I wish you'd used the word diaper in there somehow, like Pampersization or expansion diaperoscopy or something.

Ed Boyden: Well, this idea that the idea came from baby diapers is a little bit of an urban legend. It was really two things. One was ... The initial motivation for the idea is that I really wanted a way to de-crowd molecules from each other so we had rom around them to label them in interesting ways.

And then the other influence was the papers of Toyoichi Tanaka, an MIT physicist who really worked out the physics of responsive polymers or smart gels. And so in our readings, these papers shined very prominently as an amazing example of physics that could be applied potentially to novel arenas.

Julia Galef: Well I'm sure that origin story will be compelling to other scientists. But if you ever write a popular science book about your discoveries, I'm just telling you now that your publisher is gonna want you to tell a story about how you were babysitting for your friend's one-year-old, and you looked at the diaper expand, and then you had a Eureka moment. Just letting you know, that's what's gonna happen.

Ed Boyden: Well the funny thing is, if you look at optogenetics, the class of molecules that we use are these molecules that serve photosensory and photoenergetic roles for microbes. They convert light to electricity. And those molecules were originally discovered in 1971.

And then if you look at expansion microscopy, when Tanaka was working out the polymer physics of this class of materials, that was also in the late 1970s, early 1980s.

So, I think an idea that emerges from this, and it's actually something that I actively practice, is that the great ideas are buried in the past. You just have to dig enough to find out the things that have been forgotten.

And earlier, we were mentioning, what are the strategies that you could use to develop new ideas one after the other after the other? One of the things that I often think about is: if there's a hot field with a lot of activity, you know what? I might not need to know about that. I might instead prefer to go after the forgotten things, where maybe they were even perceived as failures at the time, but now the world's different.

We have gene therapy. We have supercomputers. The world's a very different place than 50 years ago. And so maybe an idea that languished or that even wasn't had back then, if we invigorate it and make it reality today, or come up with the idea today, it could be transformation.

Julia Galef: Okay, that's really cool. But is there an additional piece that's needed beyond looking for ... I'm sure there are many additional pieces that are needed, including some dash of luck.

But the process that you're following, where you're looking at ideas from the past that weren't revolutionary in their time, or maybe seemed to have failed, but might be relevant now to today's problems... what is going through your mind as you're reading about or thinking about those old ideas? Are you just saying to yourself, is there some way I could connect this to what I'm working on now? Or are you asking yourself more systematic questions than that?

Ed Boyden: I'm a big believer in having huge problems in mind, and then thinking backwards from those problems. And then, if you stumble across something, or if you are deliberately observing all the different disciplines of science and engineering to think about how you could apply it to those problems.

In my opinion, there are two really big problems in biology today. We want to see everything going on in the body, and we want to control everything going on in the body. And so I'm always trying to figure out - as I see old papers, or bump into random people and learn what they're doing, or meditating or dreaming up crazy ideas - I'm always trying to figure out, how do we check those with one of these two big goals, seeing and controlling?

I think in the years to come, a third and fourth set of problems are gonna emerge as we get better and better at seeing and controlling.

A third thing I would like to do is be able to simulate everything. As of course, what we can simulate, we can test and understand in ways that are difficult to do with only regular matter.

And then fourth, of course, is can we build? But I think for some areas like the brain, we're not really ready to plunge wholeheartedly into only simulating and building, because again, we need to map and perturb to begin with.

Julia Galef: I have another question about your insight generating process. But first, I'd meant to ask about the connection between optogenetics and expansion microscopy.

So you said with expansion microscopy, you're working with a non-living brain, because you're blowing it up, so that wouldn't be very good on a living brain. And you also said that it gives you a map of the brain's structure, which you can then combine with optogenetics, which helps you perturb and observe parts of the brain on a very fine scale.

But, it seemed to me just from listening to you that those two pieces were disconnected. That you can create a map of one brain that's not living, and you can perturb a different brain that is living - but does that map of one brain help you with the other brain? Does that make sense?

Ed Boyden: I need to tell you about a third technology, and then I can answer your question. I mean, just very briefly.

So the third technology suite that I think we really need is a way to watch the brain's electrical activity while it's alive. Suppose that every neuron would blink when it was active. And so, as a thought or a decision or a feeling or sensation or action was occurring, you could watch that dynamical and even rhythmic set of activities that percolate throughout the brain during that process.

And we just recently published, led by two postdocs in my group, Kyrill Khilkevich and Erica Young, a fluorescent voltage indicator. And we were not the first to build such a thing. But we developed an evolution strategy to make one that was very, very fast, stable, high fidelity and so forth. So it's a very all-around good molecule that we use routinely in house.

Now, the ideal experiment, I think, would go something like the following: We would use this fluorescent voltage indicator so that neurons would blink while they're active. And in the living brain, with the microscope, we would watch activity patterns during something interesting, like a decision or an emotion.

Then, we would use optogenetics to perturb different neurons and see how each neuron, ideally with single cell precision, influences the ongoing dynamics of the rest of the brain. So that would give us a bit of causal information. It would tell us the influence of one thing on another.

And then finally, when the experiment was done, we would use expansion microscopy to expand the preserved brain - ideally the very same one we just watched while it was doing a behavior - and to look at the wiring.

And so here's what we would do next in an ideal world. We would have this map of the brain, the wiring of the brain, obtained by expansion microscopy - and it would be very informative, but it wouldn't contain all the information. We might not know where all the transmitters, all the receptors, all the ion channels are, right?

Julia Galef: Yeah.

Ed Boyden: Because in part, the list is still evolving as we learn more and more about how brain functions emerge. So how do we get the information? Well maybe we could use machine learning or other computer science concepts to infer the hidden variables of that map, by looking at the dynamic data collected through voltage imaging and optogenetic perturbation.

So that's how these three kinds of technologies potentially could link up, and maybe even generate models of how a brain circuit is computing.

Julia Galef: Okay, let me try to understand this with a different approach. One of the simplest organisms that neuroscientists and computer scientists tend to be interested in, because it's so simple, is the worm c. Elegans, which has only about 300 neurons and only about 6000 connections between the neurons.

I guess it's unclear to me why we can't just simulate a full c. Elegans now, if we have the ability to do the things you're talking about, like see the roadmap, perturb the different parts of the brain, et cetera. It still seems like there's something missing, some kind of insight or understanding that we don't yet have that prevents us from being able to simulate the worm.

Unless I'm wrong, and we can simulate the worm?

Ed Boyden: Well maybe there's some insight we're lacking. There's always that possibility in science. But our group is mostly known as a tools group. But recently, we've started to actually do some basic science,

and we actually do have a couple of grad students in my group who are focused on applying these tools to c. Elegans.

So, who knows? We'd have to get the data, of course, and it could take a while to get the data. We have to build the microscopes. We have to make the worms, and we're collaborating with c. Elegans groups like Steve Flavell's group at MIT has been a great collaborator with us and helped very influentially, for example, on our voltage imaging paper. The c. Elegans community's a very warm and welcoming one.

Julia Galef: That's nice to hear.

Ed Boyden: And so, yeah, we're gearing up to think of how you could actually do that experiment. And you're right, the experiment hasn't been done yet. But somebody has to apply those technologies to it, and our hope is that we could make a contribution there.

Julia Galef: But would it be correct to say that it should be possible, logically, given the tools that you have developed, and the way that we think they work? It should be possible - we expect this to work?

Ed Boyden: Well, as a scientist I can tell you that there's always unknown unknowns.

Julia Galef: Sure, sure, sure.

Ed Boyden: I think what we can do is to do the experiment. So suppose you could image the voltage throughout a worm, and then we could use optogenetics to perturb those neurons and measure their functional strength of connection... and then at the end of the day, we could expand the worm and make a map, and then use computer science to try to stitch those data sets together into a model... we could certainly try that.

Julia Galef: Okay.

Ed Boyden: But the thing about neuroscience is that there's a lot of unknowns that are unexplored. You might have heard about this paper from ... there was one from a group in Utah and one in Massachusetts, where they were finding that when neurons undergo activity, the kind of activity that happens during learning, they manufacture virus-like particles that look a lot like HIV, the very same virus that causes AIDS.

Julia Galef: Yeah.

Ed Boyden: And those virus-like particles might be able to carry genetic payloads from one brain cell to another. And that was only reported a couple months ago.

Another series of studies over the past, really past couple of decades, has been revealing that neurons can make cannabinoids, molecules that are not unlike the molecules in marijuana.

So the number of mechanisms in the brain, the number of genes in the genome, the number of ways that cells compute, those are still emerging.

So I think that we can try, and we can do the experiments. At the very worst, the data will still be immensely informative about what kinds of mechanisms exist in the brain. But it doesn't guarantee that a complete model will be solved in the first try.

Julia Galef: And at what point would we be able to do this to an organism that's more complex than c. elegans? Like, I don't know, a mouse or something?

I guess what I'm interested in is less about the time scale and more about what would be required? Is it just the same technology, and we just would need to pour in a ton more money or time into building it? Or would it be a different technology?

Ed Boyden: Well I think the hardest part in the long run - not in the short run, 'cause right now, everything of course, is hard - but the hardest part in the long run, I think, is the imaging of the dynamics.

So, people have already imaged whole worms. In fact, we had a paper with Alipasha Vaziri a couple of years ago where we were able to do such things. That was before we developed voltage imaging of the worm, though. A larval zebra fish or a fruit fly would have 100,000 neurons, so a couple orders of magnitude bigger. A mouse would be a thousand times bigger than that, and a human would be a thousand times bigger than the mouse.

Julia Galef: I was surprised that we only have a thousand times more neurons than a mouse. I would have expected more than that. But maybe that's my human pride talking.

Ed Boyden: I think those are all ballpark numbers. But I think people commonly quote a mouse as having about a hundred million neurons, and I believe a human has about, people commonly say about 80 billion neurons. That's about a thousand, factor of a thousand, roughly.

Julia Galef: I guess I'm somewhat comforted by the fact that the number of connections between neurons goes up more than linearly as we go from mouse to human. We still retain-

Ed Boyden: I actually don't know that. Is it really a lot more than linearly?

Julia Galef: Is that not true?

Ed Boyden: I don't know, actually.

Julia Galef: I guess I was just assuming, just by combinatorics. But never mind. I'm not the neuroscientist here. I'm just trying to assuage our human ego.

Ed Boyden: Well some neurons, they have only a handful of connections. There are some neurons that I think have only three or four connections in the brain. I think we call granule cells of the cerebellum. Yeah, so it really depends on the cell types too. It's a complex question.

Julia Galef: Oh, okay. So, I wanted to go back to the thread that we were on five, ten minutes ago about your process for making these discoveries. It seems like ... I guess I'm curious why other people haven't made these discoveries earlier? Since, as you say, some of the important building blocks have been there for decades.

You named a couple of things. You named: people don't tend to revisit old insights from the past, or old discoveries from the past. And also that just almost tautologically, there's disproportionate attention on hot new technologies, which means those fields are maybe saturated, and it's harder to make advances in them. I'm putting words in your mouth, but that's how I interpreted it.

And then it also sounded like you have a relatively unique approach of starting with a big question or big problem and working backwards from it. Would you say that's unusual? I don't feel like I hear people talk like that, that much.

Ed Boyden: Well, I don't know how unusual it is, but it does seem, doesn't it, that a lot of people have an area of expertise, and then they look around with their hammer and say-

Julia Galef: They inch outward from it.

Ed Boyden: Hey, where ... Is there a nail I can hit with it?

Julia Galef: Yeah, yeah.

Ed Boyden: Whereas we try to be the opposite. We can pick a big problem. We can try to survey all the different disciplines of science and

engineering. Chemistry, physics, math, computer science, electrical engineering.

And that's helped two things. One is that, as I alluded to earlier in the conversation, I had a very broad based education in chemistry and in physics, and electrical engineering and so forth. That's kind of nice.

But then also, I find that people love to collaborate, and there are lots of experts in different areas where you might meet somebody who's the best person in the world at quantum dot engineering, or the best in the world at a certain kind of computer science, or a certain kind of chemistry, and we can connect with them.

And so the third step is what I often call constructive failures. We try lots of things out, and although a bunch of it fails, we don't just chill off the failure. We try to extract wisdom from it. We've now seen something nobody's seen before, and even if it's not directly solving the problem, it might tell us what to do next.

And then finally is what I call designer discovery, where we go forth and actually make the real design of the technology, or we make the actual discovery of what we want. And those kinds of things happen a lot. It happened with optogenetics. It happened with expansion microscopy, where there was sometimes a multi year gap between having an idea and then going through the failure phase, and realizing the actual path we wanted to go down.

Julia Galef: Your idea of working backwards from a big problem you want to solve reminded me a little bit of a blog post by Aaron Swartz about, I guess 10, 11 years ago, which was called Theory of Change. And he was basically just saying a lot of people don't ... They have some societal problem that they care about or some goal, but then they're not working backwards from what would be needed to solve this goal, like, "What's my ... Based on my best current model of what would be needed to solve this goal..."

They're just looking around for a thing that seems plausibly useful or plausibly associated with the goal that's within reach, or within the set of options in their field of vision. And then they do that.

But they're not working backwards and being strategic and asking themselves, "In order for this policy to change, who would need to be onboard with changing it, and what would convince them, and so on and so forth?"

That's a favorite blog post of mine.

Ed Boyden: Well the other big problem I see - and I haven't read this post, but I'd love to... I would love to see if he addressed it - is whether people are solving the right problem.

Julia Galef: Yeah.

Ed Boyden: So one of the things that I often tell people is: Don't take requests from people without thinking for yourself whether that's the right level of problem to work on. Sort of like the old joke, if Henry Ford asked people what they wanted, he would've tried to breed a faster horse, right?

Julia Galef: Right.

Ed Boyden: Because people didn't have ... they didn't know the concept of the automobile to request it.

Julia Galef: Right.

Ed Boyden: And I have no idea if this is true or not, that he got that request.

But also, people look at a societal problem or biological mystery or medical issue, and if you try to confront it at face value, you might find that you're making a Band Aid. When in reality, there's some deep hemorrhaging problem that's whirling beneath the surface, and that's the real problem that we should have you solve.

So that's the other issue I see when people are struggling with working backwards, in addition to what you pointed out.

Julia Galef: I want to ask you about this ongoing friendly dispute I have with some of our mutual friends, about which approach to progress is more promising? I'm gonna call the two approaches the “rationalist” approach and the “Hayekian” approach. You could also maybe name it after Michael Polanyi, if you've read him. Those are just my shorthand labels for them.

So the rationalist approach to progress would basically be: Identify which problems would be most impactful to solve, most important for understanding or global well-being, and then strategize how best to solve them.

The Hayekian, or Michael Polanyian approach, would say that instead, important progress is more likely to result not from intentionally pursuing progress and optimizing for progress, but instead, from smart and creative people playing around with ideas that catch their fancy. Some of which ultimately spark discoveries, but in ways that we could never have predicted in advance.

Now, it certainly sounds, from talking to you, like you lean more towards the rationalist approach, but is that correct?

Ed Boyden: Well, I don't know if I would pigeonhole as such, because I feel like there is always a need to pivot your strategy towards whatever it takes to get the job done, right? And the four part outline that I made earlier is not meant to be a rigid formula.

Julia Galef: Oh, sure, but I think that's a part of it —

Ed Boyden: It's meant to be a tool in a toolbox.

Julia Galef: But you're still aiming at solving a problem, right? And the problem-

Ed Boyden: But the problem is meant to be a deep enough problem that it underlies a lot of other problems. It's a foundational problem. So as I mentioned, the two problems that I often thought about the most over the last 20 years were: How do we see everything, and how do we control everything.

So is that a problem first? It's not a problem the way that, let's say a classically trained physician might want to tackle tuberculosis or brain cancer, right? I said we're trying to dig one level deeper and think about, what's the underlying problem of biology.

And as I mentioned earlier, I trained in physics and chemistry. The way I think about things is, in physics and chemistry, you have a small number of things, like protons and electrons, and a small number of ways that they interact. Like electromagnetism and the laws thereof. And of course the laws of quantum mechanics.

Now the problem in biology is you have a lot of stuff, and a lot of ways they interact. We don't even know how many cell types there are in the human body, much less the molecules within, right? Maybe there's millions and millions and millions of variants that we haven't yes described.

So in some ways, when I look at all the struggles of biomedicine and how very little's been really cured in the last several decades in terms of major diseases... And look at brain diseases and cancers and aging related diseases, and the list goes on and on... What's the underlying problem, that if we solved it, might help clear up all the downstream problems?

So, I feel like there's an element of the latter, in the sense that you have to quest for the right problem. And maybe, once you of course

find the right problem, then you should go after it full force. And I think very often the problem is in finding the problem.

Julia Galef: But do you not think that we already have a backlog of important problems that need solving? Or do you just disagree with that?

Ed Boyden: Well, that might be all the more reason to think about, is there a way to dig one level deeper and invent a technology or make a discovery that solves many of them, right? So when we developed optogenetics or expansion microscopy, these are not tools designed to solve just one disease. They're tools designed to empower everybody to really do better science so that we make progress in all the disease. And so-

Julia Galef: Yeah, that's important.

Ed Boyden: Maybe we can help clear up the backlog a little bit by digging one level deeper.

Julia Galef: I've heard that the process of developing new tools is under-incentivized in science in general. Meaning that it's pretty valuable on the margin to have new and better tools - but nevertheless, for whatever reason, people don't get very much prestige or funding for doing so.

Is that your impression too? And if so, why do you think that is?

Ed Boyden: It's changing. I think your assessment is overall correct, in part because for a long time, tools were a little bit invisible. If you discovered crescent proteins or created a new sequencing reagent, maybe millions of people would use it, but ultimately what the public sees is a cure or a diagnostic, and the tools that yielded it sometimes go unheralded.

But a couple things have been changing. First of all, and this is more recent than most people think... Departments of bioengineering, or at MIT we have a Department of Biological Engineering, it's only a little more than a decade old. This is a fairly new idea that we should go forth and build tools that confront biological mysteries, and that allow the engineering of biological systems, right? So it's not an old idea, necessarily, at many places.

The second thing is that tools have become visible. And I think it's in part because some of the tools have spread so quickly. I think everybody's heard of CRISPR as well, that they have become visible in their own right in the way that previous toolsets were not, necessarily.

In my own life I've seen this change a lot. One reason why my home base is at the MIT media lab is because a lot of traditional departments at universities turned me down for faculty positions.

Julia Galef: Really!

Ed Boyden: And it was again, serendipity that the Media Lab had a job search that they were closing, I guess because they didn't find anybody that they wanted to hire. And I had been ... the quantum computing research that I mentioned earlier, I'd actually been working on at the Media Lab when I was an undergraduate and a beginning graduate student, many, many years before. And yeah, I was coming over to the Media Lab just to chat, and they're like "Oh, you're having trouble getting a job? Why don't you apply here?"

Julia Galef: They must be so smug right now, that they scooped you up.

Ed Boyden: Well, it's kind of fun. This year I'm actually co-directing the faculty search for the Media Lab, and it's kind of fun, you know - can we do on purpose what they did for me accidentally?

And so one of the visions I have for the media lab is that it becomes safe haven for people who don't quite fit in to any traditional discipline, but it's pretty clear that there's a non zero chance they're gonna change the world for the better.

And a lot of our recent hires have been at the interface of a life science and some other science. Kevin Esvelt working at the interface of sociology and politics, and community governance, and CRISPR and gene drives in genetic engineering. And then [?] as well, who's working on new kinds of neuro-prosthetics that sit at the border of material science, and the future of humanity potentially.

Julia Galef: So that's one criteria that you're using to try to identify people with a non negligible chance of changing the world: being at the intersection of the life science and some other science. Is there any other criteria that you could name that might help identify these people?

Ed Boyden: That's actually not a criteria.

Julia Galef: Oh, it's not?

Ed Boyden: I mean the Media's Lab's been hiring people in many fields-

Julia Galef: Is that just like a revealed preference?

Ed Boyden: Like art and design. I don't have a preference there. It just happens that Kevin and [?] fit that bill.

Julia Galef: Okay, well, are there any intentional criteria you're using? Like, if you have to describe your process for trying to identify non-negligible probability world changers, what would it be?

Ed Boyden: Well I'm a big fan of first principle thinking, as much as it can be achieved. I try to ... Rather than asking for somebody's opinion and then just taking it literally, for example - which some versions of peer review, I think, are huge on that - ...but one thing that I think about a lot is: How do we use logic and first principle thinking to really deconstruct the feasibility of what somebody's proposing? And also to try to forecast the impact of what somebody's proposing.

And some of this had really been helpful in my own group as well. This expansion microscopy technology that I mentioned earlier — the first time, 10 times we submitted government grants, on peer review. I think nine times out of ten, the grants were rejected. And so that was kind of depressing, 'cause how can we get the money to fund the project?

And then what came to the rescue was the Open Philanthropy project.

Julia Galef: Oh, yay.

Ed Boyden: They ended up giving a very generous and sizable gift to our group to bootstrap this project. When we started having the conversation with them, you can read the many-page long description of this year-long investigation into the project they launched. We basically said, you know what, if the goal is just to get peer review and take people's opinions, let's save you the time.

People don't get it. Maybe it's too new. Maybe it's too unbelievable for whatever reason.

Now the technology of course works, so hundreds and hundreds and hundreds of groups use it. But at the time, opinions would be negative.

But I said, Look, if you want to really look at the technology and from first principles think about how would it work, what kind of data would we get, how would it change the daily practice of science... Let's do calculations or even pilot studies when we need to, then let's talk. And they actually ended up giving us a three

million dollar gift, and they just gave us a second one, actually, to continue the project.

So I think there are ways to evaluate projects based upon, as much as possible, first principles and logical calculation, and physical science-based ways of looking at feasibility and impact. Nothing's perfect, of course, and especially in biomedicine, things fail all the time for the least predictable of reasons. But you gotta take some shots on goal.

And so one question is, can we move beyond opinion as much as we can, and assumptions as much as we can, in order to let new ideas through?

Julia Galef: There is an example I read, I don't remember if this was in a blog post of yours or an interview, but it seems like another example of this kind of logical first principles thinking that we haven't talked about yet... It was called Tiling Trees, I think. Can you talk a little about what that is?

Ed Boyden: Yeah. So this is analogous to the strategy that Karl and I took toward optogenetics, where we were just going through all the laws of physics, mechanical, magnetic, and optical and so forth.

The basic idea is, okay, you got a big problem. Great. That's a good start. Think backwards from that problem, and survey all the different disciplines of science and engineering, and try to think of every possible way to solve the problem. Now how can you do that? Well, the answer is, you can take the space of possible solutions and split it into two sets, and then keep splitting the sets into smaller and smaller sets, until you finally end up with individual ideas.

So for example, suppose you want to take the space of all possible energy systems. Okay, you could split it into renewable and non-renewable. Then you could take renewable and split it into two subsets, like solar and non-solar. And already things are getting interesting, right? Because how often do you think about a non-solar renewable system? So already we'll gonna have to stretch our imaginations. Maybe there's geothermal. Maybe there's the tides of the oceans caused by the moon.

And so eventually the goal is to split these categories into subsets so small that they are individual ideas that you could then test experimentally, or through calculation. But it's a very powerful way to think about it.

For brain interfacing, you could try to digitize the brain information inside the brain and then beam it out. Or you could try to beam out

the information in some other way like an Interlog form, and digitize it outside. And by doing these sort of binary chops, which of course results in this tree-like diagram, which is why we call it a tiling tree. The diagram looks like a tree, but at each level of the tree, the different nodes of the tree should tile the space of all possible ideas. Like, tiles on the bathroom floor.

It's a very useful exercise in idea generation. And we used it a lot in my classes as well as in my research group.

Julia Galef: I can imagine. And it's also very aesthetically satisfying on a very deep level to me.

Ed Boyden: Yeah. It seems like a number of other people have invented similar strategies. Fritz Zwicky, who developed a lot of astrophysics ideas almost a century ago, that now are being probed as hot topics in astrophysics - like you know, dark energy and stuff like that. He claimed to get many of his ideas through this kind of morphological analysis.

Julia Galef: Oh, that's interesting. Ed, I'm going to let you go in just a minute, but as the episode closes, I like to ask my guests to nominate a resource. It could be a book or blog or even a play, that influenced their thinking in some way. What would you nominate?

Ed Boyden: Well, one of my favorite books is a book called Time, Love, Memory by Jonathan Weiner. It's about the story of Seymour Benzer who also began in physics, and how he opened up the field of the genetics of behavior. He took fruit flies, drosophila, and mutated them and found mutations that caused changes in circadian rhythms. This was the topic that won the Nobel Prize last year, direct descendants of Enzo's work. Or, that would change mating preferences or change memory.

And just this idea, this odyssey of how you go into a field that's full of ambiguity, where there is no road map and no textbook, by first principles thinking and thinking backwards.

And it's just a great story too. It's full of intrigue and competition and colorful characters, and it's very, very well written. And so I read that book again, probably about once a year, just because it's so inspiring about the path that biology - and how one can try to open up a new arena, and how do you stick close to ground truth and avoid falling into pitfalls too much.

Julia Galef: Oh, awesome.

Ed Boyden: And Seymour is one of the people who opened up that entire discipline.

Julia Galef: Well, Ed, thank you so much for joining us on the show. We'll link to your research page as well as to your rationally Speaking pick. Yeah. And thanks for coming on.

Ed Boyden: Great talking to you, Julia.

Julia Galef: This concludes another episode of Rationally Speaking. Join us next time for more explorations on the borderlands between reason and nonsense.