Rationally Speaking #205: Michael Webb on “Are ideas getting harder to find?”

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, Michael Webb.

Michael is doing his PhD in Economics at Stanford University, and he is the co-author of an exciting recent paper called, "Are Ideas Getting Harder to Find?" And that is the question we're going to be talking about today. Michael, welcome to the show.

Michael Webb: Thanks for having me.

Julia Galef: So, your paper is actually kind of a funny exception to the general rule — I think it's called the “headline in the form of a question rule”, and the rule goes, if some article, usually in the popular media, has a headline in the form of a question, the answer to the question is almost invariably “No.”

Like, you know, "Are our children getting dumber?" Or like, "Could coffee be the cure for cancer?" And the answer's always no, actually no.

But the answer to your article — spoilers — is “Actually, it kind of looks like yes.”

Michael Webb: Yes. The answer to our question is a very strong yes.

This is work that I should say is co-authored with Chad Jones, and Nick Bloom here at Stanford, and John Murray at MIT. We worked on this paper for a few years, and I can tell you kind of why we've been working on it for so long, but the summary is that exponential growth in basically anything in the economy, be it looking at it at the aggregate level or any particular case study where we can measure these things well, growth is getting harder and harder to achieve. You have to sort of throw more and more scientists, more and more R&D budget at any given real world industry relevant scientific problem to get a particular given amount of growth over time.

Just to sort of summarize it with an equation — if I'm allowed to put an equation on a podcast, it's a very simple one — you can think of economic growth, 2% a year or whatever, as the product of two terms. The first is research productivity, how much you get per scientist. The other is, how many scientists do you have, the number of researchers.

What we show in this paper is that kind of wherever you look, wherever you can measure these things, research productivity is falling dramatically and the number of scientists is rising a lot. Where we do see constant growth, it's simply because we have an incredibly increasing number of scientists, who are offsetting the exponential decrease in those scientists' productivity.

Julia Galef: So it looks to people who aren't paying attention to, or aren't aware of, the rapidly rising investment in number of researchers or research hours, it looks like progress is going great — and they're just failing to account for the dramatically rising cost that we're paying per unit of progress.

Michael Webb: Right, exactly.

Julia Galef: What would an example be? What's an example of a metric that you've looked at?

Michael Webb: So, I think a kind of classic one we start with is Moore's law. So, Moore's law is this famous law that the number of transistors on a chip doubles every 18 months. This is a very well-known fact. It's amazing how straight the line is, if you plot it on a log graph over time.

What people perhaps don't know quite so well is the fact that it just takes so many more researchers today to get that same level of growth that it did originally. Compared to 1970s, it takes about 20 times the number of researchers to get that same rate of progress.

Julia Galef: Yeah. It's interesting, I've observed some reactions to your paper which was going around my corner of the Twitter-verse a while back… And I've also observed people reacting to similar ideas expressed without a rigorous paper behind them, but you know, "is progress slowing down? Is research productivity slowing down?"

And there are these two very different reactions to it. Where one reaction is like, "No, of course it's not slowing down. Look at all the progress we've made." And the other reaction is like, well duh, of course. There are diminishing marginal returns to anything. Any endeavor, you're going to pick the low-hanging fruit and then it's going to take more investment to make more progress.

Basically, half of the, I don't know about exactly half, one group is saying, "That's so obvious, it doesn't need saying." The other group is saying, "No, it's wrong." Which is a pattern that I see a lot with any argument that people make, that one reaction is, "Yes, duh," and the other reaction is like, "No, of course not."

Maybe the reason there are those two different reactions here is just people who think it's obviously wrong aren't paying attention to the costs, the increased costs that we're paying. They're only looking at the numerator and not the denominator of the fraction.

Michael Webb: Right, possibly. The bit of economics in which this paper sits is growth theory. It's trying to address a specific hypothesis that has been embedded in the literature on growth theory for a long time, which is precisely the assumption that scientific productivity is constant, right?

Even within economics, within people who study this full-time for a living, there's a substantial disagreement. And this paper is an attempt to say, "No look guys, it really is sort of happening in its way."

Julia Galef: Sorry, it's standard practice to assume that productivity is constant? Why would you assume that?

Michael Webb: Right, right. So, it turns out that models that where you make that assumption have a bunch of appealing properties.

Julia Galef: That's not a reason to use it!

Michael Webb: I've not been working in this literature since the 1990s, which the people who sort of argue about these things certainly have been. But my understanding is that at the aggregate level, certainly no one has been believing that what we say is not true — everyone can see these sort of basic facts that GDP has been pretty constant, if not declining for the last many years. At the same time, research investment has been increasing exponentially.

But where they disagree is more about the level where you look. They say, "Look, okay. Maybe this is true at the aggregate level, but if you look within, for example, any particular product line, then maybe it's true that the number of researchers" — and they've sort of followed their productivity — "is constant."

This paper is saying, okay, let's take that idea seriously and let's go and look at a bunch of product lines, and where we can measure them really carefully, and, okay, if that were to be true, what would we find? That has been very much an open life question in economics in recent years.

Julia Galef: Okay, let's give one or two other examples of metrics, just to give people a sense of the range of different places where you've looked at this phenomenon of falling researcher productivity.

Michael Webb: Right. So, another one that I like quite a lot is looking at seed yields. So, people often know about the green revolution, but even within advanced countries like the US, the yields you get — say, the bushels you get per acre of corn or wheat or whatever it is — have been increasing a lot, by about 2% a year in fact.

Julia Galef: Due to what?

Michael Webb: Due to scientific research on things like hybrid seeds, and also better pesticides, crop protection, this kind of thing.

However, at the same time we've had hugely increasing amounts of R&D going into precisely that thing.

And then the other one we look at is medicine. Within medicine, we look at both new molecular entities — these are the things that get registered with the

FDA, and every year some small number of them will turn into blockbuster drugs. You look into the flow of those against R&D then in pharmaceuticals.

And we also look at disease mortality. We say, let's try and quantify the amount of increased life expectancy we are achieving for a given amount of research effort on a particular disease. So, if we run a thousand clinical trials against heart disease or breast cancer, what are we getting for those clinical trials?

Julia Galef: In terms of actual improvements to life expectancy?

Michael Webb: Exactly, right.

Julia Galef: And the thing we're getting is increasing a little bit, but not a lot, and we're putting many times more researchers into the problem and getting less out of it than we used to?

Michael Webb: Precisely.

Julia Galef: What's your preferred explanation for why we see these diminishing returns? Is it just like, the structure of the universe — that as you try to discover things, you find the easiest things first and then it gets harder and harder to find new things, and that's just the structure of knowledge? Or are there other candidate explanations?

Michael Webb: I don't think I have a preferred explanation. This paper is very much just documenting the fact. It’s not trying to make a case for any particular reason why it might be true.

I think that's a super interesting research agenda that is now... It's not as if we are proposing this question for the first time. People have worked on aspects of this question. I think this wide open vista of interesting research that has yet to come, that will both explain something about the sort of knowledge perhaps, and also then what we might do about this, which we might come onto later, I guess.

In terms of the kind of explanations that you can think of:

So, one is this idea, the “distance to the pit face” that you have to walk — this is a mining metaphor — to get to where you're shoveling the newest bit of, getting towards the gold or whatever, is a lot further today than it ever was. The amount of knowledge you have to have as a scientist to be able to get to the frontier, to make these contributions, is just so much larger today.

And you can see this from the amount of time of it takes to do a PhD, how old an inventor is the time they first take out a patent, the size of research teams. Ben Jones, he's a fantastic economics professor at Kellogg, has papers that document these things.

That means that for individuals, they could either end up spending more time studying, which is what you see in the PhD length, or you see that they just focus on narrower and narrower fields. For a given amount of time, you only learn something about a much, much narrower field. Which might mean that you just have less good insights if it turns out that for all you progress, the fields ... The wider field you have to be combining with some knowledge from quite distributed science.

So, that's kind of one big explanation, is the “distance to the pit face” thing.

Some others are just that we are experiencing these general purpose technology-driven waves of technological change. So you might think that we've been, the last 40 or 50 years, we've been benefiting from IT, from the computer. Before that it was electricity.

And within a given wave, when you first discover it, it's like an oil field when you first strike oil — it spurts out and it's super easy, but over time you take out most of the oil, and it gets harder and harder, and it's like trying to find the last sardine in the sardine tin. It's really, really hard.

But we have to find the next big, big oil field, which would then be — maybe that's AI. Or maybe something coming out of synthetic biology, or whatever. So that's one thing, of diminishing marginal returns within a given technological…

Julia Galef: Paradigm, or something.

Michael Webb: Paradigm, exactly.

So other quick ones: One is innovation exhaustion. This might be described as a fishing-out story. There's a few of them, and we've taken the low-hanging fruit and there's not so much left. That's consistent with what we find, but it's certainly not implied by it.

A couple of final ones: one is that it's plausible now that we're spending a lot more R&D on averting bad outcomes. You think about something like, the BP Deepwater Horizon disaster, and that has ...

Julia Galef: The spill?

Michael Webb: Yeah, exactly, the oil spill. “British petroleum” is what Obama kept calling it at the time. That has been associated now, subsequently, with loads more spending on really expensive valves for the oil and gas industry, and things to prevent these kinds of things happening. And that doesn't really show up in GDP, right? You just… don't see a disaster.

We have fewer low probability events... if you could run history a bunch of different times and take the average of those, then you would see this was worth well spending and it was increasing things we care about. But any given run that we experience through history, we don't see the bad thing happening.

But we do see all this money that went to preventing it. So that doesn't show up in quite the same way.

One final, final thing would be: A recent paper has been looking at the difference between the kind of innovation that startups do — a big grand new idea — and then the kind of innovation that more established firms do, which is much more follow-on innovations to that one idea they had originally. So insofar as we have a lot more bigger firms today and fewer startups, which we do, you might just expect to see that there's less foundational exciting innovation going on, and more incremental stuff that is less valuable.

Julia Galef: Right, that's a great rundown.

Michael Webb: Sorry, that's a huge ...

Julia Galef: No, that's great. I appreciate the taxonomy. I love me a good taxonomy.

To the last point that you made, unless I'm misunderstanding your paper, it seems like you guys were mostly looking at… Let’s use the mining metaphor. You're mining some mineral in a given mine and you start having to go deeper and deeper into the earth to get the same amounts of mineral, so that's sort of diminishing marginal returns. But you could go elsewhere and start mining in a different location. Or you could start mining a different mineral. Then that could compensate for the diminishing marginal returns per mine.

And if I'm understanding your paper correctly, it looked like you were just tracking diminishing marginal returns per mine. Diminishing marginal returns and productivity of researchers working on stuff relevant to transistors on a chip, or relevant to crop yield.

And so, couldn't it be that there are just other mines that are being started that you aren't capturing in your analysis? I'm sure you know more about computing science than I do, but if you looked at our ability to process natural language, or something — I know there've been a lot of increasing researchers into computing science recently, but there's also been huge improvements in our ability to process natural language in the last 10, 20 years. And so I would expect that “productivity per researchers” curve to look better than for crop yields. I'm just guessing.

How do you feel about the argument that you're not capturing lines of new research, you're only measuring refinements to existing ideas?

Michael Webb: So I'm sort of sympathetic to that argument.

But on the one hand, the pattern that we find is writ large in GDP overall. So if you look at the overall economy, we've had constant, if not declining, growth — yet exponential increases in researchers. And so in theory, GDP should be capturing all this amazing new stuff that's happening.

Julia Galef: I see.

Michael Webb: You gotta have a story as to why that's also failing. And there may well be. I'm not an expert on how things end up in GDP. And maybe GDP is subject to the same critique that what we're doing is, in some sense. But there is that factor of GDP that you have to account for.

Julia Galef: Right. Just to make sure I'm understanding the two different ways to approach this question:

One is looking at productivity in a specific, narrowly defined field or problem or technology. And that has the upside that that's very concrete, and objective — we can just track crop yields. There's not a lot of wiggle room in interpreting that.

But it has the downside that there's a bunch of other potential technologies that you're not capturing because you're just focusing on specific selected ones.

And then the other way you could approach this question is looking at productivity in the whole economy. And that has the upside that you're not missing a bunch of things that people are doing. But it has maybe the downside that it's less clear that productivity overall is capturing what we care about?

Michael Webb: Right.

Julia Galef: So it seems to me that might be the case because measuring GDP is tricky, or it's tricky to capture whether the quality of things that go into GDP is being captured or ...

Michael Webb: Exactly.

Julia Galef: I should let you answer because you're the economist. Why is it that GDP might not reflect increases in productivity in the sense that we really care about?

Michael Webb: Just to go back to what you were saying, there's absolutely this huge tradeoff — as everywhere, really, in empirical economics — between coverage and quality of measurement. You can measure a small number of things really, really well, which is what we do in this paper. Or you can measure the thing you actually care about, the really big thing, quite badly.

And so we do both. We also talk about GDP in this paper. So GDP, I should know a lot more about how GDP is constructed. I don't claim to be an expert. My understanding is, it varies a lot by the kind or product. If it's steel, that's really easy — it's like, count the number of steel bars, or do something that tries to approximate that. But with software or IT it gets much, much harder. Because at the end of the day we don't care about number of transistors on the chip, we care about what the thing can do and what that thing is worth to us.

Julia Galef: Right.

Michael Webb: So the people who calculate this for the US government and the OCD, and all those who publish these guides as to how all economists are supposed to do it — it's an absolute giant mess, is my understanding. A bunch of specific hacks for any particular given product, and ways of trying to get at these things.

In particular, new products. I'm remembering things from what I read, again, I'm not an expert but what I remember is that when new products come into GDP often it's with a big delay. So I'm remembering that Ford's cars did not show up in GDP for a long time after he first brought them to market. If I'm remembering correctly, that would mean that this critique that you're mentioning is still right on with us back to GDP.

Another thing that I would say on that critique is that: I'm pretty sympathetic to the idea that the very fact that you can measure something means that it got to that point where there's enough people in the field, they've kind of had the big ideas already, and you're already now sort of going down that curve of diminishing returns. And maybe at the very beginning, there's like three people in the field and they have all the very big important ideas.

To look at computing, for example, we didn't start measuring Moore's Law — so we're looking at that from the 60s or so, but most of the big ideas for computing happened in, there's a bunch of papers in, the 1930s, right?

Or look at deep learning today. You mentioned these amazing things in natural language processing and other areas — most of the really foundational stuff for that was done ... The first patent for voice recognition was in, I think, 1962. Back propagation was around from around that time.

Julia Galef: Back propagation being the basic AI algorithm that allows us to do voice recognition?

Michael Webb: Exactly. The means of learning for these algorithms. Now it actually works, now we have enough computing power. The idea was sound when it was first had. But it wasn't useful for that many things back then. And so that was when it was this tiny field.

If we now went and started measuring deep learning as one of our case studies, just the number of people in the field now is absolutely enormous. And there's a huge amount of duplication going on, and a bunch of people writing not-very-good papers, and it's like impossible for anyone to figure out what anyone else is doing because there's so much of it. It so hard to separate the wheat from the chaff. And then also really targeting these metrics. They care about things like percentage accuracy on some benchmark.

And if you look at these graphs, again I've not looked into it with numbers in detail, but I would put a big bet that they would totally reflect the same story as in our paper. Which is that you have — take machine translation, you have very slow process before deep learning comes along, then you turn that switch and suddenly it jumps up this huge amount, and then a few big jumps. But now there's thousands of people all trying to squeeze out a .5% to 1% change to this metric, right?

Julia Galef: Ah, interesting.

Michael Webb: So if you did the calculation that we do for everything else in this paper, you would see, I bet, again, probably an even bigger decline in the research productivity then we have in our case studies.

Julia Galef: Ha, okay. So that was poorly chosen example on my part.

Michael Webb: No, that was a good example of illustrating what so hard about this whole exercise.

Julia Galef: Yeah. If it's true that ideas take a long time to show up in productivity then that would just imply that things looked bleak 20 years ago or something. It's sort of like looking at the stars in the sky — the light comes from a long time ago. So we're seeing them as they looked many, I don't know, millions of years ago. I don't actually know what the right scale is here.

Michael Webb: Very nice metaphor.

Julia Galef: I feel like that doesn't actually undermine your analysis, it just means that we should be wondering about what was happening 20 years ago, or 30 years ago — why did things slow down?

Michael Webb: Right.

Julia Galef: I'm also wondering: it seems like the way people talk about the paper, I've read interviews with one or more of your co-authors, it seems like they're assuming or implying that if we see productivity slowing down as the number of researchers goes up, it can only be that ideas are getting harder to find.

But it seems to me that there could be a bunch of other things — other inputs to productivity besides number of researchers. Like, what about amount of regulation, or competitiveness in an industry, or something? What if there's been a steady increase in the rate of people developing widget-producing technology, but the widget industry is kind of unhealthy, so it's harder for new companies to get started? Or the industry has been super regulated and so their productivity looks bad, even though the researchers are not at fault?

Michael Webb: So, all those things you just said could be true. And I think that most of them do have strong elements of truth, there's a bunch of evidence that new business formation has been declining for decades now. And there's a bunch of worrying trends in that direction.

But I think the trends we document are just so big. There's just such a big increase in researchers…

Julia Galef: What’s an example, take Moore's Law for example, what's the rough increase in number of researchers working on the problem?

Michael Webb: Right, so one way putting it is that every 10 years we are doubling the number of researchers working on Moore's Law.

Julia Galef: Huh. And what's Moore's Law again? Remind me, it's every ...

Michael Webb: That's the transistors on a chip.

Julia Galef: Yeah, yeah, but what's the ...

Michael Webb: Every 18 months you double the number of transistors on the chip.

Julia Galef: Every 18 months you double the number of transistors and every 10 years you double the number of researchers?

Michael Webb: Yes.

Julia Galef: That… sounds good, actually. Wait, am I doing the math wrong in my head?

Michael Webb: Well, we're treating as the output the *percentage* change at the time.

Julia Galef: Ah, yeah.

Michael Webb: So there's a constant 35% per year exponential growth.

Julia Galef: Right.

Michael Webb: And that's not changing. But the number of researchers that are being used to create that constant 25% a year is going up exponentially.

Julia Galef: Got it, right, right. That actually was another thing I wanted to talk to you about. As you know, we chatted about this a few weeks ago.

Michael Webb: Yeah.

Julia Galef: One of the responses to your paper could be: Why look at the percent change? Why not look at the absolute change? Isn't it impressive that researchers are able to produce more and more transistors per chip in just an absolute sense? Why grade them in a different way? Why grade them by the *percent* improvement in number of transistors on a chip?

Michael Webb: We just think that's the natural unit, if you like, for that thing.

We do use different units for different case studies. So for medicine, we're looking at absolute changes in life expectancy. But for Moore's Law we totally think it makes sense, it would be very silly not to use proportions. If you just think about the nature of the ideas, right? If you have an idea for a clever way of, a new shape for transistors, so you can fit more of them on any given chip, if you hand me a chip with a million transistors then my new idea would mean that you've now got 1.2 million transistors. If you hand me a chip with a billion transistors suddenly that same idea would give me 1.2 billion transistors on that same chip. So the idea is giving a proportional improvement not an absolute improvement. Does that make sense?

Julia Galef: Yeah, it does… I honestly, I confess, I don't know how to think about whether we should care about absolute versus proportional. I'm so used to things being measured, progress and growth being measured in proportional terms but in other cases there are reasons for that. I don't know, I find it hard to think about, but that's reasonably compelling to me.

The kind of research that allows us to increase the number or transistors on a chip is more applied than the kind of research that physicists are doing, that might help us figure out what the structure of the universe is. Have you looked into / do you have any intuitions about whether progress is also slowing down in more pure research? Or just applied?

Michael Webb: I have very little intuition about the rate of progress in pure/propositional knowledge. We've been looking at these very concrete things, very applied, which have actual economic implications and are therefore easy to measure.

I would have no idea how to go about measuring the amount of knowledge we have from theoretical physics, for example. If you talk to these people, some of them think there's been a big slow down in that field, for example. I'm happy to sort of take what they believe as some evidence that I should take account of in my own beliefs.

But I think it's just not at all clear to me what the metric would be, how you'd even begin to think about… I can think of some ways, one can think of many different ways of quantifying increased knowledge, propositional knowledge in the natural science-

Julia Galef: When you say propositional knowledge, that's in contrast to what?

Michael Webb: In contrast to prescriptive knowledge.

Julia Galef: So is that sort of like pure versus applied? Or is it not quite ...

Michael Webb: Kind of. I'm taking this directly from Joel Mokyr, he's an economic historian at Northwestern, who's wonderful. He certainly didn't invent those terms; propositional knowledge is knowledge "what," and prescriptive is the "how."

"What" is kind of, how does the world work? Like, what is nature? What is the boiling point of this thing? Where is Neptune? And prescriptive knowledge is, "how" do you build a chip? Like, as in a new computer. How do you make clothes?

So prescriptive is kind of what can you patent. And propositional is what counts as a discovery in some sense. You can't patent discoveries, but you can patent new ways of building something.

Julia Galef: Got it. That's a good way to draw the line.

Michael Webb: And indeed, so the actual definition — we don't mention it in the paper but it's sort of implicit in everything — it comes from the economist Paul Romer, who founded the ideas-based growth models in a wonderful 1990 paper in the journal of Public Economics, where he defined ideas as "the instructions that we follow for combining raw materials."

There's this great quote which is, "100 years ago all we could do to get visual stimulation from iron oxide was to use it as a pigment. Now we put it on plastic tape and use it to make video cassette recordings."

… You can tell it was written in 1990! But it's a nice way of putting it. How do we combine what we already have, to get better stuff, more stuff that we care about?

Julia Galef: I like it.

Michael Webb: And that's what was, in this paper, that's the ideas that we're referring to in "Are ideas getting harder to find?"

When it comes to things physicists study, often that's incredibly important for them being able to make new ideas, as we mean them in new recipe/instruction manual, the stuff we care about.

And possibly one of the reasons that we've seen declining productivity that I didn't mention in that long, what you kindly called a taxonomy, I just call a long laundry list ...

Julia Galef: It was a simple taxonomy.

Michael Webb: A one dimensional one!

… is just less spending on precisely that kind of propositional knowledge, that pure science, foundational science. All these big breakthroughs in applied work, in ideas as we mean them, come from discoveries that were more fundamental science. And if we're spending less on the first, then we might perhaps be getting less per scientist of prescriptive knowledge.

Julia Galef: Maybe we should only care about propositional knowledge in so far that it eventually translates into some measurable improvement in productivity or human welfare or something like that. And the reason that we care about it is that we assume eventually some fraction of the propositional knowledge we accumulate will be able to be turned into prescriptive knowledge.

But we still might as well just measure the outcomes we care about, and assume that eventually the propositional knowledge, if it's worth anything, will show up in that somehow. It just means we'd have to measure these trends over longer time horizons, if we don't want to look at propositional knowledge separately.

Michael Webb: I guess that's a very sort of Baconian view. The value of this part of it is simply an enterprise that will create useful knowledge we can use for the economy. And I'm certainly sympathetic to that. But I wouldn't want to go on record and say, "Physics has no value except for if it leads to better iPads, or whatever."

Julia Galef: Just in terms of what we should ask the government to pay for.

Michael Webb: In terms of what we want taxpayer's money to do — should the taxpayer be funding us to study just stuff that we find interesting — even then, maybe.

You might think that going to the moon has value just as a thing of scientific discovery and not as ... if you've only done that, it would still be a great thing to have done. And we might think that one of the roles of an advanced economy is precisely to improve human understanding and knowledge, even if it has nothing to do with how good our iPads are.

I still think there's a huge, almost moral value and certainly aesthetic value, I think may scientists would say, to that more pure/foundational knowledge. But I'm also sympathetic to the idea that we mostly care about it, at least as economists we care about it, as an enterprise that translates propositional into prescriptive knowledge.

Julia Galef: Yeah. I might have missed this in your list of potential explanations but it seems intuitive to me that there's just all this institutional inertia in universities and academic fields and society in general.

And maybe the reason that we're seeing a decline in researcher productivity is that, unproductive subfields, there's just no mechanism to kill them off. They're just gonna keep submitting grant applications and they're gonna keep making arguments for why they're worth funding, and they're gonna keep finding things to study, that may not be that important, but it's a thing to study that justifies their existence.

So if we had a way to kill off unproductive fields then maybe we would see better productivity?

Michael Webb: It's kind of a question about allocation of resources in society. Sure, you could kill off this unproductive field, if you thought you could find one. But my question would be, how do you know that that field is in fact unproductive and is not on the cusp of something that is really important and that you care out?

Julia Galef: Yeah. Well, fields or departments, yeah.

Michael Webb: Oh indeed, departments, right.

And then also, the people in those departments, what would they go and do instead? Because if you spend 10 years getting a PhD in this super narrow thing — and you are contributing to that thing — and your knowledge is so specific that if you weren't doing that, then you'd have to be unskilled, manual labor or something… That's probably putting it too far. But sort of like, less good than what you're doing as a professor at some department. Then it's not clear that we should do that.

Julia Galef: Right, so maybe we don't care about researcher productivity, literally, to the the exclusion of all else.

Michael Webb: Right. If we were to — you're sort of talking about interventions here — if we were to go and somehow, declare we are now stopping all funding of A, B, C, D and E — then what would happen? So that's a different question from just how productive are all these fields right now?

Julia Galef: Right, right.

What other questions are you most interested in, in this general area? What do you think are the important next questions that we should be trying to answer?

Michael Webb: I think the really important questions are: What is driving both the overall trend that we see, and also the differences between fields? And then within a field, between labs in a field and so on?

And I think we're now in this amazing time in human history where we really have incredible access to the kind of data that you'd need to start to answer these kind of questions. And also I think a much increased willingness among everyone to submit themselves to RCTs for example, and for funders to fund these kind of things.

Julia Galef: Now that the increase in people in various fields paying attention to rigor, and replicability and so on, has taken root.

Michael Webb: Exactly. I think you see an increase in willingness of scientists to impose that on themselves.

Julia Galef: Right.

Michael Webb: So economists, for example, have recently been doing experiments on, how do we incentivize peer reviewers to have a fast turnaround? So we randomize you to either, we promise to pay you money if you do it in this time, or we just thank you for doing such a good service for the profession, and those sorts of things. And we just run the RCT and see what happens.

I know less about other fields, but I know economists are very serious about taking what they do seriously and applying it to themselves. The American Economic Association has committees on market design for the economics job market. They really take their own theories very seriously, which I think is great.

Julia Galef: It's like the academic equivalent of eating your own dog food.

Michael Webb: Yeah, exactly, it really is. I think that combined with the ease of measuring now — so, now that so much more is being done on fairly consistent interfaces, on computers, it's easier to measure things. And we've got, you mentioned these natural language processing advances, for example… I just feel like, many, many things that we care about are just much, much easier to measure now.

Maybe I'm being naïve, maybe they look easy to measure —

Julia Galef: Until you get up close. Yeah.

Michael Webb: Until you get up close. And this paper has been an experience of looking outside and then getting up close and realizing there's so much more going on than you realize. And it takes you an extra year just to measure who was doing some R&D in 1970, or whatever. Which we had to do for this paper.

But in general though, I'm optimistic about our abilities to measure and then test theories about what's important. And then also again to run RCTs. I know people working in Harvard right now who are people in HBS doing RCTs ...

Julia Galef: The business school?

Michael Webb: ... the business school, doing RCTs on the medical school, to encourage them to collaborate more, for example.

Julia Galef: Yeah, maybe that's the way to do it. Have each school or department do RCTs on different school departments — so they're not tempted to rig it in their favor!

Michael Webb: Yeah, I think that's a really exciting and promising area. Another thing — from the laundry list of things I mentioned, some of them are amenable to RCTs. But some of them are kind of really not. I think there's a scope for a very wide collection of disciplines. And quite possibly, and I strongly suspect, there are many people who are not economists who are working on this. Some of them I do know about, but I'm sure there are many I don't know about, who are doing really good work on this in the general area.

Julia Galef: Which of the explanations in your laundry list — I guess we're calling it that, not taxonomy anymore — Which of those explanations would you be most excited to start with? Either because it seems like the most likely to be true, to you, or because it's the easiest to measure? It's like, under the lamplight.

Michael Webb: I would differentiate theories that are easy to test, and interventions that are easy to test.

Julia Galef: Ah, that's a good point.

Michael Webb: So I think testing the theory that "the low hanging fruit gets plucked first," some of those theories are almost impossible — they have, like, no testable implications. It's not clear to me how you would test ...

Julia Galef: Yeah, that's a good point.

Michael Webb: It's like, by definition the thing that you first find is, because you found it first, therefore it must be low hanging fruit.

Julia Galef: Right, right. Is there any other measure of low-hanging other than the time it took us to find it? Which is exactly the outcome.

Michael Webb: Right. Exactly. I've not spent months thinking about this, and maybe if I thought a little more about this, we'd find out that it wasn't tautological… but it strikes me from first glance that this is a hard thing to test.

Julia Galef: Right.

Michael Webb: But in terms of interventions, it's super easy to go and say: if we find ways of getting scientists to collaborate more, some clever way of forcing them to talk to each other, that that's been done already. Or if we go and actually measure the practices of the most productive labs and see if they're doing things differently from the least productive labs, and can we spread the best practices from the good to the less good labs, that kind of thing.

No one, as far as I know, is doing that... there's no big effort to do that in any fields that I know of. And it seems like that would be an incredibly useful, productive thing to experiment with.

Julia Galef: Before I let you go, Michael, I want to ask you to nominate the pick of the episode, which is a book or article, or it could be a blog or other website — that you don't agree with, at least you substantially disagree with, but you still think is valuable and worth engaging with. Maybe that's because you think it's well argued even though you disagree with the conclusions. Or maybe you think it has a really interesting hypothesis, even though it's unlikely to be true.

Does anything like that come to mind?

Michael Webb: Yeah. I think I would go for Foucault.

Julia Galef: Ah, a bold choice, what's the case?

Michael Webb: An interview called "Truth and Power" he did in the late 1970s. Which is not one of his famous works, but it's a very nice summary of a lot of his thinking. I read it in undergrad in my political philosophy class, at the end. There was one optional class where you got to choose what you studied, and I chose Foucault.

He argues that we live in regimes of truth. So at any particular point in time, in history, the criterion for what counts as true changes. And the criteria for how you distinguish between true things and not true things. And who has the right to say those kinds of things.

For example, today RCTs are held up as the gold standard of evidence. There are these literal hierarchies that medical institutions put out, and the number one gold standard is RCT. And that's a pretty new thing, actually. 100 years ago no one had even heard of that phrase. Or even less than 100 years ago. People weren't doing them.

Before that, you have Foucault looking at mental health, for example… You just see that the people who had the power to decide to promulgate what was true, and the means by which they were allowed to assert those claims and be taken seriously by society, and the government, have just changed dramatically over time.

I don't know what he'd say about RCTs. But I can certainly have a really interesting debate in economics certainly, and I'm sure in other fields, about the validity of RCTs and what they tell us, and are they actually useful? And all the problems we've seen recently, and how effects disappear over time. I think it's quite possible we sit here thinking RCTs are amazing — well, some people do — and we'll have very different opinions on RCTs in 50 years time, 100 years time. And in general, who scientists are and the kind of work they do may be very different.

Now, I'm not sure that ... Foucault would go a lot further, I think. I'm no expert at all on Foucault. But that paper was the most mind blowing thing I read in my life so far, at the time I read it. And he makes a bunch of very strong claims, and often in ways that it's not exactly clear what the assumptions are, and what the argument steps are, and it's not the kind of thing that you would like.

Julia Galef: Yeah, I remember feeling like he was kind of slippery. But I gotta say, Foucault is actually on my list of things that enough smart people I respect have said are valuable, that I'm grudgingly willing to keep trying to get the value out of them.

And also — this is a brief tangent, it's your pick technically but just to indulge myself — I've been trying to think about how to find sources that might update my models of the world and change my mind.

And one kind of obvious, easy way to do that, is to find people that have really good epistemic standards, they are really rigorous thinkers, they have nuance where I think they should have nuance, and they change their mind, and they seem very truth seeking — but they have different object-level views from me. And so I want to know, like, "Why are you so confident that we should not allow immigrants in?" Or something. "And yet you seem to have such good standards of argumentation and thinking, this is really intriguing…" That's a high value person for me to talk to.

But the problem with that strategy is, what if my standards about what counts as “good epistemics” are wrong? How am I going to find people who can show me that my criteria are wrong?

One of the working answers I've come up with so far is, find people, thinkers, writers, etc ...

Michael Webb: Like Foucault.

Julia Galef: Yeah, well, what is the rule that generates Foucault?

And I think the rule might be: Find thinkers, etc, whose epistemic standards seem not that great to me, but who are respected by people who I respect. And using that kind of transitive property might allow me to revise my set of criteria that I'm using to pick people.

Michael Webb: Or maybe they're respected by people who respect you who respect ... more than one link of respecting, right?

Julia Galef: Yeah, but the fidelity might break down after too many links.

Michael Webb: Yeah, it's interesting. I think what's interesting about Foucault is — and others who I would regard as in the class of “consciousness-raising” writers, who tend to not be on the syllabuses of modern university classes and analytical philosophy, for example — they are like different glasses you put on. They're so foundational about the way they see the world.

And as long as you can look at them with your current glasses on and say, “Show me assumptions, here's some pieces I have and how do I build on those to get to where this person is”… It's almost like, no, you've got to take off your current glasses, and forget everything you thought you knew, and put on these other things. And just take that leap of faith. And essentially you've taken a leap of faith already with your current worldview, about what standards of evidence count.

And there's a bunch of reasons for doing that. So if you think about modern science, everyone has given us all these amazing things around it. It really seems to work, in terms of generating useful stuff, at least in generating technology. But maybe that's not the only criteria, not the only thing that matters. And I think it's just important to have a bunch of different lenses on the world, for understanding any particular thing.

I'll put out a second recommendation, the coolest book I read last year that's on this theme of different lenses. It's a book called “Images of Organizations,” by Gareth Morgan. He's an industrial sociologist. It's written in the 80s, maybe the 90s.

It's a list of 10 or so different metaphors for understanding what an organization is. So one of them is the obvious “organization as machine,” or “organization as brain,” sort of an information processing unit.

But there's also “organization as organism,” in ecology, or “organism as psychic prison.” There's a bunch of these different things. I think that book makes a really strong case that any given one of them just does not paint the whole picture. And the skill has to be in figuring out when to look through each particular lens.

I think that’s true for everything in the world. And I think the scientific lens is like the best one we have for most things, but it's not the only one. There's a bunch of things where you might actually learn a lot by trying to understand other perspectives, even if that is uncomfortable and doesn't make sense from a given worldview.

Julia Galef: Well I'll resist the temptation to start an entirely new episode on definitions of truth and lenses for looking at the world.

Michael Webb: Ha, yeah. Next week.

Julia Galef: But Michael, thank you so much for coming on the show. We'll post links on the podcast website to your picks, and also to your paper and your website with your other research on it.

Michael Webb: Thanks for having me.

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