

Rationally Speaking #204: Simine Vazire on “Reforming psychology, and self-awareness”

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, Simine Vazire.

Simine is a professor of psychology at the University of California Davis, she's the author of the blog, Sometimes I'm Wrong, and the co-host of The Black Goat, a podcast about doing science.

Simine's research is really interesting. It's about how accurately we understand ourselves — our personalities and our behavior — and why that matters. So we're going to talk about that topic. But the way that I first encountered Simine was in a different role that she plays. She's been a central participant in the conversation about methodology in the social sciences, and where the field needs to shape up, and how.

Just to give you an example and a taste of Simine's style, she teaches a seminar that's titled, “Oh, You Like That Finding, Do You? Well, It's Probably False.” So we'll be talking about that as well. Simine, welcome to Rationally Speaking.

Simine Vazire: Thanks. That's actually not really the title of my class. That was my joke title. But yeah, that's basically the theme of the class.

Julia Galef: I'm not fully calibrated yet on —

Simine Vazire: When I'm joking? Yeah, right. I'm not sure it's obvious to others, either.

Julia Galef: But that does sound like an accurate description of the class.

Simine Vazire: Yeah, yeah.

Julia Galef: I've talked a fair bit on Rationally Speaking already about the replication crisis, reasons why studies don't replicate, with some previous guests like Brian Noseck and Uri Simonsohn.

But one argument that I've been thinking about recently, *against* the idea of increasing rigor in the social sciences, I wanted to pose to you. The argument is:

Look, false positives are bad. Thinking that we've found something cool in our field that isn't actually there, and it's just an artifact of a badly done study plus confirmation bias, et cetera, that's bad. We don't want to go down a bunch of blind alleys.

But false negatives are even worse. We don't want to fail to discover real phenomena. And so maybe there's a trade-off where if we increase the standards of rigor — like we make it harder to publish things, we increase the

standard of evidence — basically, we're reducing false positives, but at the same time we're maybe increasing false negatives, and that's a bad trade-off.

This is an argument some people have made to me. I'm not sure exactly how I feel about it, but what do you think? Do you think that there's a real trade-off there?

Simine Vazire:

I think in principle there could be if we were doing everything right, but I think in reality false negatives aren't an issue. I actually have a blog post called *Why I'm Not That Worried About False Negatives*. And I lay out some reasons why, given the way we're doing things at least in psychology, but I don't think it's specific to psychology, there's actually very little risk of people abandoning an actually true hypothesis because of a false negative.

One reason for that is that I think with p-hacking and researcher degrees of freedom, the kinds of practices that we've, until recently, allowed and even encouraged, you could turn almost anything into a significant result. So if you have a hypothesis and you test it even just once or twice, and you allow yourself to look at it from a lot of different angles, you're probably going to find something that you can interpret as evidence for your hypothesis.

So if you're willing to do those things you don't understand that they're harmful, and you're motivated to find something to support your hypothesis — so basically if you're human — you probably will find evidence for your hypothesis even if it's not true. So there's very little risk that if it is true, you're going to miss the evidence, unless we start getting a lot stricter.

So let's say we do. Let's say we crack down on these p-hacking and researcher degrees of freedom, and so on. I do think then we have to start worrying a little bit more about false negatives — which basically means we need to increase our sample size, which is something that those of us pushing for reforms, they go hand in hand. That if we're going to stop p-hacking but keep our sample sizes the way they have been, then everything is going to be uninformative. So in order to not end up failing to pursue promising things, we need to increase our sample sizes.

But there's other, more sociology of science, I guess, reasons why I'm not that worried about false negatives. One is that even now, even with the reforms, very few journals will actually publish null results. So if you fail to get something, it might discourage you, but it won't discourage others because others won't find out that you failed to get it. So it doesn't have the same ripple effects that false positives have.

It also seems like people, even if we did publish some negatives, like now we're publishing replication studies once in a while, still not very much, but some are getting out there, and some of those are probably false negatives. Many of them are no results and some of them are probably false negatives, and even those are not getting very much attention. So often the original study continues

to get way more attention than even a much more rigorous pre-registered, large sample replication study.

So it seems that the null results, first, tend to not get published, and second, when they do get published people don't pay attention to them. You know, the things that are going to make good headlines and be good click bait are usually the significant results. The things that are going to make it into textbooks are usually the significant results. So I don't think there's as much potential for a false negative to change a lot of people's minds and convince a lot of people that there's nothing there and it's not worth pursuing, and so on.

This might be different in other fields — like maybe when it comes to cancer treatments or things like that, maybe given the competitiveness and so on, people will use other people's negative results to avoid going down a dead end or something like that. I don't know, I don't know what the culture is in those fields. But in psychology I don't see very much evidence that everybody concludes, oh well, there must be no effect because this one lab didn't get it, so let's abandon that altogether.

Julia Galef:

Oh, interesting. I'm now realizing that there are two different things I was conflating here when I talk about this trade-off. I think. Tell me what you think.

One mechanism by which reducing false positives could increase false negatives is the thing that you're describing. Which is: people publish results showing, actually, this effect isn't there or we failed to find this thing. And that spreads and that sort of takes root, and maybe there was a real thing there.

But the mechanism I was originally thinking of is: there is a real phenomenon and the study shows that the phenomenon was real. And the study has flaws, but — so, it's like it's an informative but flawed study. And whatever reasons the researchers had to think that the phenomenon was real were Bayesian evidence of the phenomenon, their paper itself just wasn't actually strong enough, it didn't meet the standards of evidence of the journal.

Simine Vazire:

Yeah, so then we go back to our priors, right? So I think if a paper is so flawed that we're no better off after seeing the results than we were before — which, I think many of us feel that way about some chunk of the published literature from the past, that basically, we just don't know whether to put any stock in it. Again, not the whole literature, but there is some pieces of it.

I think we definitely need to be careful that absence of evidence is not the same thing as evidence of absence. So if I think that a study was flawed to the point where it doesn't add to my certainty in the effect, I shouldn't conclude there's no effect. I should go back to my baseline level of certainty in the effect, which was whatever my prior beliefs were before reading that study.

I do think sometimes we make that mistake of thinking, “If they had to p-hack to get this result, then it must not be real.” But that's not true, right? An effect could be real, but the study was so poorly designed that they couldn't detect it

without p-hacking. Or they just p-hack because that's what they were taught to do and so it doesn't mean that they had to p-hack to get the effect, and so on.

So yeah, I think that's important. I'm not sure how big of a problem it is in reality. I do think, I've seen that happen, people slipping into "if it was p-hacked, then the effect must not be real." Some of that, I think, is legitimate because I think our priors on some effects in psych should probably be low, because we've been pushed more and more to study counter-intuitive things.

I think the plausibility of some of our hypotheses is low to begin with, so then if the study was poorly designed then we go back to our prior, which is like, "That didn't sound likely before the evidence, so I still don't think it's likely."

Julia Galef: Right. There's actually a bunch of things like this in the social sciences, broadly, where I believe there's a real effect there, and there's a bunch of research showing there's a real effect, but the research has nothing to do with my belief in the effect.

Simine Vazire: Yeah. So just to get super controversial right away ... One area where I think that might be the case — and I don't know the literature super well, but I hear people express skepticism about the studies, and let's assume that skepticism is at least sometimes valid — is stereotype threat.

Julia Galef: Can you just explain for listeners —

Simine Vazire: Yeah, so the most abstract version of stereotype threat, and I'm going to butcher this, but it's basically the idea that if there's a negative stereotype about a group that you're a member of — your gender, your ethnicity, something like that — being reminded of that stereotype is going to make you anxious. And it's going to make you perform worse in that domain.

And I think on some level, that must be true. My prior on that is quite high. And then if there are some studies that are flawed, or that used practices that were acceptable at the time but now I wouldn't consider strong, that doesn't necessarily change my prior. But it doesn't increase my confidence either.

Julia Galef: Simine, another question I've been wondering about recently with regard to increasing rigor in the social sciences, or in science in general, is: How much is it true that new research builds on past research?

Well, let's just stick to your field of psychology. For example in theory I could just go run studies today on Mechanical Turk, testing various hypotheses in psychology, without being all that knowledgeable about the literature. And if my methodology isn't very good, then the results aren't going to be great, but at least I'm not relying on previous work ... Trusting the previous work was good, and basing my investigation on that.

It seems to me that the degree to which we really need increased rigor depends in part on how much we think psychology is this pyramid when new research is

building on previous generations of research and distrusting that it's solid. This is just a fact that I don't know about the structure of psychology.

Simine Vazire: I would say that I ... I don't know if this is what you're implying, but I agree that psychology is not very much like a pyramid. I can get into it a little bit more.

But I don't think I agree that that's related to whether or not we need regular ... In psychology, a common joke that people say ... I don't remember who said it first but, there's a joke — which is not really a joke, it's basically true — that in psychology, theories are like toothbrushes. No self respecting person would use somebody else's.

Julia Galef: That's funny, gross and depressing.

Simine Vazire: Yes. It's sad. When I was an assistant professor trying to get tenure, everybody told me ... I think it was true, although I never actually got the word explicitly — but the rumor was that to get a grant from social psychology at NSF, which is the main place that we get grants, you had to have your own theory.

So I came up with my own quote-unquote “theory,” which was really, really simple and obvious, but you couldn't just build on someone else's theory. I was at a meeting once where a dean from Stanford said that they would not give someone tenure for incrementally improving somebody else's theory.

I think that's a common view, is that incremental work ... This is different in different sub disciplines of psych, I think in cognitive psych there's more valuing of incremental work than I think in social and personality psych. I don't know for sure, but that's my impression.

Julia Galef: Just a brief tangent, I'm curious: Do you think these professors, who wouldn't give tenure to someone who didn't have their own theory, do you think they agree that incrementalism is valuable? That double checking other people's work, and modifying other people's work to make it better, is valuable?

Simine Vazire: I think they think it's less valuable. I think they think somebody needs to do it, but the less skilled people should do it, not the Stanford professors. That's my guess. That's the only way I can reconcile it.

I see a lot of people talking about the importance of creativity and novelty, but when you say, "But isn't correction also important?" They say, "Yes. Yes. Of course." Then what I take away from that is they think, “Well, the really smart people are doing the creative novel stuff, and then the people that can't do that, they can do the correction.” The second stringers, as some people would put it.

Julia Galef: Really gives unpleasant incentives, or unfortunate incentives, if we've just decided that the correction stuff is valuable but low status.

Simine Vazire:

Right. Right. And that's actually a rosy view. I think many people think it's not valuable. Or they in principle think it's valuable, but when it actually gets done they think, Why are these people being so mean? They think the specific instances of correction, when you see what that has to actually look and feel like, they don't like it. I'm not even sure that it's valued even as a second stringer activity.

But I would say that even if that's the case and this is ... I have a pretty bleak view of this, and maybe I'm too pessimistic. But even if it's the case, psychology is not as incremental and not as much of a pyramid as it maybe should be, or may be that for young science.

But in any case, even if that's true I would say it's still really important that the past literature be [reliable] because we still use cumulative [work]... to decide what gets put in textbooks, who should get awards, et cetera. Things still accumulate...

But I think meta-analysis is one example of where we've got really big trouble. And the other phase I've seen myself and other people go through when you start questioning whether we're doing right in science, one of the early phases, is you start putting all your faith in meta-analysis because you realize that single studies aren't very trustworthy.

And that actually turns out to be worse. There's this later disillusionment where you realize that actually the cure is worse than the disease. Meta-analysis really assumes that the set of things being put in the meta-analysis has more signal than noise, and if not, it's just biased. When you aggregate biased things together, the bias amplifies.

Julia Galef:

I'll just try to clarify why I thought there was a relationship there — you can still disagree with me, but just to make sure I was clear... In a world in which there's not this pyramid structure someone who just decides, "Darn it, I know that a lot of my peers are using shoddy methods and that's unfortunate — but I'm going to be a really good researcher. I'm going to use good methods." They can just do that, and do good science and get trustworthy results — in this non pyramidal world.

But in a world where they have to trust other people's work, then they're screwed even if they personally want to use good methods.

Simine Vazire:

Yes. I think that's true. I think that it's easier to wipe the slate clean and start over if things weren't building on each other to begin with, or weren't building actually too much.

On the other hand, if we had had a more pyramid structure it's possible that we wouldn't have let the problem slide so much. If, many, many other people were building on your work, it's possible that they would have detected the problems at some point. Somebody would have come along and said, "No. This way of

doing it isn't good, and if I'm constrained to do it the same way as everyone else, then I'm going to critique the way that I'm being pushed to do it."

I think that it is easier going forward to just do things a different way. It doesn't disrupt some long chain of things. On the other hand, it means that the older stuff is never going to get cleaned up it just sits there.

Julia Galef: Related to this point: You had this great argument a while ago about how the standards that people apply to replications — when a team of scientists attempts to replicate someone else's study, to double check that the phenomenon they purported to demonstrate was real — that the standards people apply to replications are really different from the standards that they apply to regular, original papers. What's going on there?

Simine Vazire: Yes. I think that actually it's a broader phenomenon. I didn't really make this connection until very recently, that it's maybe not so much about replications that it's just about findings we don't like. And maybe we do the same thing if it's an original study.

I've now been an author, a coauthor, a collaborator on a number of papers where we had an idea — we planned a study, or we found data sets that we could use to test it, and we didn't find what we expected, we got a null result. And it's often really interesting to watch me and my collaborators go back and question whether the data really was a good way to test the hypothesis, whether the design was really adequate.

And I've also seen this as an editor and reviewer. I had one case where a reviewer explicitly said, "I thought the design was fine but then I saw that the result was null, I went back and looked at the design more closely to see what was wrong with it, and I found these problems."

In a way, that's great. But we should do that independent of the results. And it's really interesting to me to see how willing we are to throw our methods under the bus when we don't like the result.

And I think in case of replications, often ... In that case, the people who don't like the result are not necessarily the authors themselves, but maybe the authors of the original study or people who are fans of the original study. They can nitpick the methods. My view is let's do that, but let's do that to *all* studies, including the ones that find exciting things that we want to believe.

But it's really fascinating to watch the level of critical thinking, how people step up their critical thinking when they don't want to believe the result. It's kind of nice to see that we're capable of it, clearly.

Julia Galef: There's a quote by — I think it was Tom Gilovich, who wrote "How We Know What Isn't So." He said ... I'm going to mangle the quote, but it was basically, "When there's something that we want to believe, we ask ourselves 'Can I

believe this?' and when there's something we don't want to believe, we ask ourselves, 'Must I believe this?'"

Simine Vazire:

Right.

Julia Galef:

Implying the different standards of evidence we're using.

Simine Vazire:

I catch myself doing that, and I see reviewers doing that. When we're evaluating science we should always be asking, "Must I believe this." Or, maybe not quite that far, but, would it be foolish to believe there's something like that? I think, "How much do I have to believe this."

Julia Galef:

I'm pretty familiar with the phenomenon of, "Must I believe this" and it makes sense to me that people would be applying that to failed replications of studies that they liked, or that they were invested in.

But the phenomenon that I thought you were pointing at was different, and kind of interesting in its own right — it was more of a status quo bias. That even if you have no investment in the original study or the replication — the failed replication — there's this thing that happens, where we accept something as true, and then the replication is on a different plane of justification.

Simine Vazire:

It reminds me of Andrew Gelman, where he says imagine that —

Julia Galef:

Yes, that was what I was thinking of.

Simine Vazire:

The reversal test, [reverse the order with] the original. But actually that implies that it's about order, but I don't know if it's about order.

I think it's about prestige in many ways. I think the original often has a lot more prestige. Part of that is because it came first. And people think that replications are less intellectually innovative, because they're just copying something — we talked about that earlier. For that reason it's also often that people doing replications are less famous than the people who did the original work very often. Often replications are published in less prestigious journals than the original work et cetera.

I think that there's a bias, yes. It ends up being a status quo bias because the original came first, but I think a lot of it has to do with just people being skeptical of the ability and the motives of replicators. And being skeptical of people they've never heard of; being skeptical of journals where the replications often get published.

Julia Galef:

Is it just your intuition that that's the more likely mechanism? Or do you have any way to distinguish that from the status quo bias model?

Simine Vazire:

I guess I would predict, very strongly, that if a null result got published in a lower tier journal, as an original study; and then someone came along and replicated it and said, "No. See, there is an effect" ... And they were more famous, and they

got published in a better journal — I don't think there would be status quo bias. I don't think people would be like, "No. No, We're going to stick with the original conclusion."

I think it's a bias toward significant, rather than all findings, of bias towards more famous, more splashy people in journals and conclusions. But I don't think that order is a big part of it.

Julia Galef: That's a nice segue into another question I wanted to ask you, which is: You've made the case in several venues that the scientific community is too reliant on "eminence" as a marker of quality. Older, prestigious scientists from more prestigious institutions, et cetera. They're more likely to get attention and awards and publications, et cetera.

How can we tell that eminence isn't just a marker of quality? How do we falsify the null hypotheses that there isn't any bias towards them, and it's just that talented people become eminent, and then they get more attention and awards because they're talented, et cetera?

Simine Vazire: It's really hard. Because part of the criticism of eminence is a criticism of metrics, and I think that a lot of us ... Well, a lot of us feel like the way people get recognition is partly by just having many, many papers that are cited many, many times.

And if we are arguing, "Well, we shouldn't just be using those metrics. We shouldn't decide who's successful just based on these numbers that can be gamed, or could just not reflect quality" ... then how do we disprove that hypothesis you just said? Because you have to do it quantitatively to convince people. You have to find other metrics. But every metric can be gamed, or can reflect something other than what it's supposed to reflect.

I think that sometimes this is what's behind wanting to test the validity, through really rigorous large scale replications, of some of the most classic, most celebrated findings. Some people think that this is to take down famous people and to try to become famous by taking down a famous person yourself. But I think part of it is, to test this hypothesis that there's some calibration between the recognition and fame and attention a finding, or its authors, get — and how solid it is. I think that's what's behind some of that drive to be, "Well, let's see. Let's take the most solid, most celebrated things that we teach our undergrads. If that's not solid, then the correlation between prestige and rigor can't be that high. If the most prestigious things turn out not to be rigorous."

I don't know how we would test the whole spectrum, but I think a pretty good way to start is to test the things at the top.

Julia Galef: That makes a lot of sense to me. Although it also makes it hard ... A lot of the discourse around how to respond to criticism or failed replication of your work is, "Look, this isn't judgment on you as a scientist or person. This is just the

process of science, we should be correcting each other's findings. It's not personal."

But as you say, it also is personal, just in the sense that it's a referendum on whether you deserve your prestige and status.

Simine Vazire:

I remember in the earlier days of the replicability crisis in psychology, which has been going on now for six or seven years... Maybe two years ago, not that long ago, and I'm sure it could still happen today. But a few years ago, I heard someone say that people shouldn't do replications because they're skeptical of a finding. That that's not a good reason to do a replication. I think that was a very widely held view at the time, and maybe still is today.

I think that's crazy. Of course it's a legitimate reason to do a replication. Yes, that introduces bias, but original authors are biased in the sense that they hope that there is an effect. Someone who thinks that there's not an effect, that doesn't make them more biased than someone who thinks there is. We just need mechanisms in place to rein in that bias, like preregistration and transparency.

In that sense it can be personal, although I still think in that case it's still more about the effect. When I'm talking about testing the most prestigious findings, I'm talking about the findings still, rather than the people. But obviously those are very hard to separate.

Another kind of paradox of the replicability debate is that the critics are asked to not target people, to not name names, "Why do you even have to reference the authors names, just talk about the effect?" But then when we're getting praise, when we're citing something as supporting in a positive way — we do use the names. And it would be really weird not to.

There's this double standard, that it's counted as personal if you say, "So-and-so's effect" in a negative way. But you do want it to be called after their name when it's in a positive way.

But yes, I think it is a little bit ... We shouldn't deny that there's some aspect of choosing things that are held up in really high regard, to see if those stand up. I think that's a good way to test how deep is the problem, but yes, it's not a neutral way.

Julia Galef:

One of the many things I like about your blog, "Sometimes I'm wrong," is that you do address the human side of improving rigor and correcting findings. Which is something that I find lacking in a lot of discussions of the replication crisis that I otherwise agree with.

People will say, "Look, you've got to take criticism, you've got to let people to critique your work, you should be happy when people criticize you, because the field is progressing." I agree with the spirit of that, and I think the paramount

virtue in science has to be criticism and transparency. Even if it hurts people's feelings.

But I still think people tend to be pretty glib about ... I hear people say, "Oh, I'm always happy when I get criticism." And I'm like, "No. You're not. I just don't believe that."

Simine Vazire: Right.

Julia Galef: Do you have any suggestions for listeners, on handling — I was going to say handling "fair criticism," but maybe I want to broaden it to any kind of criticism — with aplomb?

Simine Vazire: I don't know. It's funny, because I thought you were going to ask how to *deliver* criticism in a more sensitive way.

Julia Galef: That's also a good question.

Simine Vazire: I'll start with that one. My answer to that is actually to try to remember a time when you were wrong and had to admit it.

For me, I try to pay attention to those times, because I think you can learn a lot from them, and have a lot more compassion for when you're the one criticizing someone else. I was recently at our annual SPSP conference. Which is a big conference of social and personality psych. I was in a room with my friend Alexa, and I was flying in from California, and the conference was in Atlanta. That's a three hour time difference — and she wanted to get up early to work out. I was really grumpy about the fact that she wanted to set an alarm for 6:45. I gave her a really hard time and I was just a jerk about it. Then the next day I was like, "Oh. Man. I was such a jerk. It's totally reasonable for someone to want to wake up at 6:45 at a conference. That's just normal."

I texted her, and then found her and gave her a hug. But it was hard, even though I was sure that I was in the wrong, I had to swallow something. Almost physically felt I just had to swallow something. But it was nice even though this is not at all an intellectual thing. It was a nice reminder that even something silly and easy like that... She forgave me right away, it wasn't hard, but even that was probably the most unpleasant thing that happened to me over a span of a couple of days. If you're asking someone to accept that they were wrong and something where there are serious things at stake, I think having some compassion is good.

But I do think: if I had to say what's overdone, the lack of compassion and meanness in criticism, verses the attacking people for not criticizing nicely enough? I would say the latter is overdone more. There's just way too much complaints about the way people are criticizing, when I think the much bigger problem is that there is so much to criticize. That should worry us.

But, yes. I think that if you're going to call someone out or point out an error ... I'm not sure that there's any way to do it without it hurting other person, so that's not the right standard. But just like remembering that how much that sucks, even the little things are important.

And then in terms of how to respond: I think taking a beat is such an obvious answer. But as an editor I've had more experiences now with people criticizing my decisions, or criticizing the things we've published, or things like that. And even if I think I'm totally calm, that I could respond right away and it would be totally fine — it's almost always a good idea to wait a day.

...

Julia Galef: There's a entertaining footnote on your blog — which, by the way listeners, Simine's blog is peppered with entertaining footnotes — this one said "emotion suppression gets a bad rap." And I think I agree with this, but I wanted to ask you about it.

Simine Vazire: I'm not an emotion researcher, but my understanding of the emotion research in psychology is that there's different emotion regulation strategies: there's reappraisal, where you try to reframe what happened in a more positive light. And reappraisal is often considered the most adaptive way to regulate your emotions. And then suppression, which is "try not to think about it," is considered a less adaptive way. And I think there are other emotion regulations strategies. But those are the two that come to mind.

And I always thought suppression got such a bad rap. For me, if something's bugging me and I can't do anything about it, I literally try to find something in my visual field and just mentally describe it, and I'll feel better. I think trying to distract yourself often is a really good strategy.

I think it depends on if it's something you need to respond to or do something about, then suppression might not be good. And, I mean, the other argument against suppression is that it makes it harder for others to know you, and I think that probably is true of me — that I am a fan of suppression is probably not unrelated to the fact that people find me hard to get to know.

But it has a good intrapsychic effect, that maybe the interpersonal effects are not as good, I don't know.

Julia Galef: I was actually always suspicious of the "you need to process your emotions, you can't just suppress them" — just because it sounded too... "pretty." It felt like it fit into this general narrative that our culture happens to have, at this point in history, about emotion and authenticity, and so on. That didn't necessarily by itself mean that the claim is wrong. But it made me more suspicious of it.

Simine Vazire: I feel like there's this idea that if you suppress emotions they're going to bounce back, but that's not always true. Surely sometimes they just go away. I mean, I think it depends ...

Julia Galef: People have this metaphor of a spring — you push down, it springs back. But it doesn't have to be that way.

I definitely want to make sure we leave time to talk about your object level research on self-awareness. So this is a good time to segue into that.

A recent paper of yours had a pretty clever novel way of measuring how well people know themselves. Can you tell us about that?

Simine Vazire: Sure, I think I know what you're talking about.

Julia Galef: Maybe you've had multiple clever ways of measuring this! I'm thinking of the one with the recorders.

Simine Vazire: Yeah. This is something actually when I was in grad school in the early 2000s, there was another lab in the same department, Jamie Pennebaker's lab. And his grad student Matthias Mehl had developed this technique called the Electronically Activated Recorder. Which at the time was literally a tape recorder. And they used it to look at language, and what people talk about. Especially during traumatic events and so on.

And it occurred to me that this might be a really cool tool to look at actual behaviors.

So one problem in personality research, which is my sub-discipline, is that we tend to rely a lot on questionnaires and what people say they're like in a questionnaire could be right or wrong. And we've kinda taken it ... it's not so much that we've taken it at face value. We have tested the validity of people's self-reports, and they are quite valid.

But if you want to study where people might be wrong, where they might have blind spots, there wasn't really a great way to do that. Because you could use self-report questionnaires, or you could use peer reports — which is where you ask the people's friends and family what they think — but then if those two disagree, which they do a little bit, it's kind of a glass half empty, glass half full. So they agree substantially, but there's areas of disagreement. And there wasn't really a good way to solve who's right when they disagree.

Julia Galef: You could always say, like: "My friends say I'm arrogant" or something, "...but what do they know. They're just jealous."

Simine Vazire: Exactly. And sometimes, by definition — if your friends say you're not funny, then it doesn't matter what you say.

Julia Galef: That was actually the first example I thought of. And then I was like: wait, no. That actually doesn't work.

Simine Vazire: And so there's a few where by definition, you're right. So your self-esteem, you're almost by definition right. Though you could be really deeply deluded

about your self-esteem. And then how charming you are — others, by definition, are right.

But there's a lot of stuff in the gray area. Like if you say you're friendly and others say you're not, it can be ambiguous. I always have ... in my head, there's one that I count as more valid than the other, but it would still be nice to have empirical evidence.

And the problem with most behavioral measures that we had up until that point was that they had to be administered in the lab. So you had to bring people into the lab and videotape their behavior. Put them in a situation and videotape their behavior. And there's a lot of downsides to that.

For one thing you might put people in a situation that they actually never would have chosen to be in. And so then you're eliciting behavior that is not typical for them.

So if you want to know what they're typically like, which is what we mean by personality, you need to catch them in their own natural environments. And this technique that Matthias Mehl and Jamie Pennebaker had developed, the electronically activated recorder, or EAR, would allow us to do that. At least for some behaviors. Because it's an audio recorder, we only get behaviors that you can tell from audio recordings.

So we have our participants wear that for various amounts of time. In the last study we did it for six days. And it comes on and off, so it records thirty second snippets at regular intervals. So then we had the self-report questionnaires so we knew what they said they were like — we knew what their friends said they were like — and then we had this week of their actual behavior with their friends, with their classmates, etc.

And so we have a huge team of coders. Undergraduate research assistants working in our lab. And we develop coding protocols, and they listen to the recordings and code behavior from those recordings. So then if you say you're really warm and your friends say you're not, or vice versa, we can listen to the recordings and try to adjudicate and say who's right or wrong.

That's a kinda overly-simplistic way of thinking about it though. Because here, the way I described it, we're treating the coder's ratings as the truth. But they could be wrong too because they're only getting audio. They're not seeing your face, they're not seeing your movements, etc. they're only getting thirty second snippets so they're missing some context. They're only getting 5% of the time. It's not recording all the time.

Julia Galef: But that just adds noise, right? That wouldn't ...

Simine Vazire: It adds noise, yeah. But for some behaviors, we basically can't get them from the EAR. Some behaviors are either too rare, or they're just not acoustically

detectable. So one of the things we're trying to do is figure out which behaviors you can reliably get from the EAR and which ones you can't.

Julia Galef: Interesting. So what did you find about people's level of self-knowledge?

Simine Vazire: Again, it's kind of a glass half empty, glass half full type of thing. In one study we found that the self-reports and peer reports were about equally accurate, but they were accurate about different behaviors.

And then we came up with an explanation post hoc for what kinds of behaviors the self is probably more accurate for, and that's behaviors that are more private, basically. Which is not too surprising. And then what kind of behaviors others are more accurate for, and that's behaviors that are more public or overt. And also things that are more evaluative. So things that would be hard to admit about yourself.

Julia Galef: A type of value or judgment?

Simine Vazire: Really desirable or undesirable things. And not actually because ... often people jump from that to "everybody loves themselves and thinks they're great." And that's not actually the case. Overall, your friends' rating of you tend to be more positive than your self ratings. And some of that is probably a little bit artificial.

But the problem with the self ratings is not that everyone loves themselves. The problem with the self ratings is that there's individual differences in how much people love themselves. And that plays in to how they rate their personalities. So the higher people's self-esteem the more they describe themselves in a positive way.

Some of that is valid. There's probably a correlation between having a good personality and having high self-esteem. But some of it is just positive self views, so like a halo effect.

And people with negative self-esteem actually under rate themselves. And so that messes with the accuracy.

Julia Galef: Across the board, on everything? Or are there some traits that everyone, to varying degrees, tends to be positively biased about themselves and some traits that everyone tends to be negatively biased about?

Simine Vazire: There's individual variation on every trait. So on every trait, if you ask people to self report, there's going to be a range. And that range, I think in every case that we've looked at, it correlates with self-esteem. People are more likely to overestimate in the positive direction the more they have high self-esteem. And they're more likely to under rate themselves the more they have low self-esteem.

There's definitely mean differences. So there's some traits where people on average tend to rate themselves more positively. And that's the more desirable

traits. And also the ones that are more vague, and easier to define in whatever way you want. So if I ask about your intelligence, people are going to tend to give a higher rating, and it's going to be more influenced by their self-esteem than if I ask "How good are you at math?" Or "How good is your vocabulary?" Those are going to be more accurate and less influenced by self-esteem.

Julia Galef: And what have you found about the effects of someone's level of self-knowledge, how much their view of themselves correlates with other people's, or with objective ratings?

Simine Vazire: That's the million dollar question. What we really want to know is, who has more self-knowledge and what's different about those people. Do they have better relationships? Are they happier? Are they better off in some way? Do they make better career choices?

Julia Galef: I'm personally very interested in this, because I have this strong suspicion that accuracy makes you better able to navigate the world.

Simine Vazire: I want to believe that, like you do.

Julia Galef: But I'm interested in challenges to that.

Simine Vazire: I think we just don't know. At this point, we've measured accuracy for the whole sample. So we can come up with one number that tells you on average the correlation of both self reports and their actual behavior is, say, .4. So there's some accuracy, but it's not perfect and so on.

But we haven't yet come up with a way to say "here's this person's level of accuracy. And here's that person's level of accuracy." So we have accuracy for people as a whole. Or for our sample as a whole. But we haven't come up with a way to look at individual differences in self-knowledge or in accuracy. Which that is what you would need in order to correlate that variable. Each person's level of self-knowledge, with outcomes. Relationships, and work and so on. There are ways to do it, I just don't think any of them are very good.

Julia Galef: Didn't you ... am I misremembering? I thought you had a paper where you looked at people's self-knowledge and their relationship quality. Were you dissing your own paper there?

Simine Vazire: Yes.

Julia Galef: That's sweet. Why, though?

Simine Vazire: There's three authors on that paper, and we disagree about how confident we are in the findings. And I don't really want to ...

Julia Galef: That's fine.

Simine Vazire: I think we would all say it was a single study with 80 people and a high value. So, if it's true, we got kind of lucky.

Julia Galef: Worth investigating further.

Simine Vazire: We would all want to see it replicated. I intended to replicate it, and then stupidly, in my last EAR study, which — an EAR study takes many, many years. So the one that we're currently analyzing, we collected ... data collections started 2012. We'll finishing coding the EAR files maybe in two or three years.

Julia Galef: As a former research assistant, I can only imagine how much work that must be.

Simine Vazire: Yeah, our research assistants are really ... they work so hard and they're great. And it still takes forever. So I had planned to include all the variables we would need to replicate that effect that you referred to, but I stupidly didn't. Like, I thought I had, and then I went to go start planning the analyses for when we have the EAR data and I was like "we didn't measure the right variables." I have no idea why.

And that's one of the problems with these studies that take forever is that you go back and you're like "why didn't we measure that?"

Julia Galef: So, I don't know if my next couple of questions can be answered given the lack of individual measures of self-awareness. But I'll ask them anyway.

What's your impression of how deceived people really are? Like to the extent that we have inaccurate views of our traits — do we really, deep down, believe those inaccurate views?

The argument that we don't would go something like: I might insist I'm great at fighting because that makes me look good to other people. And maybe in that moment that I'm insisting it, I kind of believe it. Because that makes it easier for me to be convincing to other people that I'm a good fighter.

But then if a fight actually breaks out, maybe I'm just going to run away, because deep down, I actually know I'm not good at fighting.

And so my brain's doing these two different things at different times depending on what's appropriate. There's the signaling system, and then there's the "actually make decisions that won't get you killed" system. Call it the navigating-the-world system. And it uses different models of me depending on what system is appropriate.

Simine Vazire: That's my intuition too. And that's one evolutionary theory about self deception, is that the function of self deception is to deceive others because you'll be better at convincing other people if you really believe what you're saying. So at least on some level it helps if you believe it.

We do have one study that gets at that. So we ask people to rate themselves on a bunch of characteristics. We recorded those ratings. They couldn't change them. Then we showed them the ratings again, and said "do you think that you overestimated or underestimated or just right?"

And we found a lot of accuracy. So people who overestimated said they overestimated. People who underestimated said they underestimated. It turns out they were just using the heuristic that if I rated myself high, then I probably overestimated. If I rated myself low, I probably underestimated. So it was not a sophisticated judgment. But it was accurate.

So I think that supports this idea that on some level you believe it, and on some level you don't. And I've definitely had that experience, and it's something I've talked a little bit with philosophers about — "Isn't it possible to kind of believe something but also kind of not?" They don't really like that.

Julia Galef: Philosophers are weirdly stubborn about that. Like, to me it seems obvious — just, having a brain and living inside that brain, that this is a real phenomenon. There are books and books written about the paradox about why people do things that aren't in their self-interest. And I'm like Well, I have different "selves" in my head.

Simine Vazire: And was it W.H. Auden or Walt Whitman, I can't remember which one, but some poet wrote "do I contradict myself? Very well then, I contradict myself. I am vast. I contain multitudes." And I've had that experience all the time. I feel like I contradict myself all the time. Including about myself. And I think that that's adaptive to some extent. Like, it's nice to be able to bring up different self views depending on what's going to be functional in what context I'm in.

And maybe one way to think about that is, in terms of confidence intervals, for any given characteristic I have a confidence interval in my head. Maybe for some characteristics it's pretty narrow, I'm pretty sure I'm around the 80th percentile. And for other characteristics, I think my best guess is the 80th percentile, but it could be anywhere from 20th to 99th or something like that.

Julia Galef: There's also, just to complicate the picture even more: I might acknowledge and genuinely believe that I'm probably wrong about 20% of things that I believe. I could genuinely believe that. But if you point to any specific thing that I believe, I feel much more than 80% confident in that thing. Across all the things.

Simine Vazire: This goes back to the replicability issue. Lets say that we think that at least 30-40% of our publish studies are false, which I think that shouldn't be too controversial. Let's say 30%. I don't think that should be controversial. But we don't know which 30%. And so some of us take that to mean, let's be skeptical of everything. And other people take that to mean, I'm not going to worry unless I have a reason to doubt a specific paper. And I think when it comes to our own beliefs, we're more like the latter. Like, "I'm not going to worry about which ones are wrong. Unless I get feedback that tells me I'm wrong, I'm going to assume I'm in that 80%."

Julia Galef: Can I run my own pet theory about self-knowledge past you?

Simine Vazire: Yeah.

Julia Galef: My pet theory is — this is specifically about is self-knowledge helpful. Or accurate knowledge about yourself. And the theory is that there is an interaction effect. So it's helpful, if you have some other trait or important cluster of traits. And an example of a trait I could imagine filling that role is low neuroticism. So accurate knowledge of yourself is helpful if you are able to not freak out over the flaws that you see in yourself, and not dwell on them, and react constructively to them.

Simine Vazire: So if you're good at emotion suppression basically.

Julia Galef: Exactly. But if you have high neuroticism, then maybe you're better off not fixing your flaws and going on.

Simine Vazire: I think there are a number of what we would call moderators of Self Knowledge, and one of them has to be whether you can do anything about it.

But “do anything about it” is a broad category, because I have plenty of characteristics that I know I can't do anything about — but I actually wore the EAR and I listened to myself. And I learned how flat I am. And I didn't know that about myself. And people had told me, but listening ...

Julia Galef: What do you mean?

Simine Vazire: I don't express very much. I'm not emotionally expressive. I sound uninterested in things. I think I've gotten a little bit better, but not much. So listening to myself, I realize I come across as if I don't like anyone and I don't like anything they're saying and I'm not interested in anything. I also say much less than I thought I did. So I would listen to these conversations and I would remember the thoughts I had and be like "I'm sure I expressed that thought." And I would wait and nope, I never expressed it.

I can't really change that. Or maybe this is also me being a personality psychologist about it. I'm pretty pessimistic about people's ability to change. But now that I know I come across that way, I will sometimes go out of my way to tell people, "you probably think I don't like you, but actually I do."

Julia Galef: Well, I have a very good friend who was like that for years. He never really smiled. I always was confused about why he kept wanting to hang out with me because he always looked so miserable when he did hang out with me. And he has very low ... well, I don't know if you'd call this neuroticism. But he's very good at acting on ...

Simine Vazire: Feedback.

Julia Galef: ... knowledge about his flaws. So when he found out he was giving people that impression, he just learned how to smile. A lot. And got into that habit. And now he gives off this very warm vibe to people he hangs out with. So there's hope. I felt like you were very engaged in this conversation, for whatever that's worth.

Simine Vazire: I think I've gotten better. And having a podcast myself helps a little bit too, because I have to force myself to sound more expressive. The thing is, it's not that I don't feel it, it's that I assumed it was coming across when it wasn't.

I still find it hard to do in person, but I've learned to do it more in writing. Like, I'll follow up with an e-mail and be like "I really enjoyed our conversation" or whatever. So I'm a little bit better at that. So I think that it's hard to think of a case where there's nothing you can do with the self-knowledge. But I think there are some cases where it does more harm than good. And I think neuroticism is probably a good predictor or what might make that difference.

But also just other resources, right? There's more you can do to fix your flaws the more resources you have. Whether that's low neuroticism or money or time or people to support you or things like that.

Julia Galef: Well, Simine, before I let you go, I wanted to ask you for your pick for this episode. For you, I think my question is: is there a book, or article, or blog post, or something else that you've consumed over the course of your career that you don't agree with, but you nevertheless think is a valuable thing to read? Or well argued. Or worth engaging with.

Simine Vazire: What comes to mind is actually something I did when I was thinking about self-knowledge a lot in the early days of my research. I taught a class on self-knowledge and one of the things we talked a lot about was differences between how people see themselves and how others see them. And actually, I didn't do this in my class, but I had the idea of eventually developing a seminar where all we did was read autobiographies and then biographies of the same people.

Julia Galef: How interesting. I like that.

Simine Vazire: I never got around to doing it with students, but I started it on my own and one of the first ones that I did, and I don't remember how far I got beyond this.

The one I remember doing was Clarence Thomas. And so I read his autobiography and then I read a biography of him. And it was really fascinating. I would recommend that in particular. I think that's an interesting person to see through their own eyes, and then someone else's eyes. And I think I read a pretty critical biography.

But in general, I guess I would recommend, when people you really don't understand where they're coming from write autobiographies, that seems like a really unique opportunity. And obviously, especially politicians or people involved in the political arena. There's probably a lot of filter. But I still think it

can be really fascinating to hear what they want the world to see about themselves versus what a biographer would write.

Julia Galef: That's great. Do you remember the autobiography and biography of Clarence Thomas so that we can link to them?

Simine Vazire: I don't remember but I can ... I'll try to find them. I think the biography, there's only one. I'll try to find that and the biography.

Julia Galef: Well, Simine, thank you so much for coming on the show, it's been a pleasure.

Simine Vazire: Thanks, it was a lot of fun.

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