

Episode 256: How to be a data detective (Tim Harford)

Julia: Welcome to *Rationally Speaking*, the podcast where we explore the borderlands between reason and nonsense. I'm your host, Julia Galef. And today's episode features Tim Harford, an economist and journalist who writes the Undercover Economist column for the Financial Times and the author of multiple books, including most recently *The Data Detective: Ten Easy Rules to Make Sense of Statistics*.

The idea of the book is to help you evaluate statistical claims that you read or hear with a skeptical eye, but not a reflexively skeptical eye. So we talk about that and also about how to overcome motivated reasoning, which is something I wrote a whole book on and Tim devotes a chunk of his book too, as well. So it was interesting to compare notes and see where our approaches agree and where they don't. Here is my conversation with Tim Harford.

...

Julia: One thing I loved about your book was just the framing of it as a corrective to this classic book that I'm sure a lot of listeners have heard of, called *How To Lie With Statistics* - which came out in, I guess, the fifties. And it was just delightful to me because the book's very famous and I've never seen anyone criticize it before. It actually hadn't occurred to me to be critical of the book before.

But you made this great case for why we need a corrective to *How To Lie With Statistics*. So, could you just summarize what that book is about and why you felt a corrective was needed?

Tim: Absolutely. *How To Lie With Statistics* is a great book. It's a slim, a very witty guide to all of the ways in which statistics can be used to mislead us. It's got cartoons, it's got vivid examples, lots of great stories; it's by a journalist called Darryl Huff. And I read it when I was a teenager and I learned a lot from it.

And so, I don't want to be too critical, but there were two problems. The first problem is simply that all of it is a warning against misinformation. And if the overall tone of a book about statistics is "There are so many different ways in which statistics can be used to mislead us," and there are no examples that I can think of in which statistics can be used to tell us some useful truth about the world... Well, that's a bit of a problem, because the overall tendency then just reinforces the cliché of "lies, damned lies, and statistics." And then from there, you're into alternative facts. You're into, "We've had enough of experts." It goes to a pretty dark place. So, that's my first criticism.

The second criticism is *ad hominem*, but I think it's relevant. It's that Darryl Huff, after writing this book, fed in with what you might call a bad crowd. So, he started working for the tobacco companies, trying to criticize the statistics and the

epidemiology showing that smoking dramatically increases your risk of cancer and heart disease and so on, even testifying front of a Senate committee.

And it's quite revealing the way he performed that testimony. So, he comes up with this cute example, which is in *How To Lie With Statistics*, of how there's a correlation between storks and babies. And it's just one of these spurious correlations. And it's all very funny and it's clever and "Hey, look. Statistically, I can prove that storks bring babies." And then he drops the bombshell and says, "Well, it's the same with smoking and cancer. Sure, there's a correlation, but it's just one of these spurious correlations."

Which is such a shocking thing to say. And as we now know, very wrong. And for me, I realized... it can still be a brilliant book despite that very dubious career choice, that Daryll Huff embarked upon. And I think it still is a brilliant book, but I think it's relevant that that's where he ended up, because I do think that is also where this systematic, excessive application of doubt can also get you. You start with, "Well, the numbers can sometimes be used to mislead us." And you finish up saying, "Yeah, I know there's this epidemiologist telling me that smoking will give me cancer, but you don't expect me to believe any of that nonsense, do you?" So, that's my criticism.

Julia: It really was shocking to me. I had never heard that about Daryl Huff before, and it does throw the book in a whole new light. The analogy that occurred to me as I was reading your introduction was, if there was this series of beloved crime novels by some author, about serial killers -- and then you discover the author himself was a serial killer -- it makes it hard to enjoy the books again.

Tim: Yeah. And Hitler did these wonderful watercolors. Yes. It shouldn't influence your assessment of the book, but it's not irrelevant.

Julia: His points about how one can lie with statistics are still accurate. It doesn't really cast those in doubt. But it still just makes the tone of reading the book a lot different, if you felt originally like you were being led by this kindly guide who wants to help you see through the misinformation -- then it turns out, he's actually the culprit himself. It just is a different book, is all I'm saying.

Tim: Absolutely. And I think the reason that it is relevant is because there was this process over the second half of the 20th century of, the phrase that's often used is "weaponizing doubt," right? So, you start with this idea of, "Well, statistics aren't always accurate. Science is imperfect. The experts sometimes get things wrong." All of which is absolutely correct. And you just take that and you turn it up to 11 and you suddenly get to this idea of "The experts are always wrong. Don't believe the numbers," et cetera, et cetera. That's been a very powerful way of casting doubt on scientific evidence over the decades.

Julia: Right. What it seems to me increasingly is that there's just so much wiggle room in what you can get away with believing. And the definition of motivated reasoning -- which I cite in my book, and I think you also cite in your book, because it's just so good -- comes from Tom Gilovich, where he says, "If you're evaluating something you don't really want to believe, then you evaluate it through the lens of *must* I believe this, looking for any excuse to reject it. And if you're evaluating something you want to believe, you instead evaluate it through the lens of, *can* I believe it, looking for any reason to accept it."

And so, for something like, Should we suspect this correlational link between cigarettes and cancer is a sign of a causal link, and not something else? Well, it sure seems to me that a reasonable person would conclude: "We have this very plausible mechanism by which cigarettes would cause lung cancer. So, it's not just random, we didn't just pick two random things and find a correlation between them."

And so, that really should be a prime hypothesis in your mind for why we see this correlation -- but it's not definitive proof. There's always still at least a little bit of wiggle room where someone who is motivated to reject a claim of causal links between cigarettes and lung cancer can say, "Well, correlations are often spurious. Here's a bunch of examples. We can't prove it. We shouldn't jump to conclusions."

You can do that. You're not *forced* by the evidence to conclude that there's a causal link. And so, it's just very easy to take advantage of that, consciously or unconsciously.

Tim: Yeah. And it is, in fairness, it's very hard for us talking in 2021 to think ourselves back to the mindset of 1951 when this was quite a radical new hypothesis. And the main advocate of the idea that cigarettes were bad for you was a gentleman called Adolph Hitler. And really, with enemies like that, who needs friends. In terms of the sort of emotional valence of that, wasn't really very helpful.

And so, although some German scientists had... And no one really wants to look too closely into their experimental methods, but some German scientists in the thirties and the forties had found a link between cigarettes and cancer. But it's just a blank slate; you just have to start again, because nobody wants to pay attention to any of that work for a second.

But there's a wonderful... I remember the Gilovich example early in *The Scout Mindset*. I think it's in the, when you're telling this wonderful story about the Dreyfus affair and the different modes of reasoning that different people on different sides of that argument were using.

But there's a nice example of the contrast between being in this motivated reasoning mindset and being outside the motivated reasoning mindset, in the history of cigarettes and cancer, and that link. And it's Austin Bradford Hill, who... Austin

Bradford Hill is this British economist who became a medical statistician. And the footnote about his life which I love, is: He was nearly killed by tuberculosis and then ran one of the world's first controlled clinical trials demonstrating an effective treatment against tuberculosis. So, it's evidence-based medicine as revenge, which I liked.

But he was one of the people doing this early work, investigating what's causing lung cancer and coming up with the idea that it may be something to do with cigarettes. And he is confronted at a party by a Doctor. Because Bradford Hill and his colleague, Richard Doll have sent questionnaires to, I think, 70,000 or 80,000 doctors around the UK, quizzing them about this. Most of those doctors are smokers. Most of them know what this is about. Most of them don't like the idea that their habit is killing them, and so, they're looking for reasons to reject it and they're taking it personally.

And so, he's confronted at this party and this doctor says, "Oh, so you're the chap who wants us all to quit smoking, are you?" And Bradford Hill's response is perfect. He says, "No, I just want to see how you die. If you quit smoking, I want to see how you die. And if you don't quit smoking, I want to see how you die. So quit or don't quit. I will chalk your death up anyway, and it will be very useful to me." And quite apart from being just the most devastating comeback, it's real scout mindset stuff, isn't it? That's like, "No, I don't have a hypothesis. I'm just exploring. I just want to see the data." And it was really so good, so good.

Julia: Yeah. I was so jealous that you found that story for your book and I didn't find it for mine. It's perfect. Well, let's talk a little bit more about motivated reasoning, because this is the focus of your first chapter, your first lesson about how to evaluate statistical claims. And you basically say, "Just learn to notice and try to correct for influence of whether you want claims to be true or false and what that's doing to your way of reasoning about them." Could you just talk about how you actually do that in practice? Because I feel like that's really where the challenge is, it's easy for people to say, "Oh yeah, I should correct for my motivated reasoning." but to actually do that in practice is a much different beast.

Tim: Yeah. I think it's a habit of mind that one needs to try to cultivate. And I'm not sure exactly how you cultivate habits, but I think I'm managing to do it. So, the piece of advice that I give specifically... The book has 10 rules for thinking more clearly about using numbers. And so, this is rule number one, and it's notice your own emotional reactions. I think I say, "search your feelings," for the Star Wars reference, but notice your emotional reactions.

Now, I suppose my untested hypothesis is, if you notice your emotional reactions, they will then have less unwelcome influence on the way you reach conclusions. If you say to yourself, "Huh, I saw this thing on Twitter and it really made me feel frightened." or "It really made me feel, yeah, I knew those bastards would do that."

or "It made me feel joyful." Or "It made me feel defensive. That can't be true. That's fake news. Yeah, no, there's no way that's right. There must be some trick."

... Just getting into the habit of noticing that you're having an emotional response, which is very common because, of course, social media and media, it's all designed to produce emotional responses. That's what this whole system thrives on. My hypothesis, which I can't prove, or at least I haven't... Let me rephrase that. My hypothesis, which I don't think I've tried to prove, and I'm not sure how one would prove it -- but my hypothesis is, if you take a little moment to notice your emotional reaction, if you get in the habit of noticing your emotional reaction, you will, at that point, go back with a little bit more calm and you will be more likely to evaluate on a rational basis. You'll be able to set the emotion, at least to some extent, to one side. And I might be wrong about that.

Julia: So, what's interesting for me about that, is that what you just described totally matches my experience, that I often notice that I am eager to disprove something, or I noticed that I feel really validated when I read some claim that confirms what I already believe. And I do think I've gotten better at noticing, I think, really just by wanting to notice and wanting to get better at noticing, it just makes it more salient to me and I'm more likely to have that moment of self-awareness.

And just as you say, I find that this often leads to me deciding to be a little bit more skeptical or just taking a moment before hitting the retweet button to double check the claim, because I'm aware, "Oh, this is one of the things that I often realize in retrospect I was too hasty to share before checking because I wanted it to be true." Which is just as you described, but --

Tim: Well, on the flip-side, sometimes a bit less skeptical or maybe a little bit less cynical and a little bit less reflexively dismissive, because there are certain claims that you look at and you go, "There's no way that's true."

Julia: That's right. That's the flip side, exactly.

Tim: And that also, maybe one should try to correct for that, because that's also an instinctive knee-jerk reaction. It's not straightforward because, of course, often your knee-jerk instincts are correct. But still, I think it is worth observing them.

Julia: Yeah, no. So, I completely agree in the sense that that matches my experience, but I've really struggled to get this to work for other people. So, I, just to share my personal experience with this, I used to teach these classes at... So, I co-founded this nonprofit about improving reasoning and decision-making skills.

And one of the classes that I taught at several workshops was on just noticing your own motivated reasoning, and it bombed. It was my least successful class ever. And one of the centerpieces of this class was, I had my students go through this list of headlines or claims that were, they were all at least a little bit politically or

ideologically charged. And I just wanted them to notice as they read through the claims, to notice, "Do I want this to be true or false?" just to have that practice of noticing their preferences about the claims.

And there was just such a difference. Some people, but a minority of the class, had the experience that I did when I read through these claims of like, "Oh yeah, I have a preference about most of these claims." Like, "Oh, I want that to be true. I don't want that to be true." But then another, an even larger section of the class were just like, "I have no preferences about these claims. I just want to know what the truth is."

And some of these people, I've seen them arguing before, and I know that they have motivated reasoning. They're not just rare exceptions to all of humanity. But they weren't sort of conscious of that while they were reading. And I don't know, maybe just the claims I chose, somehow, I didn't manage to hit ones that they had motivated reasoning about. That's possible.

But it came to seem to me like a lot of people just don't have the conscious experience of the emotion of wanting something to be true or false, even though on some level that must be working in the background of their mind. Anyway, I guess I'm just wondering whether you've noticed any variability in people's self-awareness about motivated reasoning.

Tim: It's a really interesting question. There must be, but a lot of what I do as a journalist is to put ideas out in the world. Some of the ideas are fact-checking, and some of them are just kind of, "Here's a thing you might want to think about," or "Here's an idea I have."

And so, there's not much comeback. People occasionally send me emails telling me I'm brilliant or I'm an idiot, but I'm not out there giving seminars. There's not much people coming back and going, "I tried your method and it works for me." or "It doesn't work for me." So yeah, I would say it remains untested.

Although I find your instinct interesting. And it probably reflects the fact that you're a better person than I am, Julia. What you're saying is, does this work for other people? Whereas my instinct is, I just want it to work for me. I've given up on the project of improving anybody else, I'm just trying to get my own head straight.

And it is intriguing -- How To Make The World Add Up is the UK version of the Data Detective. And that came out in September. So, I've had several months of giving talks, and then I do get this sense of feedback. So, people have questions. And a lot of the questions are, "How can I get my dumb friend to stop being dumb?"

And my response is, "Well, I don't know. It's a hard problem. I think if you can get yourself to stop being dumb, that seems to be the project of a lifetime. I'm just

trying to get myself to stop being dumb. You get yourself to stop being dumb, and then maybe we can worry about stopping other people from being dumb at some other occasion."

But there's an appetite to fix other people. I find it interesting that Adam Grant's book *Think Again*, which is a very good book, came out the same time as the *Data Detective*, which overlaps with the *Data Detective*, has some overlaps with *The Scout Mindset*, but also has a lot of stuff, which is like, "Here's how to persuade other people who are thinking wrong thoughts."

And Adam's very good at this, and he explains it very well and it's all evidence-based, but it is intriguing. There is a much bigger appetite to change other people's faults than there is to fix our own faults it seems, at least when it comes to reasoning. When it comes to diet and exercise and makeup tutorials and so on, we want to fix ourselves. When it comes to fixing our reasoning, suddenly we seem to flip it around and go, "No, I just want to fix everyone else's reasoning. My reasoning is fine."

Julia: Yeah. Well, I assume just because it's so much less visible. We can see our own paunch growing during the year of COVID, but it's much harder to see the ways in which our reasoning is flawed or biased.

Tim: Yeah. It's hard to see your stupidity in the mirror.

Julia: Yeah, exactly.

Tim: Well, if I may briefly name drop, I was interviewing Daniel Kahneman earlier this week, this is the joys of writing for the *Financial Times*. You get to do these things, but Kahneman was mentioning naive realism, which I think is a very interesting idea. Which for those listeners who are not familiar with the idea, naive realism is just your strong, intuitive sense that you are perceiving reality in a straightforward and unfiltered way. And therefore, anyone who thinks differently must be wrong.

The old joke is, you're driving along the highway and anyone going faster than you is a maniac, and anyone going slower than you as a moron. And, of course, because you are driving at what you perceive to be the appropriate speed for the conditions. If you didn't think that was the right speed, you'd be driving at a different speed. But naive realism is this very powerful illusion, and it does make it hard to imagine having to fix your own cognitive processes.

Julia: Yeah. Really the only... Well, I don't know about the only. One of the more effective ways I've found to combat that is just imagining taking this outside view where... Well, I'll give an example. I was in an argument with a colleague of mine when I was doing this nonprofit in Berkeley. We were arguing about, I think it was about whether we should run a workshop very soon or whether we should wait a longer time and just do more preparation and practice.

And she thought we should run the workshop soon. I thought we should wait and we weren't ready. And we were arguing about it and it seemed so clear to me that I was right and that she was wrong. And, as it turned out, she ended up overruling me and we ran the workshops early and they actually went really well.

And I was shocked in retrospect that she had been right about that. But then I took this outside view and I imagined just being a third party looking at our disagreement and trying to decide to myself like, "Which of these two people is more likely to be right?" And as the third party, I noticed, "Well, Julia doesn't have any more experience in this topic of running workshops than Anna does, and they're both smart people. And so, I don't know, it's not clear to me that either one of them is more likely to be right than the other."

And so, to be clear, I didn't do that thought experiment at the time, I did it after the fact. But if I had done it at the time, I think it might have broken down some of my sense of naive realism that "Well, clearly the way I'm perceiving things must be right." I didn't actually have any real outside view reason to think that my intuitions would be more likely to be right than hers.

Tim: It's very interesting though, because what you're describing there, which makes perfect sense is, you've already taken one step down the process to clearer thinking. Which is, you're having a disagreement with somebody. So, if the two of you had both been of the opinion that you should delay the workshop -- which you're now saying was probably not the right decision -- It wasn't necessary to delay, it was just time and expense. And you could have brought it tomorrow to market-

Julia: Better to fail fast, basically, and get more information.

Tim: Yeah. You could learn and it would have been fine. So, if the two of you had been in agreement, you might never have really realized that there was any alternative view even to consider. And so, that step of taking the outside view is, it becomes harder to do and that's where naive realism dominates -- because either you're operating by yourself and so, there's no check on what you're doing, or because the person that you're talking to is in agreement with you.

Julia: Shares your sense. Yeah, exactly.

Tim: One of the things that I've become more aware of since finishing the book, another possible check on our cognitive errors is this idea of just paying more attention to the truth, or to the question of whether something is true or not. So, this field of psychology is, I think, most famously occupied by people like David Rand at MIT, Gordon Pennycook at University of Regina, and they've done various bits of work where they just get people to slow down, or they prime people to evaluate whether a headline is true or not and see how that piece of priming then makes people more sensitive to the truth of all the other headlines they then consider sharing on

social media. So, in a way, I think that just reinforces the advice I give early on in the Data Detective, which is to stop and observe your emotional reaction.

But I think Pennycook and Rand would say, "And by the way, while you've stopped, take a moment to just think about whether this headline is true or not." Because we might well be sharing stuff on social media that if we thought for a second, we'd go, "Actually, that can't be right." But we don't even think. That's what their research suggests, we don't even bother to try to evaluate whether it's true. We're in a different mode of thinking.

So yeah, it suggests there are three basic reasons why you might be wrong about something, a claim that you're pondering amplifying on social media. The first reason is because you're not smart enough. You don't have the expertise. You don't have the knowledge. It's super-confusing. And therefore, you need more education, more training, more information. That's the most obvious reason.

My hypothesis is actually this stuff is not in fact as complicated as it often seems. And so, the second reason that you might get stuff wrong is because of motivated reasoning. You just really want something to be true, and that turns out to be very powerful.

But then the third reason -- this is the Pennycook and Rand idea is -- actually, you could spot the truth and you're perfectly willing to search for the truth, it just never even occurs to you to try. You're distracted. You're just scrolling through Twitter and you see something and it seems vaguely appealing, and you just retweet it and you don't even stop to think for a second. And if you did stop to think for a second, you'd go, "Oh yeah, 500 migrant caravaners arrested on the Mexican border wearing suicide vests. No, I don't... Why would migrants from Guatemala have suicide vests?" But we just don't stop and think. And that's the Pennycook and Rand idea.

But fortunately, although I think that's a partial challenge to what I'm talking about early on in the Data Detective, I don't think it changes the cure, which is a pause for reflection.

Julia: Right. No, I think that's a really important point. And it actually reminds me of a story that came to mind as I was reading your book, which happened, it was a few years ago now. It was on MSNBC, and Brian Williams, the news anchor was talking to some journalist from the New York Times about... Oh yeah, this was during the Republican primary, Mike Bloomberg's campaign, where he spent \$500 million running for president.

And so Brian Williams and this journalist were talking about what an insane amount of money that was. And they said, "Yeah, there was this tweet that went viral about this. Let's put it up on screen." And they put the tweet up on screen, and this person had tweeted, "Mike Bloomberg spent 500 million in his presidential

campaign. With that money, he could have given every American \$1 million and still had money left over." And so, they showed this tweet and Brian Williams was like, "That's just crazy that this is possible in America." And the journalist was like, "Yeah, it's really disturbing."

And it was shocking to me that, it wasn't even just a brief thing where they saw the tweet and went like, "Oh yeah, that's crazy." They planned this news segment. They took a screenshot and put the tweet up on the air and never did it occur to anyone on screen or in production that this makes no sense. I think the tweet might have said at some point, "Yeah. There's like 300 million Americans. And so, Bloomberg could give a million to every American if he spent 500 million on his campaign." And so, yeah, go on.

Tim: This is a very interesting kind of units error. My friend, Matt Parker, who wrote a wonderful book called *Humble Pie*, which is all about maths mistakes and their consequences, which I thoroughly recommend to everybody... He was trying to cope with people making this mistake. And it turns out it's quite common. It's to do with what the brain does with that million and where you carry it or you don't carry it. People find it confusing.

But yeah, what you should be doing is going, "Okay. There's 500 million dollars. There's nearly 500 million Americans, not quite, but ballpark. So, it's about a dollar each, right? A dollar 50 each. That really shouldn't be that hard.

Julia: That's the thing. I can kind of see a unit error thing happening in the human brain... but that sort of basic first step, of that little red flag going off, going like, "Well, that doesn't sound right," or "That seems like orders of magnitude too high," or orders of magnitude too low... It was kind of surprising to me that that red flag didn't go off. And I don't know if that's a matter of people just being less inclined to stop and ask, wait, does that make sense or is that true? Or if it's some kind of mathematical, I don't know what to make of that. I was kind of disturbed by that.

Tim: No, it's very interesting. So there's clearly some cognitive misfiring going on with the way that some people do that little piece of mental arithmetic, but I think there must also be some motivated reasoning going on.

So my hypothesis is -- this is speculative -- that there's a certain kind of person who imagines that the wealth that people such as Mike Bloomberg have is so enormous that if you were to redistribute it, everybody would have loads.

Julia: Yeah.

Tim: And as an economist, I know that's just not true. I mean, they do have enormous amounts of wealth, but there are just a few of them and there are enormous numbers of people in the world. And if you were to say, take Bill Gates' wealth, and

Jeff Bezos' wealth and so on and redistribute it – I mean, people would notice, they'd be like, oh, that's nice that I got a little bit of extra cash in my pocket, but it's not going to be life-transforming for most people.

So I think there's this mental model that some people have, that says: Of course, if you just split up Mike Bloomberg's money, you could give a million dollars to everybody. And it's absurd, that's not how things are... but if that's your mental model of the world, then you're not going to question the maths. Whereas if you've thought about this a bit more and you know as a sort of professional economist how the numbers tend to stack up, then you're immediately going to say no, that can't be right. So you will stop and think.

So you're right, it's not just about can you do the calculation? It's about whether the answer that comes out of the calculation feels intuitive or feels crazy.

Julia: Yeah.

Tim: There's a famous class of problems called cognitive reflection problems made famous by Kahneman. They're actually developed by the behavioral economist, Shane Frederick, and the most famous one is the bat and the ball question. So a bat and a ball together cost a \$1.10, the bat costs a dollar more than the ball, how much does the ball cost? And you just go, 10 cents. It's the instinctive answer, but not the right answer. It's not hard to work out what the right answer is – but the question is, do you stop and check or not?

So there's that tendency to stop and check, but you're much more likely to stop and check if your worldview says that doesn't seem like it could be true versus if your worldview says, yeah, that's absolutely true. Obviously with the bat and ball, no one has any preconceptions, but with a lot of the questions out on in the media, on social media, we do have a sense of what the answer should be like. And that may mean we double-check or we don't.

Julia: Yeah. Well, this is actually a good segue into another thing I wanted to talk to you about, which is, I guess a different kind of reality check, which is reality checking the statistic you read against your personal experience. Like, if your personal experience doesn't really match the official statistic you're reading, when should that make you suspicious of the statistic, and when should it cause you to go, “Oh, huh, I guess my personal experience was unrepresentative”?

Tim: Yeah. I think it's not a straightforward problem, but I think it's something that we should be aspiring to do. If you see a number or you see some claim and you go, “Well, that's not how it is for me,” then that's pretty lazy thinking. But the flip side of that, which is “I will just accept any statistical claim regardless of whether that seems to conflict with my personal experience” – I don't think that's the best way forward either.

The reason is that the statistics are... as you can tell, I'm a big numbers fan, but the statistics are often quite thin. They're not very descriptive, they're giving you a very partial view of the world. And sometimes they're just wrong. But even when they're not wrong, sometimes they're misleading in one way or another that's worth thinking about. And often your own personal experience can help you get into, what is this statistic telling me and what is this statistic not telling me?

Julia: Right.

Tim: So an example I gave in the book is just the experience of traveling in absolutely packed conditions on the London underground, versus the statistics that say "Actually on average the London underground isn't that busy. So there's typically, I think it's something like 80 people on a London underground train that kind of thing."

Well, that's like sort of a dozen or even less each carriage. So that's pretty roomy, whereas it's standing room only when I'm there. So what could explain that? So maybe my personal experience is completely unrepresentative, but I don't think so. I'm just traveling in the rush hour and it's busy. Maybe that maybe the numbers are wrong, but the numbers are probably right as well -- so you look into where the numbers come from and you go, that's probably right.

And actually what you realize is the issue is, do you measure it from the point of view of the train or from the point of view of the passenger? And if you, if you imagine a very unevenly loaded underground system where people are traveling on these very, very, very busy trains, and then the rest of the time the trains are pretty much empty, well, the typical passenger is traveling on a very busy train, but the typical train is mostly empty.

And it seems paradoxical, but if you think about it, you go, no, it's not paradoxical at all. And then you can go, well, which is the right number? Should I be measuring the experience of the typical passenger or should I be measuring the experience of the typical train? And, well, it depends what you're trying to measure. Customer satisfaction, you probably want to measure the typical passenger's experience. But if you're trying to think about whether the network is efficiently used and what the environmental impacts are then you need to understand what's going on on the typical train.

So that's just an example of: My personal experience conflicts with the numbers, I need to think it through. And when you think it through you go, oh, I now just understand the whole system a bit better.

Julia: Yeah. That was a great and instructive example. And then you have this other example, which is a little bit more of a caution against rejecting stats that conflict with your personal experience. And that was the type of thing that happens when someone reads a statistic saying that well, cigarettes cause lung cancer, and then

they're like, "Well, but my grandmother smoked cigarettes her whole life and she never got lung cancer, and so I don't know about this statistic about cigarettes and lung cancer." And you correctly say that's not a very strong reason to reject the statistics.

And so I was comparing those two examples and thinking about what is the difference? Like why is the first case one where you should be a little suspicious or at least curious about the discrepancy, and the second case is one in which, no, your personal experience really should not cast much doubt at all on this statistic. Well, I have some thoughts about what kind of statistics would you give?

Tim: Well, I think one thing to simply think through is, is it possible that both of these things could be true? Is it possible that an old lady could smoke all her life and not die of smoking related diseases but at the same time that cigarettes could increase your risk of various diseases and shorten your life expectancy? And the answer is, well, yeah, there's nothing particularly troubling or paradoxical or problematic about that.

And it is worth pointing out that one of the difficulties with demonstrating the cigarette cancer link is that lung cancer is actually pretty rare. So although cigarettes increased your risk of lung cancer by I think it was 16 fold, 16 times a very low risk becomes a risk you don't really want to take, but it's still quite low. So a well under 1% risk goes up to somewhat over 1%, I forget the exact number.

So most cigarette smokers don't die of lung cancer. Heart disease is more common, but then there are lots of people who don't smoke cigarettes who still have heart disease. This was what really foxed people when they were thinking about it -- what is the nature of cause? If I can smoke cigarettes and not die of lung cancer, or I can die of lung cancer and not smoke cigarettes, what does it even mean to say that the cigarette smoking is causing lung cancer? Leave that one to the philosophers if you like, but this is what we're dealing with.

Julia: So maybe another way to rephrase that might be: If you imagine the statistic is true, how likely would it be that you would have encountered the personal data you have that conflicts with it? And in the case of the lung cancer, if we imagine it's true that cigarettes increase your probability of lung cancer by whatever amount, how likely is it that you would know someone who smoked cigarettes and didn't get lung cancer? Actually very likely.

Tim: Yeah.

Julia: But it's maybe not the most intuitive calculation for people to do.

Tim: Another thing that I think is worth asking is, and it's worth asking in general, but in this case in particular, where did this number come from? I'm being asked to take this statistic seriously, where did it come from? How did they gather it?

Was it just some stupid survey that somebody did in order to try to sell ice cream, which is where some statistics come from, or was it serious work done by serious people with resources? The National Health Service in the UK got data on how everybody died and there's a lot of data out there. You could start taking it quite seriously.

One of the examples of which I think statistics are often misleading is when they are used as targets or control variables. So you've got some manager of a system who looks at some statistics and says, "All right, we're going to use these statistics to motivate behavioral change."

So you've got the US news world report rankings of colleges, or you've got the UK government's example of trying to use statistics to increase the response times of ambulances. And these are cases where actually there are lots of ways to game the numbers. If the manager only observes the numbers and can't actually observe the behavior, there are lots and lots of really unhelpful ways to change your behavior to make the numbers look better, which I discussed in the book. But there's a mechanism why at that point you would say, look, you should be trusting the personal experience, you should be ignoring these numbers, because the numbers are just hopelessly distorted by this gaming behavior.

Julia: That's right. So if there's a motive for people to try to juice the numbers then that should increase your suspicion, your willingness to be suspicious.

Tim: And the conspiratorial minded among us would say there's always a motive. Somebody is always got a motive-

Julia: Yeah, but sometimes the motive is much stronger and more obvious than other times.

Tim: Yeah, absolutely. Sir Angus Deaton, who was of course, one of your interviewees- And I mean, that's obviously his career highlight, but he also has a Nobel Prize in economics as another-

Julia: Secondary highlight.

Tim: Yeah, secondary accolade. He has made the point that economic statistics start to get really tricky once serious sums of money are riding on what the statistic is. So for example, you can measure a poverty line and send out some geeks into the field, survey people's standards of living, trying to measure a poverty line... he's a real expert in this kind of thing.

The moment you say, oh, we're going to start sending money to everybody who's below the poverty line, suddenly, enormous pressure gets put on that number to move it down, to move it up, people really care where that number is. And it will

start to bend under the strain. It doesn't mean you shouldn't be trying to send money to people who are poor, but it's a real cautionary note.

Whereas when it's just a few academics just trying to measure a thing and that there are no stakes, you can have much more trust in their efforts to do that.

Julia: Yeah. I kind of think of Angus Deaton and actually another economist I've talked to recently, Russ Roberts, as examples of people who kind of had the reaction to statistics that you were describing when you were talking about *How to Lie With Statistics* which is that they've seen so many examples of randomized controlled trials being held up as this gold standard of proof and turning out in retrospect to have been really flawed, that they've kind of rejected the whole enterprise entirely in some contexts and I've kind of gotten into interesting debates with both of them about like, can you learn anything even hypothetically, from a hypothetically really well conducted survey or study?

And so I was on the side of arguing like, no, if it's like really well conducted and you have a lot of different independent confirmations then that should increase your confidence somewhat in the claim. And both of them were like, no, you learn nothing.

Tim: Yes. I don't fully understand the reasons for Angus's skepticism about randomized controlled trials --

Julia: It's possible that I don't either.

Tim: Yeah. There's one line of skepticism that basically says they don't ever measure anything that really matters. This is Lant Pritchett's idea. Like sure, run some randomized trials, but you're only ever going to identify niche variables of niche interest, and you'll never move the needle on all the stuff that really matters. Maybe, I don't know. And then there are other objections that we don't really learn anything from them, I don't fully understand.

I have to say though that Russ and Angus I would place in a slightly different point on the spectrum. And I admire them both as thinkers, but Russ, I think tends to my taste too much towards the radical skepticism. I'm not sure Russ has ever been persuaded by any number on anything. I'm probably placing that too strongly, but that's very much... and I think Russ would agree, I don't think I'm misrepresenting him. But his go-to view of any economic study, any economic data is, "Yeah, well maybe, but it's complicated, isn't it? It's complicated and can we really believe that?"

Whereas Angus is a real data specialist and has spent a lot of time really studying the numbers. And I think obviously that breeds a certain degree of skepticism. But I sort of feel if you felt that you could never learn anything from the data, then why has he spent so much time looking at the data? And his work with Anne Case on *Deaths of Despair*, that's data-driven, he didn't just do some interviews and say,

well, we don't really know, but this is our speculation. It purely came from the data. So he's skeptical about randomized trials, but he's not skeptical about data in general, or at least he's not radically skeptical about data in general.

Julia: That's right. No, that is a great distinction to make, because that's right. My debate with Russ was about, if you had a survey that found, hypothetically, that 99% of parents wish they hadn't had children, does that tell you anything about whether you would be likely to regret having children if you had children? And he said, no. And so that's not a randomized controlled trial that we were talking about that's just a measurement. Whereas Angus Deaton was specifically objecting to the idea of learning from randomized controlled trials.

Tim: There is a wonderful example, I don't think Angus would reject this because it's in medicine, but I don't know, maybe. There's a wonderful example of randomized trials being used really well during the pandemic, where a couple of Oxford academics set up something called a recovery trial. And the recovery trial is... it's sometimes called a platform trial. So normally when you think about randomized trials in medicine you think of, there's a new drug and we need to prove it's safe, so we do the safety trials. And then we need to prove it's effective. So we do a randomized clinical trial to show that this drug works in some way.

The platform trial flips that. Platform trial basically says, we've got this condition, severe COVID, people in intensive care. So we're going to start throwing possible treatments at them and we'll see which ones work. So you still run a randomized trial.

And the way the recovery trial was designed was to try to make this completely routine. So you've got doctors and nurses at the bedside of the severely ill patients. They get consent either from the patients or from the patient's family. And they're giving them various bits of treatment that, this is the treatment that we would recommend, and there's this other thing which is like a hydroxychloroquine, and some people think it would work and some people think it wouldn't work and we don't really know. So some of the patients will get it and some of the patients won't, and some of the patients will get dexamethasone and some won't.

And it's just all part of the natural patient pathway. You just click a couple of buttons on a website and it's all very low bureaucratic or administrative work. But because every patient who comes into any hospital in the UK with COVID is potentially enrolled in these trials, you're very rapidly learning.

And what they learned was dexamethasone, a very cheap steroid, works incredibly well. Now they think that's probably saved a million lives. I talked to Martin Landry who was one of the people who set up his trial and said, what does it feel like to have saved a million lives? And he was like, "Well, no, I'm not sure it's a million actually, I think it might be 800,000 or something." I mean, he wasn't really interested in the question.

He was interested in, how do we get this so that when patients come in to hospitals with any condition, we're automatically throwing in a bit of randomization? Where normally people would be given some antibiotics, but we don't know... like for three days or for seven days? Nobody really knows, so let's make it some people get it for three days and some people get it for seven days. And we collect the data and do that with enough patients... In weeks, you find out the answer to your question.

So it is possible to scale this up in a way so as to make a real difference, and one could imagine doing that for social programs too. But I'm sympathetic to Angus's criticism that if you run one trial somewhere of how to get people to use bed nets to deal with malaria, it's not clear how much you really learn about human nature or about policy interventions all over the world.

Julia: Right. Well, that version of the claim I would agree with too. So yeah, at that point we're in agreement.

Tim: Yeah. It's interesting because then you start thinking, actually, this is about data infrastructure or statistical infrastructure. Like, what are the systems we build to collect the data that we need to generate answers on a routine basis? Because if it's actually just about scale, well, scale is something that we're pretty good at, computers are everywhere, the maths behind this are not very complicated. It must be possible. It must be a solvable problem. And that's what the recovery trial, I think gestures towards.

Julia: So Tim, I had one more question I really wanted to make sure to ask you. I was gratified to see that you, in your book, struggled with a question that I too struggled with in writing my book, which is... you talk about the replication crisis in the social sciences and all of these reasons why seemingly solid randomized controlled trials or other often famous studies turned out to fail to replicate and were kind of discredited.

And so you're clearly very aware of the fact that it's hard to just trust a study just because it was published in a psychology journal or whatever. And yet you also do cite a bunch of psychology studies in the book, in making points about motivated reasoning and numeracy and all sorts of things.

And so I was just curious to talk about what heuristics you landed on for when to trust a paper enough to cite it in a book, versus not. Because I thought a lot about this too, and clearly you did as well.

Tim: Okay. So let me give you two answers. There's the kind of approved answer, and then there's the honest answer.

Julia: Oh, I love answers like this, go on.

Tim: I think the respectable answer, the approved answer is: Well, look, I would not lean on a single study to give me a counterintuitive result. Like “You might think that X is true, but hey, I found this study and it says that Not X is true, isn't that surprising and amazing, and give me a book deal.” That's not the right way to think about things.

So instead what I'm looking for is, okay, well, this study seems to support this other study which seems to support this third study, and the three of them are pointing in the same direction, but they're measuring slightly different things, they're measuring them in different ways. So even if one of them doesn't add up, they're all kind of pushing in the same direction. And they're also pushing in the same direction as my own experience, my own judgment. And so, add them together and you've got a collage of evidence that supports the case that I'm making.

So that I think is the intellectually respectable answer. It's not so very different from the honest answer, which I find slightly embarrassing, which is: “Well, look, I think this thing is true, and I found this study and it does too.”

And I realize there's a real danger there in that you ended up finding the studies that support the argument you were making and maybe not looking hard enough for disconfirmation. So it's always a risk.

What I think is interesting is that the intellectually respectable version of the answer I gave you, and I think the more human and more honest version of the answer I gave you... They're not very different, are they? I'm describing a similar process of combining multiple studies plus my own intuition, experience, and judgment. But in the one way that I framed it, it's all, “I have five reasons to believe this is true.” The other way of looking at it is, well, how hard did I search for reasons to not believe it?

Julia: That's exactly what I was just going to say. I think that's what distinguishes whether you're doing a kind of careful process of looking for multiple different sources of evidence before you're willing to cite something, versus the confirmation bias thing - is, Do you ever have the experience of saying, “Well, this thing seemed true to me, but the evidence actually isn't that great and so I won't cite it”?

Tim: Yeah. And there are examples of work that I've cited where it turns out it doesn't replicate.

Julia: Yeah. I'm not sure any science journalist has escaped that. Not sure there's a single one.

Tim: Yeah, from my book *Messy*, there's this really fun study that shows that there's a disfluency effect. So if people are given harder fonts then they actually learn more. Turns out it probably doesn't replicate. I wouldn't say that the entire thesis of *Messy* was based on that, and it wasn't, but it was kind of a nice example of the kind of

thing I was talking about. And so it's no longer a nice example of the kind of thing I was talking about.

These things do happen. It is also worth bearing in mind with all of these studies... This goes back to what we were talking about with randomized control trials... whether they replicate or not, they're always investigating quite a narrow claim.

Julia: Yeah. I'm glad you brought that up.

Tim: So going all the way back to this idea of, well, why do people get things wrong? Do they get things wrong because they're too complicated or do they get things wrong because of motivated reasoning, they're trying hard to reason in a tribal way? Or do they get things wrong because they're just not paying enough attention, they're not thinking hard enough?

I think the answer is, well, it's a combination of all three, but it depends what you mean by "things." Like, there are some things that we get wrong because we're not Einstein, they're really hard. There are some things that we are absolutely determined to reason about in a tribal way. And there are certain statements that might seem superficially plausible, but are obviously ridiculous on closer examination, and maybe those are the things that you would discover if you cause people to pause and reflect.

So with Gordon Pennycook and David Rand's work where they're focusing on people just not stopping to think and if you persuade people to stop and think for a moment, and in particular, you get them thinking about the question of truth, people are really good at spotting fake news headlines. I mean, that is sensitive to the kind of fake news headlines that you test. So people will spot the "Pope endorsed Donald Trump" thing really, really quickly. But I think if you gave them a more complicated question, it might not be the case that simply stopping and thinking and priming people to reflect on truth would solve the problem.

So all of these studies, whether they replicate or they don't replicate, they're always investigating a narrower question than we as human beings just moving through a complex information ecosystem are interested in. We're always looking for the big general truths, and that's not how social psychology or cognitive psychology works.

Julia: Right. That's kind of what I landed on in thinking about this problem myself that yes, you do want to, as you say, look to see if there are other studies that contradict the one you were going to cite that are not obviously worse than the one you were going to cite, and you want to see if it is very counter-intuitive and if so, maybe be suspicious, and so on.

But you also -- in reporting the study and in citing it, you don't want to make it sound like this one study is definitive proof of this very general claim about human nature or how things work in general. You want to report it with the right epistemic

status, essentially. So you want to make it clear in your wording that this is a *little* bit of evidence, for this general claim, but not --

Tim: Absolutely, and there are very famous studies -- sorry to interrupt, but like the Asch conformity experiment, the famous study Solomon Asch did with asking people to evaluate whether the lines were the same length. And actually there were a lot of actors in the room and they would all agree on the wrong answer. And then there's this one person, and does this person conform to what the majority have said?

And that study is from the 1950s. It's often reported as "People always fit it in with the crowd." That's not what Solomon Asch found at all. Arguably those are not even studies of conformity, because arguably very often what he found was people figuring out how to not conform in what was a socially awkward situation. And managing this difficulty.

But it's not just journalists who report them as conformity, they are often reported in psychology textbooks as studies of conformity that found conformity.

Julia: It's true.

Tim: So, now we've been dealing with the way studies get kind of boiled down and summarized. I've been dealing with this problem for a very long time and it's not going away, because we think in stories, and what's memorable and what's complicated and fiddly.

Julia: Well, I can't think of a better way to end this conversation than with the word fiddly. So I think I'll stop there and let you go. But Tim, thank you so much for coming on the show and discussing your book with me. It's been a delight.

Tim: Well, I'm a huge fan of the show.

Julia: Thank you so much, that's just wonderful to hear.

[musical interlude]

Julia: That was Tim Harford, author of the Undercover Economist column and the book, *The Data Detective: 10 Easy Rules to Make Sense of Statistics*. One thing I wanted to add to our discussion at the end of the episode about how do you decide which studies are trustworthy enough to cite, I think I'm even more skeptical of most studies than Tim is. I don't know if that came through in our conversation or not, but I really appreciated his recognition of just how narrow most studies are and I wanted to underline this because I think this point is neglected, even by most smart and knowledgeable people I know. When I was going through the relevant literature for my book, I discarded most studies that I read, and there were a few recurring reasons, but one of them was that the study was just done on such a

narrow and contrived scenario, I didn't really feel comfortable generalizing from it to anything in the real world.

So for example, one very widely cited study had people try to switch on and off a light bulb and then ask them whether they felt like they were successfully controlling the light bulb and then generalize from that to conclusions about people who feel like they have control in their lives. Another study gave people a little geometry puzzle to solve and measured how long they spent working on it and then tried to generalize from that to conclusions about whether or not people persevere in general in life. I just don't believe that those very specific and weird scenarios tell us very much about life or human behavior in general. It pains me to see people, very smart people with PhDs, citing a study like that to say, "Well, as research shows, people tend to persevere when blah, blah, blah."

That's my little rant about social science research for this episode. But anyway, you should check out Tim Harford's book, *The Data Detective*. It's very thoughtful and easy to read and there's a lot in there that we didn't cover in our conversation. And also his column, *The Undercover Economist*. I'll link to both of those.

That's all for this episode. I hope you'll join me next time for more explorations on the borderlands between reason and nonsense -- and for more angry rants about social science, I'm sure.