Rationally Speaking #210: Stuart Ritchie on “Conceptual objections to IQ testing”

Julia Galef: Welcome to Rationally Speaking, the podcast where we explore the borderlands between reason and nonsense. I'm your host, Julia Galef, and my guest today is Stuart Ritchie.

Stuart is an expert on human intelligence and cognitive aging. Maybe to introduce him, the best thing is just to read you his Twitter bio: “Stuart is a postdoctoral fellow at the psychology department at the University of Edinburgh, and looks like a cartoonish, startled hedgehog.”

Stuart, welcome to the show!

Stuart Ritchie: Hi! How's it going?

Julia Galef: I just have to tell you that what I love about your Twitter ... I mean, I love your Twitter presence overall.

Stuart Ritchie: Thank you. That's very kind.

Julia Galef: But one particular thing that I love is how the photo of you, in the photo that you posted on Twitter, you're sort of staring, horrifiedly, down at the lower right hand corner of your screen. Which makes it look, with every tweet you post, it makes it look like you're horrified by what you've just written, because you're staring at it in dismay.

Stuart Ritchie: Yeah. Well it is often something, you know, annoying or depressing that I'm putting up there. You know, someone misinterpreting a paper or there's a rubbish scientific paper, or some-

Julia Galef: Yeah.

Stuart Ritchie: Some kind of-

Julia Galef: And you're, like, covering half your face almost in a, "Oh, God" Expression.

Stuart Ritchie: Yeah. I also, like, I couldn't bear to put up a serious picture because I just can't bear the thought of posing for a serious picture and putting it online. So, that one seemed to work.

Julia Galef: Well you definitely avoided that failure mode, I'll say.

Oh, I should also add that Stuart recently published a great book with a somewhat misleading title. The title is 'Intelligence: All That Matters.' Now the reason that's misleading, I learned, is that when you first read the title you assume it's saying that intelligence is the only factor that matters in determining someone’s success. Or, a society’s success.

But in fact what I discovered is this book is part of a series of short explainer books called 'All That Matters’ — so they have titles like,

'Bioethics: All That Matters', or 'The Future: All That Matters' and so in context it's clear that “all that matters” refers to, “We're going to tell you everything important there is to know about this topic,” and not “This topic is the only thing that matters.”

Stuart Ritchie: It's extremely unfortunate. And I've had to, at the start of many conversations, get that out of the way. Yeah, so I'm glad you did in this case, yeah.

Julia Galef: Yeah. I was surprised at first. Because I was like, “I... don't think that Stuart Ritchie believes that intelligence is the only important thing determining success? Just based on what I've seen him post on Twitter, I'm surprised that he chose this title for his book."

Stuart Ritchie: I have a section in the book that's like, "All that matters? Actually, no, no absolutely not.” But that's just a slight unfortunate aspect of the-

Julia Galef: I know, it's just especially unfortunate for intelligence just because-

Stuart Ritchie: Yeah.

Julia Galef: People are so sensitive to that... Like, a lot of people think that that's something researchers of IQ science actually believe — that intelligence is, in fact, all that matters. So, it's kind of a straw man that the title inadvertently props up.

Anyway, so there's a couple of things I wanted to talk to you about. The first is conceptual objections to the idea of IQ testing. As opposed to, for example, ethical objections to the idea of IQ testing.

The first, and maybe most basic, objection that many listeners will probably have heard before is: we can't even define intelligence, so it doesn't make sense to have a test that's claiming to measure it. What do you say to that?

Stuart Ritchie: Yeah. Well, I think the standard objection to that is that you can have something which is useful, and a useful tool, without actually having a definition of exactly what it measures.

The kind of classic example is the example of temperature, right? So, we knew that the mercury rises in the thermometer when it's hotter, and falls when it's colder. Long before we had the kinetic theory of gases that tells us exactly why that's happening, that the movement of molecules is what underlies temperature. Thermometers were extremely useful way before we had that theory. So, having a definition isn't necessarily useful for all purposes.

Of course, it would be great to have a brain-level, or synapse-level, definition of intelligence. And that's something we're kind of working towards. But that doesn't mean that the actual IQ tests aren't measuring something that we can then use to predict people's success in education,

or in occupations, or in every day life, and so on, to some degree. They absolutely do measure that.

So, when people say “You don't have a definition of intelligence, and intelligence is just defined as what the tests are testing” — That's kind of okay to me. That seems absolutely fine. We know that this is a useful thing, and we can use it to predict stuff. That's kind of all we're claiming at this point.

People have come up with definitions. They've tried to give definitions that range from quite vague, to trying to be more theoretically sound... But at this point I think the data just aren't good enough to really make a strong theory of exactly what intelligence is. And I think we would be kind of wasting our time if we try to do that. At some point we'll have the data that we can really make a good stab at that, but I don't think that's now.

Julia Galef: So, a variant on that objection would be the argument that there are multiple different types of intelligence. That IQ just captures one facet of what we colloquially see as intelligence. So, for example, there's also emotional intelligence. There's creativity, spacial reasoning, et cetera. So, this argument goes, IQ scores may predict some important things in life — but we shouldn't talk about it like it's the only, or even the most, important kind of cognitive trait.

Stuart Ritchie: Yeah. You know, that's ... If you just want to define stuff as intelligence and call stuff intelligence, then that's totally fine. I think that's absolutely okay.

So, if you look at someone like Howard Gardener, who is a professor at Harvard who came up with the idea of multiple intelligences, who has all these different intelligences... like there's analytic intelligence, which is kind of more like the stuff that we talk about when we talk about IQ tests. But then there's also a whole bunch of other stuff like interpersonal, or intrapersonal intelligence. And he's added new ones, like ... I think there's one called naturalistic intelligence, and existential intelligence, and all these kind of-

Julia Galef: Existential intelli ... I'm sorry. I have to pause you and ask what that is?

Stuart Ritchie: Well, it's not like it's defined on the basis of data or anything. I mean, Howard Gardener kind of has just decided that these intelligences are intelligences.

So, this one, he says that it's the use of collective values and intuition to understand others and the world around them. It's about seeing the big picture. That is existential intelligence.

There's also a whole bunch of other ones. There used to be seven, I think there are nine now. And instead of collecting loads of data and running factual analysis to see where the patterns in the data lie, which is what generally people in the world of intelligence do, Gardner has kind of come up with these things which he kind of thinks are intelligence.

And then there's other people who say things like, "Well, there's also common sense, which is separate from intelligence. And rationality, which is separate from intelligence. Wisdom."

Julia Galef: Street smarts as opposed to book smarts, for example.

Stuart Ritchie: Yes. Yes, exactly. And you can totally define these as intelligences if you want, and that's fine. But I think a lot of these involve things which we wouldn't necessarily consider intellectual abilities.

So a lot of them can involve things like personality. So, for instance something like emotional intelligence — clearly there's going to be like an analytic intellectual aspect to that. Which is your speed at understanding other people's emotions, and your ability to represent their mental states. Some people are going to be better at doing that then others.

But also there's an aspect of personality in there. An aspect of how much you like talking to other people, and how much you can intuitively empathize with other people, and so on.

So, it seems to me that there's a useful separation between the analytic type stuff and the personality type stuff. Those are kind of the two prongs of individual differences in psychology. There's intelligence on one hand, and there's personality on the other.

Emotional intelligence is kind of like a perfect mix-up of the two of them, in a way. It kind of includes aspects of both.

I think it's probably more useful to talk about them separately. In fact, there are studies showing that once you take into account personality, as measured on personality tests, and intelligence as measured on IQ tests, emotional intelligence doesn't predict much more of people's job performance and so on. So it's a kind of combination of the two of those.

If you want to call lots of other things intelligence than that's fine. I guess in my view it's more useful to describe a lot of these things as “skills.”

So, the sort of naturalistic intelligence that Gardner talks about is like an intuitive understanding of the natural world, and enjoyment of the natural world. That's something you can work on, and some people clearly have more interest in that than others. But I would call that a skill or an interest more than intelligence.

That is totally, like, immaterial to much of the science here. This is just people playing around with definitions. What I'm really interested in, and what most intelligence researchers are interested in, is the really rigorous stuff, where you've collected loads of data, you've done analysis to work out what goes with what.

In intelligence it really is the case that when you give people pretty much any cognitive task, they do correlate positively together, and you get this general intelligence factor, which explains about half of the reasons

people differ across all mental tests. That general intelligence is something which people don't really like to think about too much. But it's there, and this is what Gardner's theory is often held up in opposition to.

So, I think it's absolutely fine if Gardner want's to say there are different, multiple intelligences. But it's when you say there's "no such thing as general intelligence" that you get into trouble, and you come up against the data.

Julia Galef: So, to push back on that a little bit: Let's say we just agree that all the cognitive traits, verbal skill, math skill, creativity would count as a cognitive trait I think. Let's say we agree those are all correlated with each other, and with IQ scores. Still, isn't it the case that you could design a variation on the IQ test that would give more or less weight to, say, creativity compared to math skill? Then the new score — call it IQ Prime — would still be correlated with all these different cognitive skills or traits. But it would be more correlated with some, and less with others.

So, isn't it true that we're implicitly prioritizing some types of cognitive trait or skill over others, just in the way the test is designed?

Stuart Ritchie: To some degree, but not once you get to a really, really good intelligence test which includes loads, and loads, and loads of different cognitive abilities.

So, there are two papers by Wendy Johnson, who is a professor here at Edinburgh. She found these amazing data sets with people who had taken dozens, and dozens, and dozens of cognitive tests, that covered everything from how quickly they could check letters off a sheet, to vocabulary tasks, and every thing in between. Including arithmetic and mental rotation, and absolutely every thing you can think of that would come under the rubric of a cognitive task.

And she looked at the general intelligence factor that you can extract from these data sets. So, you know, you had Intelligence Test One, which has maybe 12 tasks. And then you had Intelligence Test Two, which has another completely separate 12 tasks. The general intelligence factors that you extract from those data sets are almost identical. That is, when you take just what is common to lots of different cognitive tests — even if they differ quite a lot — when you take just what is common, that general intelligence factor seems to be something very similar across all different tests.

So, the "g," general intelligence factor that we're extracting from those batteries, those very different batteries of tests, correlated together at like .99 or sometimes one. They were identical for all intents and purposes.

So these two papers are called... there's one called 'Just One g" and then she replicated it again in another sample and it's called 'Still Just One g'.

Julia Galef: I really hope she keeps this up.

Stuart Ritchie: Yeah.

Julia Galef: "As ever, just one g."

Stuart Ritchie: Yeah. Yeah, yeah. Exactly. When you get to the very high level there really does seem to be just one g.

At lower levels, if you just give people a couple of tests — you know, if I gave one sample of people just verbal IQ tests, you know, a vocabulary test, and like a similarities test where you have to say, like, "What do a dog and a rabbit have in common?" that sort of test... Then I gave another sample of people just tests that were to do with speeds — so, like, press your finger on a button when a light goes on, that kind of test.... Then obviously we would expect that those would be similar general factors that came out of those tests, but we wouldn't expect them to be the same.

But the point is, that if you aggregate across a whole range of mental skills, there really does seem to be something that's common to them all. That's what is theorized to be the general core of intelligence.

Now, there's lots of different theoretical reasons as to where that g comes from, and there's developmental theories of that. People have been arguing about this literally since 1904. But, at a high level the general factor does seem to be the same.

So, in answer to your question, yes it is right that if you give someone a test that's biased towards one particular kind of item, that is, one particular type of cognitive skill, then sure you're going to have problems.

But on the other hand the best intelligence measures are kind of a "the more the merrier" perspective, where you're just adding in as many cognitive skills as possible.

Julia Galef: I think the part I'm still having trouble with is this idea of what's "common" between the different tests.

So, like, I could imagine two people, Devon and Taylor. And Devon's two standard deviations above the mean on creativity, but just average at mathematical reasoning skill. Taylor's the opposite — two standard deviations above the mean at math, and average on creativity.

Now, you give them both a bunch of tests that test both creativity and math skill. Do their IQ scores end up the same? Or is there a way ... Is the question of who gets a higher IQ score sensitive to how many creativity tests you give one versus the other?

Stuart Ritchie: Yeah. So the type of test is important too. So, as you know, there are some tests which seem to measure general intelligence better than others.

So, something like vocabulary seems to have a very high, what we call a "g-loading." So it's very strongly correlated with the eventual g-factor that you get out of that test, plus a whole bunch of others. It seems to be a very good indicator of general intelligence.

And there are others which are less good. Something like a simple reaction time, just pressing a button when a light goes on, doesn't have a very high g-loading, because it's a very simple task. It's not ... It doesn't hit that complex thinking that g is supposedly measuring.

But, you know, another way of looking at that question is people have different interests, and people have different skills. The general factor only accounts for half, around about half of the variance in most batches of cognitive tests.

There's a whole bunch of variance that's left over, which might be to do with people's specific skills on particular kinds of tests. Their interests might have lead them to develop a particular skill more strongly then others. They maybe let their math skills slide a little bit while they've been focusing on verbal stuff.

That might be the result of a kind of initial preference for that stuff. Or an initial higher skill level in those areas. But through a kind of set of feedback loops it means that they end up spending most of their time doing verbal stuff instead of math stuff. Other people end up doing more maths stuff and more verbal stuff.

So, it looks like they have big differences in their cognitive abilities. But actually it would be possible if you had ... Theoretically, it would be possible if you had gone back and maybe nudged that person more towards math stuff, they may well have come up with the same level of ability at the end.

So, that's, I think, a common kind of misconception — is when people talk about general intelligence, they think that that's all the variation, when actually it's just about half of it. There are specific skills too. Those are interesting to think about as well...

Julia Galef: Okay. I'm going to give you one more variant on this class of objection.

I'd say this is probably the most frequently cited critique of IQ science among people that I know. It's a long essay by someone named Cosma Shalizi in which he basically argues IQ is just a construction. I have a quote from him here:

"The correlations among the components in an intelligence test, and between tests themselves, are all positive because that's how we design tests. So, making up tests so that they're positively correlated, and discovering that they have a dominant factor, is just like putting together

a list of big square numbers and then discovering that none of them is prime. It's a necessary side effect of the construction, nothing more."

So, this is basically a more formal way of saying the old quip that the only thing IQ tests measure is how well you do on IQ tests. What's wrong with that reasoning?

Stuart Ritchie: I think anyone can come up with a scientific finding and say, "Well, that's a boring finding" if they want to. That's, you know, that's their opinion. But actually I think it is interesting that no matter what people come up with, no matter what kind of cognitive skill they try to measure or test, it's correlated with other cognitive abilities.

People have deliberately tried to come up with new cognitive tests that don't correlate with IQ. As I say, I know there are some that have better or worse g-loadings, that is, their correlations with the general factor. But, in general that's really, really difficult to do. You can't find cognitive abilities that aren't related in some way to g. It just gets into everything.

And I think that's a really interesting finding about human psychology — that, to a very large extent, intelligence seems to be general ... Or, ability seems to be general. You can say, "Well, we designed the tests that way," but actually, people have tried their best to design tests that aren't that way, and it still falls into this correlation.

Julia Galef: I think that's the crux of disagreement between you and the Shalizi camp on this issue, is that they seem to think that there are tests that we have devised, or we could have devised, that are not correlated with the IQ tests.

Stuart Ritchie: And you know, go ahead and collect some data on it. I'd love to see that.

Julia Galef: Yeah, it's interesting.

Noah Smith, the economist, who blogs at Noahpinion and writes for Bloomberg, he proposed a theory a few years ago in response Shalizi's essay, for why it could be possible for there to be multiple different intelligences that are not actually correlated with each other, despite the fact that all these cognitive tests are correlated with each other.

His reasoning was: suppose that there are these various simple mental skills, or basic mental skills that are substitutable for each other. So, quote:

”Suppose that any simple information processing task could be solved using spatial modeling, or solved using symbolic modeling, or some combination of the two. That would result in a positive correlation between all simple information processing tasks without any dependence between the two mental abilities." End quote.

Does that seem plausible to you? Do we have any evidence of whether or not that's true?

Stuart Ritchie: So that's not Noah Smith's theory. That's Godfrey Thomson's theory from the 1930s.

Julia Galef: Maybe it was independently discovered.

Stuart Ritchie: Yeah, possible. But that's known as ... Godfrey Thomson, who was an intelligence researcher and statistician at the start at the 20th century, and was very influential, for instance, on the British education system... he came up with what he called the 'Bonds Theory of Intelligence,' which was in opposition to Spearman's g-Factor theories.

Charles Spearman was the guy who first noticed the positive correlations of the tests, and his idea was, 'Well, there's just one underlying general intelligence,' some kind of, what he called “mental energy” that causes the factors to ... the skills that correlate together.

And that is a perfectly viable explanation of the correlation from the tests. It's just that there are other explanations, too. And that theory that you just described there in the quote from Noah Smith is what Godfrey Thomson came up with.

And they kind of argued with each other for a long time about it. So the brain might have lots of different kind of very basic processes, and that they are different, what he called samples, by the IQ context that are given.

Essentially, the debate has continued, and we still don't know whether Spearman or Thomson was correct about that, because there doesn't really seem to be a good way of testing that.

I guess the first thing I should say is that this doesn't actually matter in terms of the practical consequences of intelligence tests, right? So we know that Thomson would have completely agreed. And even someone like Cosma Shalizi would agree that there is a general factor, that it's there and summarizes peoples' performance across lots of texts.

Now the reason for why it happens is a different question — but that general factor that's there is still predictive of all the stuff that we know IQ tests are predictive of. And so it's still useful as a measure.

This is a kind of an orthogonal question to that, which is, "Why do these texts correlate together?" And we're not gonna have evidence on that until we have a better brain level understanding of intelligence.

There's papers from ... mathematical modeling papers from 2009, 2010 that my PI worked on. He's a big fan of Godfrey Thomson, and so he thought, 'Well, we'll try to run some mathematical models that compare Thomson's theory to Spearman's theory.'

And the conclusion of those papers is they are both completely viable models of how the g-factor exists. So, we don't really have evidence to kind of clench that debate, on either side. But I do think it's a mistake to

put that up as a response to someone who's saying, "IQ tests correlate with loads of stuff, and are predictive of loads of stuff, and are correlated with biological variables and so on."

I think a lot of people link to that Shalizi essay, and link to the idea of the Bonds Theory of Intelligence and so on, as a kind of a knockdown against the practical validity of IQ tests. And that's completely irrelevant to that practical validity.

Julia Galef: Yeah. One thing that I've noticed in the discussion of there being multiple intelligences is that there's two things that people are conflating. One is the idea of these different intelligences being distinct, and the other is them being uncorrelated.

And so people will often say, like, "Look, there are different intelligences," and they'll give evidence that these different cognitive skills are distinct from each other, they don't all correlate perfectly.

And then there's this slippery thing they do where they act like they've just shown that those intelligences are uncorrelated with each other. So they'll say things that imply that those things are uncorrelated, like, "You can't learn anything from an IQ score. The fact that there are these multiple components debunks IQ," or something like that.

But in order for that to be fair, you would have to show that there's no common factor, right?

Stuart Ritchie: Yeah, absolutely. And it's a common problem, actually in papers. I review a lot of papers where they ... Maybe it's a newer science paper where they've got brain scans, or a genetic paper, and they say, "Well, we've measured," say, "working memory. Or we've measured mental rotation. Or we've measured some specific cognitive ability."

But the question is: have you? If you've only given one test, it's very hard to say whether you're actually just measuring some general ability, or you're measuring something that's not necessarily the name of that test. Just because we call a test 'Mental Rotation' doesn't mean that it doesn't also include some sort of visual perception aspects, some reasoning aspects, and so on.

We haven't gotten tests that can really pinpoint specific cognitive processes at this point. Or I suspect ever, because there's gonna be lots of different cognitive processes that are involved in performance on any individual cognitive test.

And so it's very mistaken of these papers to say, "Well, we've found the brain areas that are responsible for mental rotation." Well, okay, but you have to check that those brain areas aren't just involved in all cognitive tasks as well.

So the fact of the correlation between cognitive tasks confuses even scientists. It's a common thing that I see in reviews.

Julia Galef: Interesting. I'm gonna shift tracks now to a different kind of objection, which is: How do we know that what the test is measuring is cognitive ability, as opposed to other things, like motivation to do well on a test, or familiarity with the kinds of questions that you get asked on an IQ test, like abstract questions?

Stuart Ritchie: I think there's different ways of looking at that. I think, first of all, there is gonna be an aspect of motivation, and there is gonna be an aspect of familiarity, with any test. A good tester will make sure that the person that they're testing is not looking out the window and completely uninterested in what they're doing and we'll try to focus peoples' attention on the test.

Much of the data from IQ tests comes from high stakes testing, so you'd expect peoples' motivation to be fairly high for that. But if you ask people to fill in questionnaires about their level of motivation, and also stuff like self control and things like that, concentration levels, and parcel that out of the IQ tests, when you then get their eventual score, some of the variation is likely due to motivation. It makes perfect sense.

And also, when you give people the same people the same test again and again and again and again and again, the tests show practice effects. So, people get better when they practice the specific kind of ... the specific test and also the specific kind of test. You can raise your IQ by four or five points, I think the most recent meta analysis shows, if you just do some tutorials on that particular kind of test.

So anyone who does IQ testing is well aware that these are potential objections and will try in predictive tests to try and account for that by measuring things like motivation and self control and so on, and try to partial that out of the analysis.

It's definitely something ... It's actually similar to my previous answer, where it's very hard to pinpoint one particular cognitive process with a cognitive test. Well, it's actually hard to pinpoint cognitive processes, full stop, with a cognitive test because you're also picking up on all this other stuff.

So the ways around that are just: Measuring the motivation partialing out, as I say. Making sure that the testing environment is not full of distractions and so on. But also, testing people on multiple different tests, and taking this general factor from them that has just has to do with all of the tests, and removes any error that might be to do with one test...

Another way is testing them multiple times as well in multiple different scenarios. And this increases the reliability of the measures.

In all psychology research we have measurement issues. And we know that humans are complicated and messy, and when you try to collect data on them, there's lots of different reasons for the way that their performance might be on any test. And so, making sure that you're increasing the reliability of your tests by using well-known standardized tests, rather than just ones you made up.

Sadly, lots of psychologists use tests that they just made up in their experiments, and then we don't know the reliability of. And by taking latent variables and taking variables across time, you can, to some extent, get around these problems.

But yeah, these are totally ... I completely accept those objections. They don't mean that there isn't a core of cognitive ability that you are actually measuring, but when people say things like, "Well, if you get better IQ tests, then that means that the IQ test is not measuring intelligence." Sure, but some people are gonna be able to practice the tests more efficiently, and that might be something that we might want to call intelligence.

Julia Galef: Yeah, I think a useful thing to do with all of these conceptual objections to IQ tests is to try to apply those same objections to other things besides intelligence.

Like physical fitness, for example, has a bunch of different components that are correlated but distinct, you know? Like running speed or upper body strength, or things like that. That doesn't necessarily mean that it's meaningless to talk about physical fitness as a general thing.

And then also, in this case, with practice you can get better at the long jump, or you motivation ... or energy levels might vary from day to day. That doesn't mean that there isn't a core underlying thing that it's meaningful to call “physical fitness,” that we could measure across multiple days, et cetera.

Stuart Ritchie: Absolutely. And with any human data, as I say, there's always gonna be noise, and some of that noise is gonna be due to other factors.

But testing large samples of people on large samples of cognitive tests is really the best way out of that problem. Intelligence testing has really been pretty impressive in recent years — increasing the sample sizes of the tests, and increasing the quality of the research. [There’s a lot] still to be desired, as there is in all psychology, I think.

But yeah, people are well aware of these objections. If you look through some discussion sections of papers you'll see limitation sections for people who are like, "Well, you know, the people might not have been that motivated because it wasn't a high stakes test," and we need to take that into account and so on. But it doesn't mean that there isn't a core there of actual cognitive ability.

Julia Galef: We were just talking about changes, things that could cause variation in individuals' IQ score, test to test, or over time. But then there's also variation over time in society's average IQ score. So the Flynn Effect, which I'm sure many listeners have heard of, is that IQ scores seem to have gone up ... I think it's like 30 points, over the multiple generations since we've been measuring it. Feel free to correct that number.

But do you think that that undermines ... Do you think that that undermines, not the idea of IQ as thing we're measuring that's meaningful, but does it undermine the “IQ is an innate, I guess, genetic thing”?

Stuart Ritchie: Well it's often used as an argument against IQ being related to genetics. It's something that we haven't talked about, but there is a lot of evidence that variation in IQ scores is explained to some quite substantial extent by variation in genetics. We can talk about that in detail if you'd like.

But yeah, the Flynn Effect is often held up as a, 'Well, here's the environment influencing intelligence, and therefore it can't be anything real or anything biological.'

Yeah, it's gone up 3 IQ points per decade across the 20th century. And there's some evidence that the Flynn Effect is kind of leveling off in some rich, industrialized Western countries. Not so much in the developing world, where it seems to be continuing at pace, which is fascinating to see.

That doesn't really say anything about the actual ability of IQ tests to tell us something about peoples' cognitive abilities or peoples' success in daily life, because it's a mean level effect. There's still variation. And the variation in IQ, if you imagine the sort of ... the normal curve ... the Bell Curve of intelligence. The Flynn Effect is simply that curve shifting along to the right as the generations go on.

There are many explanations that have been put forward for this: better nutrition, more effective schooling, better healthcare, and so on. There's lots of different reasons why the Flynn Effect might be occurring and both of the societal and kind of biological or nutritional level.

That doesn't say anything about the actual size of the variation around that mean. So that mean is increasing, but there's still variation, and that variation still tells you that people who are at the high end of that normal curve of intelligence are, on average, are gonna be the ones that do better at school, that better at the jobs, that live longer ... You know, there's evidence of IQ being linked to longer life span.

And so it's fairly irrelevant where the mean is around that. And we always knew, from twin studies and so on, that there's a large environmental portion to the variation in IQ — environmental being just, stuff that isn't genes in this case.

But it still is the case that if you measure IQ at any point along that Flynn cycle ... If you have twin studies of IQ from the 1940s and twin studies of IQ from now, you get fairly similar results, which is that monozygotic twins are much more strongly correlated with each other than the dizygotic twins are. And from that you can derive the fact that IQ is heritable. And that mean difference ... The Flynn Effect doesn't really [have anything to] say about that in any way.

Again, just like you did there with the physical fitness thing, think about other traits that have increased. Your height has increased across the 20th century. People got taller and taller. Probably for similar reasons. You know, the nutritional reasons, for instance, that might have stopped the kids being stunted and so on, and generally increased height across even healthy people. It doesn't say anything about whether height is heritable, to make the observation that height has increased across the 20th century.

Julia Galef: Yeah, I guess it suggests that ... I mean that maybe the missing piece, or the piece that a lot of people might be missing, is that a trait can be heritable, but its expression in a particular individual can still be influenced by the environment. So you could think about the "potential for height" as the heritable thing, and then, whether you reach your full potential can be influenced by, like, whether you get adequate nutrition as a child.

Stuart Ritchie: Yep. It's the most common misconception when talking about heritability, that people think heritability is the same thing as set in stone, immutable, never gonna be able to change it.

And I think that's where a huge amount of the anxiety comes from when people talk about heritability. Not just of intelligence, but the other areas that are controversial about heritability. Psychiatric disorders and so on. They've been extremely controversial when people have talked about their heritability. The heritability of schizophrenia for instance ... It's something which has debated hugely, although the evidence is very strong, that it's highly heritable. That doesn't mean that you can't change it.

So thinking back to height, height is strongly heritable, up to 80% heritable. Eighty-percent of the differences in height between people are due to differences in their genes. And yet, you can cause someone to be stunted if you malnourish them as a kid, and that's completely environmental effect on their height, which we know is heritable.

With the other classic example that's given — it's slightly different because it involves technology — is short-sightedness. We know that myopia is really strongly inheritable. 80 to 90% heritable. And yet, it can be cured instantly by one environmental input, which is wearing a pair of glasses. It's immediately gone effectively. It's no longer part, no matter what your genetic propensity for myopia is. We can sort it out.

Now that doesn't mean that we have these abilities to suddenly boost somebody's IQ. We don't have those yet. That's not in our technological abilities right now, but it's not theoretically ... The possibility of those technologies is not at all spoken against by the fact that IQ is heritable.

Julia Galef: There's a book called "Intelligence and How to Get It" by Richard Nisbett. I haven't read it, I'll admit, but the summary is: he is arguing that we don't have good evidence that intelligence is significantly or mostly hereditary. Do you understand why you disagree?

Stuart Ritchie: Yeah. There are objections to twin studies. Up until very recently, the evidence that we had that IQ was heritable was twin and family studies. As I said, the correlations of IQ for the monozygotic twins were compared to the correlations for the dizygotic twins, that is, identical twins and fraternal twins. The identical twins always had much higher correlations, and from that, you can work out that a portion of the variation in IQ is due to genes.

There are objections to that twin methodology. There are problems, for instance, with — maybe it's the case that identical twins are actually treated differently from non-identical twins. Maybe people dress their identical twins similarly. I've seen that objection. Maybe that has some aspect of an effect on how they see themselves, and that might affect their IQ.

I've never quite got the logic of that. I don't know why dressing people similarly makes their IQ more similar. But that's an objection that's often made to twin studies. People have written at length on the weaknesses of twin studies, misclassified zygosity where you think you've got dizygotic twins and they're actually monozygotic.

There are a whole range of debates, and objections, and back and forth in this sphere. But I think the consensus in the world of genetics is that whereas there are some reasons that the heritability of traits — not just IQ, but you can use twin studies to examine any trait, whether it's a disease like schizophrenia, say, or a trait like height, or a trait like IQ — there's some reason to believe that heritability might be slightly inflated in twin studies.

It might be a bit higher than can actually be associated with genes, but the constellation of evidence that we've got, not just from twin studies but now these days from DNA as well, from direct testing of DNA, shows that the twin studies were really on the right track all along.

Nowadays, we have evidence where we can take people who are completely unrelated, or as unrelated as any randomly selected humans are, and give them a DNA array that checks the variation at maybe 400,000 or 600,000 points in the DNA, and give them an IQ test. And essentially say, "Are the people who are more similar in their DNA more similar in their intelligence?"

That gives a positive heritability number as well. So you get a good chunk of the variation is associated with the genetic differences.

We're actually finding the specific genes that are related to variation in IQ, and those can be put together to produce a heritability number as well. You have twin studies, family studies, the study of twins adopted away into different environments, and so on.

The whole, as I say, constellation of evidence hangs together really well, that there really are genetic influences on variation in intelligence.

As I previously said, that doesn't say anything about whether we can change intelligence, or whether it's immutable, or whatever. That's just irrelevant to that. But it does mean that when we're looking for the reasons that people are similar right now or different from each other right now, some of that has to do with their genetic makeup.

We're making progress, as I say, in identifying the specific points in the DNA. There are papers coming out every few weeks now with large studies, we're talking hundreds of thousands of people, that are illustrating the points in the DNA where there's variation which is related to differences in cognitive test scores.

And also other traits that are related, like education. It's really easy to measure education. You just ask people how many years of schooling they had, and then you can get ... The next study that's about to come out has over a million people in it, with the DNA test and information on education. It's finding tons of points in the DNA that are related to better educational outcomes. A lot of those overlap with IQ, as well.

This is a really exciting and fascinating field now. To make arguments that actually there's no genetic basis to IQ is really out of date. That was an argument that was had, and I have to say won by the “IQ is heritable” side, in the 1970s.

Julia Galef: Is there any way for us to tell now what percentage of IQ, or related things like educational attainment, will eventually be explainable — like, we'll be able to find the genes for?

Stuart Ritchie: Yeah. We can take the twin study estimate as the highest point. The ceiling on how much variation in IQ is explained by genetics has to come from something like twin studies, because twin studies measure all possible genetic influences. They're fully formed people. We're not trying to take a DNA test and actually measure all the different aspects of their DNA. We're just saying, "It's a matter of relatedness, and how genetically related you are to someone or not."

So actually, that's the upper limit. That's the upper boundary. Say for IQ on average, across studies, that's about 50%. That means that we won't get higher than that in looking at the DNA studies, in terms of heritability.

The heritability from the DNA studies, at the moment — if you just look at common variation... Say a quarter of people have one letter in the DNA where everyone else has the opposite letter. If you just look at that kind of variation, then you get up to 20, maybe 30% of the variation in IQ is explainable by that kind of variation.

If you look at rarer genetic variants, you're getting up much closer to that twin ceiling. There has been big progress made on this, and we're getting towards the point where we can certainly measure using DNA most to all of the genetic differences.

Now, finding specifically what those variants are is the next question. We know even less of that. And then finding out the actual biological processes, the actual functions of all those genes, and how they relate to other genes, and how they relate to the environment, and so on, is the massive landscape of work that genetics has to do in the next few years.

I think it's been a pretty exciting thing over the past couple years. Even since my book came out, where I just made the prediction that we would, in the next few years, find some replicated variants that are associated with IQ scores. That prediction was correct. But there's been so much more besides.

Julia Galef: I have a question about your Twitter account again. There was an exchange on Twitter in which someone said, "Here's a tip for young people interviewing for jobs. Just take an IQ test and put that on your resume. If they still care about your degree or other things, then ignore them. They're not worth your time."

Setting aside the strategic wisdom of that advice or lack thereof, you replied that you would not be interested in someone's IQ score if they applied to a job for you. Why is that, if IQ is so predictive of all these important things?

Stuart Ritchie: The first thing to say is that at the group level, we have really strong evidence from big epidemiological studies that IQ is predictive of all sorts of good stuff, from education to occupation. And as I said, people with higher IQs tend to live longer for various reasons, and all that. Having a higher IQ is going to be better in that respect.

When I saw this person's CV when they were applying for a position here and I was asked to comment on some of the CV aspects, when they include their IQ, there's lots of things that come to mind.

First of all, it's just socially inept to say near the top of your CV, "My IQ is ..." I can't even remember what the number was. It was relatively high. I can't remember exactly what the number was.

Setting aside the fact that there was no validation of that, it was just a number on a CV. If maybe there was some transcript attached to it from some independent body that had verified this person's IQ, you would maybe take it a bit more seriously.

But I think it goes back to this question of, is it all that matters? We know that people, that individuals who can have a super-high IQ can do really badly in lots of areas, and people who are maybe lower can be totally competent in all sorts of stuff, too.

But the discussion on Twitter did make me reconsider that. There were some good arguments.

Clearly people who have not had all the educational opportunities that they might have wanted, and they maybe didn't have all the right contacts,

and they maybe didn't have the right location in the world to get the education in the world that they wanted or that would have been appropriate for them, maybe taking an IQ test is a kind of way into showing that they have lots of potential.

That is, in fact, the argument that people like Godfrey Thomson and the other early IQ folks gave at the start of the 20th century for changing the UK's education system. From one where essentially, if your parents knew people or if your parents were rich, they got you into a good school — and if not, you were in a really broken-down old building with poor teachers, and all harsh discipline, and all that stuff. IQ testing allowed people who were from poor backgrounds to get into the grammar schools, so the high-quality schools.

Now, that system is no longer in place in the UK, mainly because the schools that people were sent to who didn't pass the IQ test, essentially, were so poor, and it caused a huge amount of resentment, and so on.

Yet that kind of idea, that IQ tests can be a way of being more socially equitable, is a fairly strong argument. And that actually made me reconsider, why did I react like that to someone who was putting their IQ on their CV? I think it just comes back to the social ineptness. I think it stands out in a bad way that someone is showing off about their IQ, in a sort of Donald Trump. He often tweets about his high IQ and so on.

I think it's more a matter of, for the individual, having IQ tests in your dataset, if you're a scientist like me, is fascinating. It's great to see on average how people do, but bragging as an individual about your really high IQ is ... Yeah.

Julia Galef: Maybe we could solve this problem by allowing people to include their IQ score on their resume and then also include their score on a test of social skills, so that they can avoid sending a negative signal of bragging about their scores.

Stuart Ritchie: Sure, yeah. No, I think in a world where everyone had their IQ on their CV, it would be totally fine-

Julia Galef: And informative.

Stuart Ritchie: ... and if it was verified in some way, it would be fine.

But it's the fact that this person stood out by doing that. I think fundamentally, the thing that you're actually interested in is how people have used their IQ. For instance, if you're applying for a job in science, then what you want to see is lots of really cool papers, and lots of good courses that they've taught on, and lots of grants that they've applied for and got with those cool ideas. You read the papers, and they're full of beautiful writing and interesting ideas, and so on. That's what you want, rather than someone just saying, "My IQ's 145. Please hire me."

Julia Galef: Stuart, before I let you go, I wanted to ask you to recommend a book, or blog, or article, or something like that that you have substantial disagreements with but you still respect, slash like, slash think is worth engaging with for various reasons. Do you have anything like that you could recommend to our listeners?

Stuart Ritchie: The thing that immediately comes to mind is The Mismeasure of Man, which is the Stephen Jay Gould book from 1981, updated in I think '95-ish. Everyone's going to think, "Weird. This guy is an IQ proponent. Why is he recommending this —"

Julia Galef: Yes, that is what I thought.

Stuart Ritchie: But I had cause to read it again recently, and it's beautiful. It's beautifully written, as all Stephen Jay Gould books were.

Actually, the big one about the structure of evolutionary theory is pretty dull. But other than that, they're all pretty exciting. It includes loads of really important critiques of intelligence tests.

Now, a lot of them are not accurate. You go and look at the actual papers, and the historical record, and so on, and you find that they are not the case. It's been criticized for not fully accurately representing the data in the case of the skull sizes that he talks about, the skull size comparisons that were made by Samuel Morton at the end of the 19th century.

There's criticisms there, but there's a whole lot of stuff in that book that is really important to know.

And it has this rubric of, I think he calls it, learning by debunking. Which is: you can learn a lot about a subject by looking at how it has gone wrong, or it has been misconceived. In the same way that — I think Gould actually uses this example — you can learn a huge amount about evolution from watching a debate with creationists. Creationists are saying all these things, and then biologists come in and say why they're wrong. I find that really fascinating and an engaging way to learn new stuff.

Ironically, you can learn a lot about IQ tests from learning what Gould himself got wrong. I think there's a kind of meta-debunking going on there. You can read his book, which is about debunking, and learn how the evidence is actually fairly strong in many cases for IQ testing, but there are also some good arguments in there.

As I say, it's an entertaining and fascinating read, anyway. I think anyone who is truly interested in this subject should read that, but in the knowledge that any science book that was written in 1995 or originally in 1981 has to be out of date. There's a huge amount of good evidence that didn't exist back then. But also even then, there were mistakes and critiques of it. I think to fully appreciate the debate around this topic, that's one that you've got to read.

Julia Galef: Great. Well, thank you for that, and thank you so much for being on the show. I'll post a link to your excellent book that's my favorite introduction to the field of IQ sciences-

Stuart Ritchie: Thank you.

Julia Galef: ... “Intelligence: All That Matters.” Yeah, thanks so much for being on the show, Stuart.

Stuart Ritchie: It's my pleasure. Thanks.

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