

IDEAS AND DATA

HOME

Nassim Taleb on IQ

JANUARY 8, 2019

Nassim Taleb has published [an attack intelligence research](#) that is getting a lot of attention and so I thought I would respond to it.

As summarized in this useful chart from [Strenze \(2015\)](#), meta-analyses of hundreds of studies have demonstrated that IQ is predictive of life success across many domains.

Advertisements

Table 25.1 Relationship between intelligence and measures of success (Results from meta-analyses)

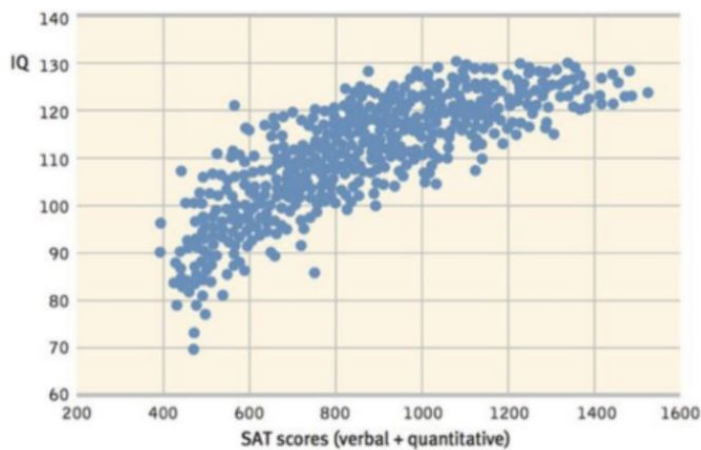
Measure of success	r	k	N	Source
Academic performance in primary education	.58	4	1,791	Poropat (2009)
Educational attainment	.56	59	84,828	Strenze (2007)
Job performance (supervisory rating)	.53	425	32,124	Hunter and Hunter (1984)
Occupational attainment	.43	45	72,290	Strenze (2007)
Job performance (work sample)	.38	36	16,480	Roth et al. (2005)
Skill acquisition in work training	.38	17	6,713	Colquitt et al. (2000)
Degree attainment speed in graduate school	.35	5	1,700	Kuncel et al. (2004)
Group leadership success (group productivity)	.33	14		Judge et al. (2004)
Promotions at work	.28	9	21,290	Schmitt et al. (1984)
Interview success (interviewer rating of applicant)	.27	40	11,317	Berry et al. (2007)
Reading performance among problem children	.26	8	944	Nelson et al. (2003)
Becoming a leader in group	.25	65		Judge et al. (2004)
Academic performance in secondary education	.24	17	12,606	Poropat (2009)
Academic performance in tertiary education	.23	26	17,588	Poropat (2009)
Income	.20	31	58,758	Strenze (2007)
Having anorexia nervosa	.20	16	484	Lopez et al. (2010)
Research productivity in graduate school	.19	4	314	Kuncel et al. (2004)
Participation in group activities	.18	36		Mann (1959)
Group leadership success (group member rating)	.17	64		Judge et al. (2004)
Creativity	.17	447		Kim (2005)
Popularity among group members	.10	38		Mann (1959)
Happiness	.05	19	2,546	DeNeve & Cooper (1998)
Procrastination (needless delay of action)	.03	14	2,151	Steel (2007)
Changing jobs	.01	7	6,062	Griffeth et al. (2000)
Physical attractiveness	-.04	31	3,497	Feingold (1992)
Recidivism (repeated criminal behavior)	-.07	32	21,369	Gendreau et al. (1996)
Number of children	-.11	3		Lynn (1996)
Traffic accident involvement	-.12	10	1,020	Arthur et al. (1991)
Conformity to persuasion	-.12	7		Rhodes and Wood (1992)
Communication anxiety	-.13	8	2,548	Bourhis and Allen (1992)
Having schizophrenia	-.26	18		Woodberry et al. (2008)

r correlation between intelligence and the measure of success, k number of studies included in the meta-analysis, N number of individuals included in the meta-analysis

This is the basic validating fact when it comes to IQ: the use of IQ tests can help us predict things we want to predict and to explain things we want to explain.

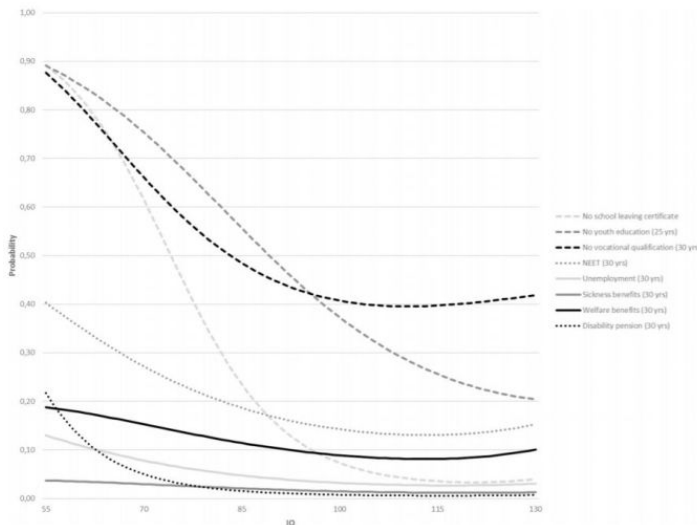
Does IQ Linearly Predict Success?

Some people wonder if IQ's relationship with success weakens above a certain threshold such that it is better described by a curvilinear trend rather than a simple linear one. Taleb bring this up and displays this graph:

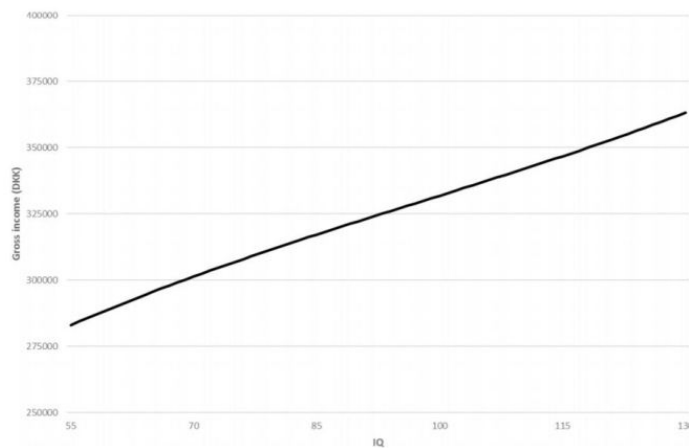


This graph does show a decrement in IQ's predictive validity as we move up the IQ scale. But there is still a positive correlation between SAT scores and IQ among those with IQs over 100. Just compare the distribution of scores among those with IQs of 110 and 130.

We can find other examples of this. For instance, [Hegelund et al. \(2018\)](#) analyzed data on over a million Danish men and various life outcomes. For several outcomes, IQ made little difference among those with IQs over 115.



However, for income the relationship was entirely linear.



We see the same thing in America if we look at the relationship between [IQ and traffic incidents](#):

MOTOR VEHICLE FATALITIES
(FROM STUDY OF AUSTRALIAN VETERANS UNDER 40)

IQ	Death rate per 10,000
>115	51.3
100–115	51.5
85–99	92.2
80–84	146.7

So this happens sometimes, but other times it doesn't. Importantly, these situations do not arise with equal frequency. [Coward and Sackett \(1990\)](#) analyzed data from 174 studies on the relationship between IQ and job performance. A

non-linear trend fit the relation better than a purely linear one only between 5 and 6 percent of the time, roughly what one would expect on the basis of chance alone. Similarly, [Arneson et al. \(2011\)](#) analyzed four large data sets on the relationship between IQ and education or military training outcomes and found in all four cases that the relationship was best described with a linear model. Thus, IQs relationship with occupational and educational outcomes is normally adequately described with a linear function.

I'll say more about this below, but here note in passing that Taleb never explains why a non-linear trend would invalidate IQ in the first place.

IQ and Job Performance

Often times, IQ tests are used by employers in their hiring process because IQ scores are a good predictor of job performance. Taleb doesn't see the point in this and writes that "If you want to detect how someone fares at a task, say loan sharking, tennis playing, or random matrix theory, make him/her do that task; we don't need theoretical exams for a real world function by probability-challenged psychologists."

This argument has a lot of intuitive appeal and is probably convincing to people who aren't familiar with this field of research. Within the field, however, it has long been known not only that IQ adds to an employer's predictive ability even if they've also administered a work sample test but that, in fact, IQ is sometimes a better predictor of job performance than work sample tests are.

Predictor	Correlation with job performance
IQ	0.51
Employment interviews (structured)	0.51
Peer rating	0.49
Job knowledge tests	0.48
Job tryout procedure	0.44
Integrity tests	0.41
Employment interviews (unstructured)	0.38
Biographical data measures	0.35
Work Sample Tests	0.33
Conscientiousness tests	0.31
Reference checks	0.26
Job experiences (years)	0.18
Years of education	0.1
Interests	0.1
Age	-0.01

([Roth et al., 2005](#); [Schmidt and Hunter, 1998](#))

Given this, Taleb's argument against using IQ tests in hiring is not compelling.

On Normality

Taleb also writes the following: “If IQ is Gaussian by construction and if real world performance were, net, fat tailed (it is), then either the covariance between IQ and performance doesn’t exist or it is uninformatinal.”

Taleb is correct to say that the distribution of many real world measures depart significantly from normality, that IQ scores are normally distributed by design, and that departures from normality can cause problems in statistical analysis. However, his conclusion from these facts, that IQ research is essentially meaningless, seems totally unwarranted.

Firstly, not all distributions are non-normal. Secondly, not all departures from normality are large enough to cause serious problems for standard statistical models. Thirdly, when departures from normality are large researchers typically do things like running variables through log transformations to achieve acceptable levels of normality, or run a different sort of analysis that doesn’t depend on a normal distribution. For Taleb’s criticism to be compelling, he would need to cite specific studies in which normality was departed from in a way which renders the actual statistical analysis done invalid and show that the removal of such studies from the IQ literature changes an important conclusion of said literature. He does nothing of the sort.

Moreover, Taleb’s conclusion, that the results of IQ research are meaningless, is clearly wrong. If such results were totally “uninformatinal”, they wouldn’t follow a sensible pattern. Yet, IQ correlates with job performance, and correlates better within jobs where IQ would be expected to matter more, and these correlations are consistent across studies. IQ correlates more strongly among identical twins than fraternal twins. IQ predicts performance in education. Etc. The probability of this theoretically expected pattern of relationships emerging if the analyses were so flawed that they were utter nonsense is extremely small, and so we are warranted in thinking that Taleb’s conclusion is false.

Taleb’s Measurement Standards

A consistent theme in Taleb’s article is that IQ tests don’t meet his standards for measurement. However, his standards for measurement are not standard in psychometrics, not justified by Taleb, and intuitively implausible.

Taleb writes that IQ is “not even technically a measure—it explains at best between 13% and 50% of the performance in some tasks (those tasks that are similar to the test itself), minus the data massaging and statistical cherrypicking by psychologists; it doesn’t satisfy the monotonicity and transitivity required to have a measure. No measure that fails 60–95% of the time should be part of “science””.

Let’s break this down. First, Taleb says that a measurement must explain more than 50% of the variance in tasks it is used to predict. That is, if we have a measure the use of which reduces our degree of predictive error by 50%, said measure is invalid according to Taleb. Taleb gives no argument justifying this standard. I’m going to give two arguments to reject it.

First, reducing our error by such a degree could be very useful. Actually, its hard to think of any situation in which a 50% reduction in error wouldn’t be useful.

Secondly, if real world behavior is complex in the sense that it is caused by many variables of small to moderate effect then it will be impossible to create measures of single variables which explain more than 50% of the variance in behavior. In the social sciences, single variables normally explain less than 5% the variance in important outcomes, suggesting that human behavior is, in this sense, complex. Given this, Taleb’s standard would be totally inappropriate for the behavioral sciences.

A related aspect of Taleb’s standards is that a measure not fail 60% of more of the time. Unfortunately, Taleb doesn’t define what “fail” means and it isn’t obvious what it would mean in the case of IQ research. It’s equally unclear where he got this number from.

However, even without knowing any of this it seems clear that Taleb’s standard is problematic. Consider a case in which your probability of correctly solving a problem is 1% without a given measure and 40% with said measure. This measure thus increases your probable of success by a factor of 40 and would be extremely useful. Yet, it has a fail rate of 60% and so, according to Taleb, can’t be used in science. This seems clearly irrational and so rejecting Taleb’s standard seems justified.

Finally, let’s consider Taleb’s standard of montonicity. This is getting back to the idea that IQ’s relationship with an outcome, say job performance, needs to be the same at all levels of job performance. As I’ve already reviewed, IQ’s relationship with important outcomes is largely linear. But this standard seems

unwarranted to begin with. IQ is useful in so far as it let's you make predictions. If IQ has a non-linear relation with some outcome, one merely needs to know that and IQ will still be able to help us make useful predictions.

In fact, IQ can help us make predictions even if its relation with an outcome is nonlinear and we think its linear. For instance, if IQ's relationship with some outcome becomes non-existent after an IQ of 120, it will still be predictive in the vast majority of cases and so our predictive accuracy will probably be greater than if we hadn't used IQ at all.

Against Taleb's standards for measurement, I prefer a practical standard. Firms and colleges are trying to predict success in their respective institutions and social scientists are trying to explain differences in interesting life outcomes. IQ tests help us do these things. Even with IQ tests, prediction is far from perfect. But it is better than it would be without them and that fact more than any other legitimizes their use.

Are High IQ People Pencil-Pushing Conformists?

Taleb also attributes various negative attributes to people who score highly on IQ tests. He says that people who score highly on IQ tests are paper shuffling obedient "intellectuals yet idiots" who are uncomfortable with uncertainty or not answering questions. Such people also lack critical thinking skills. In fact Taleb goes as far as saying that IQ "measures best the ability to be a good slave." and that people with high IQs are "losers".

Taleb treatment of this issue is entirely theoretical. He cites no empirical evidence nor does he make reference to empirical constructs by which his claims might be tested. However, it seems reasonable to suppose that, if Taleb is right, we should see a positive correlation between IQ and measures of conformity and risk aversion, and a negative correlation between IQ and leadership as well as critical thinking. But this is the opposite of what the relevant literature suggests.

First, consider conformity. [Rhodes and Wood \(1992\)](#) conducted a meta-analysis and found that people scoring high on IQ tests were less likely than average to be convinced by either conformity driven or persuasion driven rhetorical tactics. People who score high on intelligence tests are also more likely to be atheists and libertarians ([Zuckerman et al. 2013](#), [Carl 2014](#), [Caplan and Miller 2010](#)). These are minority viewpoints and not what we would expect if IQ correlated with conformity.

With respect to risk, [Andersson et al. \(2016\)](#) show the majority of research linking cognitive ability to risk preference either finds no relation between the two variables or a finds that high IQ individuals tend to be less risk averse than average.

TABLE 1. Studies reporting correlations between cognitive ability and risk aversion.

Results based on	Results	Significant negative	Significant positive	Non-significant
Incentivized tasks	20	7	2	11
Non-incentivized tasks	13	3	1	8
Risky behavior	9	4	3	2

Notes: Results included are from studies listed in Table E.1 (Incentivized tasks), Table E.2 (Non-incentivized tasks) and Table E.3 (Risky behavior) in Online Appendix E. Significance level required is p -value < 5%.

[Beauchamp et al. \(2017\)](#) found that intelligence is positively associated with people's propensity to take risk in a sample of 11,000 twins. This was true of risk seeking behavioral in general as well as risk seeking behavior specifically with reference to finances.

With respect to leadership, [Levine and Rubinstein \(2015\)](#) find that IQ is positively correlated with the probability of someone being an entrepreneur. In a meta-analysis of 151 previous samples, [Judge and Colbert \(2004\)](#) found a weak positive relationship between a person's IQ and their effectiveness as, or probability of becoming, a leader. This is hardly what we would expect if IQ measured a person's ability to "a slave".

With respect to critical thinking, IQ is strongly correlated with formal tests of rationality which gauge people's propensity to incorrectly use mental heuristics or think in bias ways ([Ritchie, 2017](#)).

And finally, with respect to real world problems as measured by situational judgement tests, [McDaniel et al. \(2004\)](#) found a .46 correlation between people's scores on SJTs and IQ tests in a meta-analysis of 79 previous correlations.

Thus, Taleb's assertions about the psychological correlates of IQ are entirely at odds with what the relevant data suggests.

Population Differences in IQ

Taleb also makes four remarks Taleb made about population differences in IQ.

First, he says "Another problem: when they say "black people are x standard deviations away". Different populations have different variances, even different skewness and these comparisons require richer models. These are severe, severe

mathematical flaws (a billion papers in psychometrics wouldn't count if you have such a flaw)"

It is true that Black and White Americans differ in their degree of variance in IQ. Specifically, the Black standard deviation is smaller than the White standard deviation. This has been known about, and written about, for decades. But this doesn't pose a problem for talking about the distance between groups in standard deviation units both because you can simply aggregate both groups into one and use a pooled standard deviation and because you can simply specify which standard deviation you are using.

Taleb's second remark is that "The argument that "some races are better at running" hence [some inference about the brain] is stale: mental capacity is much more dimensional and not defined in the same way running 100 m dash is."

I think the argument Taleb is imagining can be more charitably stated as follows: there are genetically driven differences between ethnic groups for many, indeed nearly all, variable physical traits outside the brain, so, unless we have specific reason to think otherwise, our default assumption should be that the same is true of the brain.

Put more precisely, we might say that the presence of genetically driven differences for most variable traits outside the brain increases the prior probability of genetically driven differences for variables traits within the brain. We might further explain that the distinction between brain and non-brain, while important to us, is not important to evolution, and that the same processes which cause non-brain differences can also cause brain differences. Thus, in the absence of other evidence, the prior probability of neurologically variable traits differing between ethnics groups due to genetics is high.

Whatever one may think of this argument, Taleb's response, that we define mental traits differently than physical traits, is impotent. After all, Taleb doesn't explicate why the difference in how we define physical and mental traits should be relevant to the logic of the argument. Nor, in fact, does he specify how said definitions differ at all. He merely asserts that some unspecified difference in definition exists and implies that this difference is relevant to the argument in an unspecified way. Obviously, this is not a compelling rebuttal.

Taleb's third remark is as follows: "If you looked at Northern Europe from Ancient Babylon/Ancient Med/Egypt, you would

have written the inhabitants off... Then look at what happened after 1600. Be careful when you discuss populations.”

Taleb is correct in the sense that the populations who are most developed today are always not the ones who were most developed in the ancient world. However, it is nonetheless true that we could have predicted which populations would end up being more economically developed if we had a more compelling model. Specifically, you can predict the majority of modern day variation in national economic development on the basis of ecological facts concerning, for instance, potential crop yield and animal domesticability, of a region in pre-historic times ([Spalore et al. 2012](#)).

The relationship between this fact and the idea that long run national development is influenced partially by genetically driven population differences is complicated since such ecological differences might directly cause differences in development, but might also cause differences behavior via impacting selective pressures, or may do both.

Thus, the relationship between ancient and current variation in national development poses no obvious problem for partially biological narratives.

Finally, Taleb remarks” The same people hold that IQ is heritable, that it determines success, that Asians have higher IQs than Caucasians, degrade Africans, then don't realize that China for about a Century had one order of magnitude lower GDP than the West.”

This comment suggests that Taleb simply hasn't read the authors who argue that IQ is an important driver of national differences in wealth. The most famous proponents of this hypothesis are, easily, Richard Lynn and Tatu Vanhanen. In their 2012 book [“Intelligence: a Unifying Construct for the Social Sciences”](#), they report that IQ can explain as much as 35% of national variation in wealth. They go on to posit several variables which might explain when nations strongly deviate from their expected wealth based on IQ, including, for instance, possessing large oil reserves and having a socialist economy.

Like individual differences, national differences are not caused by a single factor. Many variables are involved and IQ is only one of them. The fact that some variation in national wealth cannot be explained by IQ does nothing to diminish the proportion of variation in national wealth that can be explained by IQ.

Can We Believe Psychological Research?

Now, Taleb actually admits that what he said had no evidence behind it. He gives a reason for this, stating that: "I have here no psychological references for backup: simply, the field is bust. So far ~ 50% of the research does not replicate, and papers that do have weaker effect."

Presumably Taleb is referring to the Open Science Collaboration results from 2015. [OSC \(2015\)](#) replicated 100 psychological experiments and in only 47% of cases did the replications find the same thing as the original study. We might therefore think that the probability of some hypothesis being true is roughly 1 in 2 if it has been previously confirmed by a novel psychological study.

It's important to realize that this has nothing specifically to do with psychology. [Camerer et al. \(2016\)](#) replicated 18 experiments in economics and found that 61% of them replicated. In fact, both psychology and experimental economics have far higher replication rates than do several other fields. For instance, [Begeley and Ellis \(2012\)](#) found that cancer research replicated only 11% of the time. Even worse, an attempt to replicate 17 brain imagining studies [completely failed](#). That is, not a single finding replicated, suggesting that the replication rate in brain imagining research is, at most, 5.5%.

I am unaware of any attempts to directly measure the replication rates of most physical sciences, but [Nature](#) conducted a large survey of scientists and asked them to estimate the proportion of work in their fields that would replicate. I've average the results by field and as you can see, in no field do researchers expect work to replicate as much as 75% of the time.

Discipline	Estimated Replication Rate
Physics	0.73
Other	0.52
Medicine	0.55
Material Science	0.60
Engineering	0.55
Earth and Environmental Science	0.58
Chemistry	0.65
Biology	0.59
Astronomy	0.65

Now, Taleb doesn't tell us what replication rate he requires to care about what a science says. Still, one can easily imagine that his argument against caring about psychological data could also be used as an argument against caring about scientific data in general.

Regardless, let's suppose that the probability of a social scientific finding replicating is roughly 50% and the probability of a hard science finding replicating is roughly 60%. How should we react to this purported fact?

First, it's important to realize that the probability of some randomly formulated hypothesis about the world being true can be construed as being less than one half. This requires a certain way of looking at probability, but it doesn't seem unreasonable to say that there are lots of ways the world isn't and only one way the world is, so the vast majority of possible descriptions of the world are false. By contrast, replication research might be taken to suggest that something like half of hypotheses that have been confirmed by an initial study are true. Looked at this way, such rates actually represent significant epistemic progress.

More importantly, we can easily guess ahead of time which studies are going to replicate. Consider, for instance, what happens if we use a single metric, p values, to predict whether a study will replicate. That 2015 study on replication in psychology found a replicate rate of only 18% for findings with an initial p value between .04 and .05 and 63% for findings with an initial p value of less than .001. Similarly, that 2016 study on replication in economics found a replication rate of 88% for finding with an initial p value of less than .001.

Using these as similar clues, multiple papers have found that researchers are able to correctly predict which of a set of previous findings will successfully replicate the strong majority of the time ([Camerer et al., 2018](#); [Forsell et al., 2018](#)).

Thus, if we consume research intelligently, we can be a lot less worried about buying into false positive results.

Returning to psychology, and intelligence research in particular, it is important to note that a lack of statistical power is one important cause of low replication rates which does not apply to IQ research to the degree that it applies to most disciplines.

Specifically, while no field has the sort of statistical power we would theoretically like it to have, intelligence research comes a lot closer than most fields do.

Citation	Discipline	Mean Power
Button et al. (2013)	Neuroscience	21%
	Brain Imaging	8%
Smaldino and McElreath (2016)	Social and Behavioral Sciences	24%
Szucs and Ioannidis (2017)	Cognitive Neuroscience	14%
	Psychology	23%
	Medical	23%
Mallet et al (2017)	Breast Cancer	16%
	Glaucoma	11%
	Rheumatoid Arthritis	19%
	Alzheimer's	9%
	Epilepsy	24%
	MS	24%
	Parkinson's	27%
Nuijten et al (2018)	Intelligence	49%
	Intelligence - Group Differences	57%

Thus, intelligence research should replicate better than most research does. Given this, whatever our general level of skepticism about social science is, our skepticism about intelligence research should be lesser.

Of course, low power isn't the only reason that research fails to replicate, and the most important solution to this problem is to simply not rely on un-replicated research.

There are other concerns one might raise related to p hacking and publication bias, Taleb didn't mention these issues so I won't deal with them here, but these are all real problems. However, they all have at least partial answers, psychology is improving with respect to many of these problems with time (e.g. the rise of pre-registered research) and none of them warrant thinking that psychological research, when analyzed carefully, can't be epistemically useful.

Advertisements



Share this:



Loading...

Related

[Scientific American on IQ/Genes](#)

[Race, Wealth, and Economic Freedom in North and South America](#)
In "Demographic Change"

[The Alt Right and Nazism](#)
In "Race"



Uncategorized

[← Predicting Party ID by Race](#)

[Scientific American on IQ/Genes →](#)

4 thoughts on “Nassim Taleb on IQ”



johnnyjmcivorv

JANUARY 11, 2019 AT 7:11 AM

Hey Sean, where can I get in contact with you? I'm interested in doing some online polls for my own curiosity of the alt right, was curious if I could get your thoughts on it.

★ Like

REPLY



Chris

JANUARY 12, 2019 AT 6:58 AM

Thanks for this post. Despite, or perhaps because of, Taleb's poorly-constructed essay, many of his critics don't actually address what he said.

★ Like

REPLY

Pingback: [Nassim Taleb on IQ: How To Miss The Point – Dissenting Academia](#)



monerty 01

JANUARY 13, 2019 AT 7:57 PM

Not very related but are you a realist or antirealist/instrumentalist about science?

★ Like

REPLY

Leave a Reply

Enter your comment here...

- [January 2019](#)
- [August 2018](#)
- [April 2018](#)
- [March 2018](#)
- [February 2018](#)
- [November 2017](#)
- [August 2017](#)
- [July 2017](#)
- [May 2017](#)
- [April 2017](#)

- [Home](#)

Privacy & Cookies: This site uses cookies. By continuing to use this website, you agree to their use.
To find out more, including how to control cookies, see here: [Cookie Policy](#)

Close and accept