HBD at Lightspeed

by: Werkat

Version 1.0.0

Changes:
-New title and cover
-Ch.2 Substantially revised
-Miscellaneous changes e.g. typo corrections, new replications of already known findings listed next to the old ones, etc.
-[Link to a pdf file of the book] for phone users to be able to see formatting properly, and for generic users to not have to use google docs if they wish. Note that you won't see new content if you never check [the live version] for updates.
-All figures have links so that users can save the images as they please.
Planned Additions:
-Revision of Ch.7 with new evidence
-Chapter 1 will be about methods for statistical and causal inference
-More detailed overview of what molecular genetic methods miss, some improvements that have recently been made to GWAS-based heritability estimates, etc.
-Chapter 8 about sex differences and evopsych

Some notes to the reader:

1. If sci-hub links ever break, go to the [source list section] for instructions on how to gain access.
2. Go [here] to download everything in the project:

(archive of all sources, pdf of this document, instructions, etc)

3. This is a work in progress, expect new content in the future.
4. If you want to talk to me (doesn't have to be formal or whatever):
   a. wehrkatzer@gmail.com
   b. Discord: wehrkatz#0264

   Any suggestions, corrections, questions, etc are welcome. If you just want to call me a racist, that's fine too. If you want to do your own work and add it, do file-make a copy, make your suggested changes, and share the new document with me.

Table Of Contents

2. Academia

Navigation:

Previous Chapter | Table Of Contents | Next Chapter

Summary:

Degrees Are (Mostly) Zero-Sum:

Education pays, people with Bachelor's Degrees are paid 73% more than highschool graduates [1189, Table 3.1]. This, however, leaves open the question as to why this is the case. There are essentially three competing explanations which offer to partially complete the picture:

1. Explanation 1 (E1): Education increases peoples' productivity, and employers pay a premium for the extra productivity.
2. E2: Innately productive people want to, or are enabled to seek out more education than less innately productive people, and employers pay a premium for the innate productivity.
3. E3: Employers pay educated people more money than their productivity justifies.

When explanation 1 accounts for a paucity of the payment differences, education is a zero-sum game, and this unfortunately seems to be the case [more here]; all told, explanation 2 explains at least 18% of the story, and explanation 3 likely accounts for around 80% of the story. Proposed positive externalities do not seem satisfactory to make the system a net benefit in spite of its flaws [see 1189, Ch.6].

On Science:

The fundamental goal of science is to deduce sound theories empirically [more here]. To do this, it is essential to have sound operationalizations and sound statistics so that informative analyses can be done with tools which provide a clear view of the various aspects of reality.

Reasonable priors on the value of the education system should not inspire hope for academic competence [more here]. But priors aside, how do experts actually perform? Even experts who believe in convoluted theories should be able to predict reasonably well the things which they're knowledgeable about, but there isn't reason to believe that their training enables them to perform very well at this sort of task [more here]. More objectively and easily testable, and of paramount importance, is statistical literacy. Unfortunately, academics are often breathtakingly statistically illiterate in terms of tools that are widely used and easy to understand [more here].

What about the academic environment? Many imagine the peer review process as an objective one, but interrater reliability is quite low [more here] which allows for publication bias against certain results [more here] and authors [more here]. Academics' careers are dependent on publishing a large quantity of papers with results which are pleasing to publishers [more here and here]. More prestigious journals are objectively worse than smaller journals due to these incentives being felt more starkly [more here]. This distorts what general picture the research literature gives of what is true [more here]. Transparently bad papers are accepted through the filters at alarmingly high rates [more here]. Questionable research practices have led to alarmingly low probabilities that a given result can be replicated by another paper following instructed procedures [more here]. The system doesn't even seem to ensure that references are written correctly, that cited results are accurately represented, or even that transgressions as major as plagiarism are warded off [more here].

Beyond just its effects on the quality of society's researchers and research literature, the system has caused the literature to be even less accessible to the layman than the inherently esoteric nature of the scientific endeavour necessitates; it has done so in three ways [more here]:

1. It increases article quantity and length beyond what rigor necessitates.
2. Unnecessarily esoteric language is shoehorned into the literature to impress reviewers.
3. Tangible paywalls prevent free access despite authors being unpaid by journals.

If the journal system filter doesn't ensure quality, how are we to tell science from quackery? Well, it is only since the middle of the 20th century that our modern practices have spread widely and that external reviewers have been given such visibility within academic journals [1187 & 1188]. Perhaps experienced researchers can tell quackery for themselves without a middleman to tell them. After all, good papers are easily filtered; maximum expected replicability is achievable for anybody who consumes research intelligently by looking for good research practices such as the following: rigorous transparency in methods and data, pre-registration, high statistical power, and good study design [more here].

Degrees Are (Mostly) Zero-Sum:

Education pays. United States Census data shows that on average, people with Bachelor's Degrees are paid 73% more than highschool graduates [1189, Table 3.1]. However, this raw figure is merely correlational in nature. The crucial question, as always, is **why** this is the case. There are essentially three explanations which we may take as helping to explain the overall relationship:

4. Explanation 1 (E1): Education increases peoples' productivity, and employers pay a premium for the extra productivity.
5. E2: Innately productive people want to, or are enabled to seek out more education than less innately productive people, and employers pay a premium for the innate productivity.
6. E3: Employers pay educated people more money than their productivity justifies.

Note that E2 doesn't necessarily require innate, genetic, unchangeable qualities, but merely whatever exists prior to education which can explain the earnings differences.

When explanation 1 accounts for a paucity of the payment differences, education is a zero-sum game. Provided that the externalities aren't enough to make up for tuition and opportunity costs [see 1189, Ch.6], investing in education is like standing up in a football stadium: When one does it, they get a better view, but when everybody does it, their legs just get tired.

Explanation 2 accounts for at least 18% of the picture [more here]. As for explanation 3, there are a few lines of evidence we can take as assessments of its contribution:

1. Individual differences in educational attainment are greatly rewarded, but national differences are not [more here].
2. Educational returns do not come year by year, but are instead largely distributed around graduation years [more here].

Caplan's book [1189] also assesses a few extra softer lines of evidence:

3. Employers pay good money for degrees irrelevant to the occupation.
4. Irrelevant classes are rewarded as much as relevant ones are rewarded.
5. Forgetting the material is not financially punished by employers.
6. Students care about easily graduating with the most marketable diplomas, not about learning marketable skills.
7. Employers devalue diplomas as they learn about employee productivity.

Overall, the evidence seems to paint the picture of E3 being ~80% of the story. Taxpayers and kids are throwing their money and youth down the drain. Externalities do not make up for this [see 1189, Ch.6].

Personal Vs National Returns:

This sort of analysis is most directly analogous to the football stadium analogy. In the analogy, individual differences in standing should be related to individual differences in view quality while stadium differences in mean amount of standing should not correspond to mean differences in view quality. If the analogy holds true, then national education differences should not strongly correspond to national income differences even if individual income differences are related to educational attainment.

Obviously, the two variables are indeed strongly related on the individual level:

[1189 - Table 3.1]:

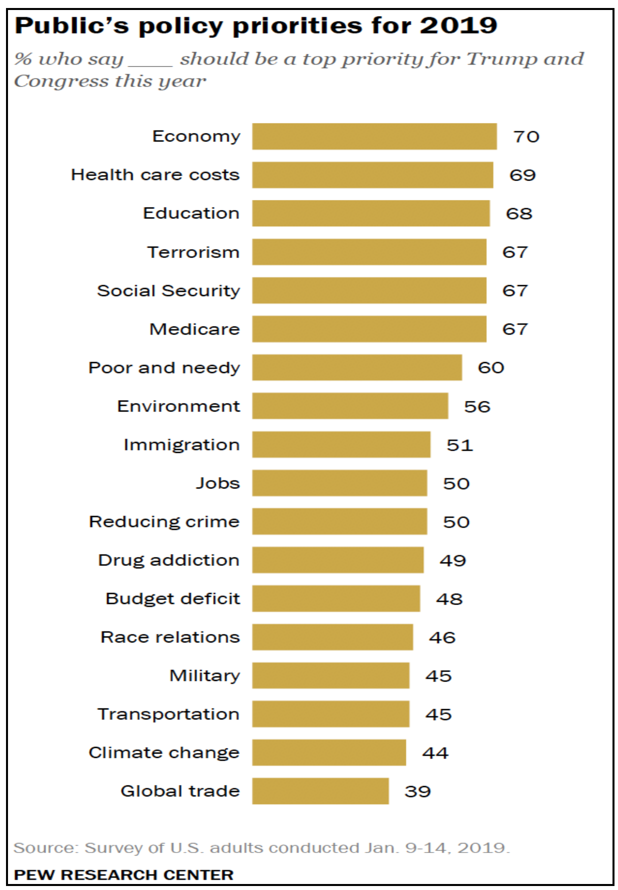
Table 3.1: Average Earnings by Educational Attainment (2011)

	Some High School	High School Graduate	Bachelor's Degree	Master's Degree
Average $ Earnings	31,201	40,634	70,459	90,265
Premium over H.S.	-23%	+0%	+73%	+122%

Source: United States Census Bureau 2012a.

But what about nationally? Correlationally, there is a large amount of heterogeneity in results, with effects ranging from slightly negative to modestly positive, giving us an overall effect size of national incomes being +1.3% higher per year of education the mean citizen receives [1189, Figure 4.3]. Already, this is much smaller than individual effects, with individuals, on average, making +10.9% more than somebody who has 1 less year of education [1189, Table 4.1].

As always however, causality is an issue. Shifting focus from individual level results to national level results only eliminates the influence of E3, not E2. Just as greater individual level income can plausibly enable more education spending, and just as individual level ability can enable graduation, these are also potential concerns on the national level. After all, the majority of the tab is picked up by the state [more here]. It could just be, for example, that increases in national tax revenue prompt increases in educational spending. After all, education is highly prioritized [1203]:



The classic bumper sticker muses that it will be a great day when our schools get all the money they need and the air force has to hold a bake sale to buy bombers. However, this great day arrived long ago; the air force may not hold bake sales, but military spending has long since been surpassed by educational spending [1189, p.200].

The best evidence on the question of national level causality comes from a natural experiment found in Russia [1204]. Recently, their standard degree program shifted in line with the rest of the world. The shift in average educational attainment did not correspond to a shift in average employability. The most educated individuals are still paid best, but the shift in average education did not correspond to a shift in average income. This approach is nice because of the straightforward interpretability which comes from its apples to apples comparison, and because the within-country approach sidesteps previous concerns of international comparability and result heterogeneity.

Graduation Years:

One straightforward thing we can do to assess the contributions of the three explanations is to break the data down into which years of education are the most rewarded. Doing this, the effect of individual years is more than cut in half, and they are dwarfed by the premiums paid for graduation years with over 60% of education premiums being accounted for by degree years rather than the raw count of school years people complete [1189]:

Source 1189 - Table 4.1:

Table 4.1: Effects Of Education On Earnings In The GSS:

Education:	If Only Year # Matters:	If Diplomas Matter Too:
Years Of Education:	+10.9%	+4.5%
High School Diploma:	—	+31.7%
Associate's Degree:	—	+16.6%
Bachelor's Degree:	—	+31.4%
Graduate Degree:	—	+18.2%

Notes: All results are corrected for age, age squared, race, and sex; are limited to labor force participants; and are converted from log dollars to percentages.

Presumably, if E1 were the predominant reason that education is valued, then compensation should linearly increase as people learn more skills. Instead, the fact that degrees are valued so much suggests that E3 is of paramount importance.

The most important objection to this sort of analysis is to bring up the role of E2. Such an analysis may be a misleading assessment of E3 if the causal influence of E2 is disproportionately concentrated upon diploma years and absent from raw school year count. Pathetic as the GSS' measures of cognitive abilities (such as wordsum) may be, they can be used to correct within-person returns somewhat downwards in order to assess the relative effects of such corrections. Such adjustments in the GSS affect all years of education equally, leaving relative premiums for degree years unaffected [1189 - Table 4.2].

Further research also reaffirms the same general finding on the relative influence of E2 on schooling premiums [1192, pp.48–50; 1193, table 3, column 2; 1194, table 4, OLS column 6; 1195, table 3; 1190, table 5; & 1191, p.606]. Note that E2 doesn't necessarily require innate, genetic, unchangeable qualities, but merely whatever exists prior to educational attainment which can be used to predict earnings.

Given the distribution of ability effects on educational returns, we should be able to take diploma effects as being signalling effects which are consistent with explanation 3. However, the role of diploma effects should be taken as an underestimate of the role of E3, as there are smaller employability spikes at course enrollment and completion [1205].

Finally, there is one more interesting pattern we can see in the raw returns data: Given a group with Bachelor's Degrees, those who took the longest to obtain them are those who earn the least [1206]. Positive correlations between non-degree school years and income is thus a dropout phenomenon. Presumably, E1 should predict that the people who take their time to learn as much as possible end up with the greatest quantity of marketable skills, and end up with the highest incomes. However, this seems to suggest that within this sort of context, E2 overpowers any such effects.

Overall, 80% is a reasonable figure for the importance of E3, and is broadly consistent with external lines of evidence [more here].

The Role Of Pre-Existing Abilities:

Given the previous discussion [more here], we can say with a good deal of confidence that the role of pre-existing abilities in explaining the education-income correlation is concentrated on the year to year 'returns' rather than the sudden spikes people get from diplomas. However, this sort of evidence doesn't tell us the actual degree to which the year to year differences in income are due to pre-existing earning ability because it is difficult to comprehensively account for every single pre-existing trait of relevance. Luckily, there is, available to us, the appealing approach of looking at identical twins.

Doing family controls will account for the degree to which family members are similar in every trait there is to measure, not just the things we've figured out how to measure. A recent meta-analysis of every twin study ever done [490], assessing 2,563,627 pairs of identical twins and 9,568 traits, finds identical twins to correlate with each other at about .636 for most traits. Given this, we can get a decent idea of just how much juice there is to squeeze out of pre-existing abilities if we assess the degree to which an identical twin who gains more education ends up wealthier than their cotwin.

Sources [1197, pp.1846-1852] & [1201, pp.219-222] review such studies, and estimates are that up to 50% of the raw education-income correlation could be accounted for with this approach. Unfortunately, noting various considerations for between-study differences, the author chooses a 10 to 15 percent figure as his preferred estimate for the role of pre-existing ability in the raw education-income correlation. Of course, such an approach only gives us an idea of the ballpark we're working with if there are pre-existing abilities to account for in which identical twins are not equal, but the paper explicitly endorses the assumption that identical twins are equal in abilities. This however, is demonstrably false. Going back to the meta-analysis [490], twins tend to correlate at about .636 for most traits. Given this, ~59.5504% of variance in traits in general is not explainable by identical cotwins; so at maximum, —the least charitable possible estimate which linearly projects twin trait effects onto the non-twin variance—, using the 15% figure, would result in an ability bias figure explaining ~37.08% of the raw correlation between education and income. Moreover, looking at IQ alone, identical twins are not completely equal in IQ, and the —identical twins who are higher in IQ than their cotwins prior to the emergence of educational attainment differences— end up with higher education attainment than their cotwins [1198, 1199, & 1200]; this leaves us with an ability bias about 15% higher than indicated by the raw twin results [1198].

The 10 to 15 percent figure is smaller than what we can get from the abilities we can actually measure; IQ alone, measured prior to school, is enough to explain about 18% of the raw education-income correlation [1202] (note: such an approach can account for the effects of between-twin IQ differences), and predicts educational attainment at ~.49 [253]:

Source 253 - Table 1:

Table 1
Predictors of socioeconomic success

	k	N	r	rw	p	S.D.r	S.D.p	CV 95%	CI 95%
Correlation with education									
Intelligence (all studies)	59	84,828	.46	.48	.56	.12	.10	.36/.75	.53/.58
Intelligence (best studies)[a]	20	26,504	.49	.48	.56	.10	.07	.42/.69	.52/.59
Father's education	72	156,360	.40	.42	.50	.14	.13	.25/.75	.47/.53
Mother's education	57	141,216	.37	.40	.48	.13	.13	.22/.73	.44/.51
Father's occupation	55	147,090	.34	.35	.42	.09	.07	.27/.56	.40/.44
Parental income	13	64,165	.29	.31	.39	.10	.11	.17/.61	.33/.46
SES index	17	69,082	.41	.44	.55	.12	.10	.35/.75	.50/.60
Academic performance	27	49,646	.48	.47	.53	.09	.07	.39/.68	.50/.56
Correlation with occupation									
Intelligence (all studies)	45	72,290	.37	.36	.43	.13	.08	.28/.57	.40/.45
Intelligence (best studies)[a]	21	43,304	.41	.38	.45	.09	.05	.35/.54	.42/.47
Father's education	52	132,591	.27	.26	.31	.08	.06	.19/.43	.29/.33
Mother's education	40	116,998	.24	.23	.27	.08	.07	.13/.41	.25/.30
Father's occupation	57	146,343	.28	.29	.35	.10	.08	.19/.51	.33/.37
Parental income	12	60,735	.19	.21	.27	.07	.10	.07/.46	.21/.32
SES index	16	74,925	.30	.31	.38	.08	.08	.22/.54	.34/.42
Academic performance	17	54,049	.33	.33	.37	.09	.07	.23/.51	.33/.41
Correlation with income									
Intelligence (all studies)	31	58,758	.21	.16	.20	.09	.11	−.01/.40	.16/.23
Intelligence (best studies)[a]	15	29,152	.22	.19	.23	.08	.06	.10/.35	.19/.26
Father's education	45	107,312	.16	.14	.17	.09	.08	.01/.32	.14/.19
Mother's education	37	93,616	.13	.11	.13	.10	.07	.00/.27	.11/.16
Father's occupation	31	98,812	.16	.15	.19	.08	.10	.00/.38	.15/.22
Parental income	17	395,562	.16	.16	.20	.06	.07	.06/.33	.16/.23
SES index	14	64,711	.15	.14	.18	.07	.08	.03/.33	.14/.22
Academic performance	14	41,937	.11	.08	.09	.07	.08	−.07/.24	.04/.13

Note. k — number of independent samples, N — number of individuals, r — average correlation, rw — sample size weighted average correlation, p — sample size weighted average correlation corrected for unreliability and dichotomization, S.D.r — standard deviation of r, S.D.p — corrected standard deviation of p, CV 95%–95% credibility intervals of p, CI 95%–95% confidence intervals of p, SES — socioeconomic status.
[a]Best studies are the ones where intelligence is tested before the age of 19, and socioeconomic success is measured after the age of 29.

Whatever we are to think about IQ, anything measured prior to school is indicative of at least **something** which existed first, and so cannot be due to later schooling.

All told, flooring the role of pre-existing ability at such an 18% figure seems like it should be a generous underestimation given that there should be traits of relevance other than IQ. What specifically these traits are though, is not immediately clear. A thin body of research investigates various non-cognitive abilities like personality; such things do confound the returns to education [1189, p.74], but research on the possibility of casualty going from education to these other traits is thinner, and mixed in results. Some may want to correct for family background variables and/or socioeconomic standing, as these things are indeed confounders [1197, pp.1843-1844]. However, there is a high enough degree of collinearity among —background, cognitive ability, and returns— that correcting for cognitive ability alone suffices [1209].

On Science:

An often boasted positive externality of our educational system is that regardless of its effects on individual skill in acquiring personal resources, it may advance societal progress by generating knowledge through advancing science. Supposedly, it should do this by training people to have the skills which are necessary to do science well, and by providing these people with an environment conducive to doing good science. To assess how well Academia helps society do this, we should first assess what science even is and how it should be done in order to contrast with what Academia actually does.

What Is Science?

Science, at root, is the art of deductively theorizing about how reality works. The overarching goal of the endeavor is to identify real phenomena and explain them with good theories and models. No theory is ever exactly correct because we will likely never have identified all existing phenomena, but some are useful because they elegantly approximate the well established aspects of reality.

Observation:

Without observations grounded in reality, we are left only with pure logic, mathematics, and philosophy. Mathematicians can perfectly formulate their logic in that they can know with certainty what exactly —the implications of any given set of premises— are, but they have no idea what they are talking about; without sound premises to work from, the implications that they derive are not likely to be applicable to anything. For example, if it is proclaimed from the heavens that people named y are on average three times as wealthy as people as x, and that people named z are on average three times as wealthy as people named y, then mathematicians can, correctly, tell us that people named z are on average nine times as wealthy as people named x. However, if we don't have sufficient reason to believe that the heavens have proclaimed to us accurate premises (in this case being the true relations between the variables), then any logical conclusions we draw from such premises do not accurately describe reality either.

Phenomena are identified through observation, and in order to see anything at all, we must make sure that our measurements (our senses) work adequately. This means designing good operationalizations (good measures) of whatever we've decided to observe. Given proper operationalization, any patterns that emerge before us from our observations will be expressed in the language of statistical terms. Statistics and measurement are paramount in identifying genuine phenomena.

Relativism:

It can be difficult to know how good our measurements are. People often see what they want to see and will keep measuring until they see patterns which fit with theories of a certain character. Given a good taste of this problem, scientists often despair and descend into relativism, the ultimate end of which is solipsism. This is an ultimately useless endeavour. Biases have effects, but reality does too. Confidence in the existence of various phenomena can be increased if their observation is robust to (unresponsive to) various biases, meaning that the same general patterns are observed from repeated measurement by multiple people with multiple different operationalizations. It is a good sign when the reduction of biases is shown to increase the clarity with which phenomena are observed. Given the demonstration of the influence of biases, the correct response is to search for analyses with the ability to robustly discriminate between biases and reality.

Elegance Vs Convolution:

Warning: $P(H|E) \neq P(E|H)$. Oftentimes, the known phenomena of a field can be explained by multiple different theories which are sometimes very different in character. In order to obtain the most justifiable possible view of the world, we must figure out how we should discriminate between them. Given a set of phenomena, the explanations most likely to be correct tend to be the simplest ones. One theory could state something akin to that —the laws of the universe, as of the year 2000, dictate Ron to be 175 centimeters tall and 13 years old, Karl to be 183 centimeters tall and 15 years old, and Charles to be 191 centimeters tall and 17 years old—. By contrast, a more elegant theory could state something akin to the following: —people tend to get taller as they age, so all else being equal, the older person should be the taller person. So, given that Charles is the oldest person and Ron is the youngest person, the rank ordering of their height fits our theory—. The elegant theory of aging is not able to explain the data as well as the convoluted one despite it intuitively seeming to be somehow better. In fact, the convoluted theory will always be the one which is best equipped to explain existing phenomena. In other words, the ability of the convoluted theory (TO) to explain the data (E) is much higher than the ability of the elegant theory (TE) to do so. However, a theory of aging should generally be thought of as better than the convoluted one due to its elegant ability to inexactly approximate the data before having been exposed to all of it. If TO and TE are hypotheses (H), then P(E|TO) is much larger than P(E|TE) despite P(TE) being much larger than P(TO) (Note that "P(x|y)" is read as "the probability that x is true given that y is true".).

Bayes' Theorem:

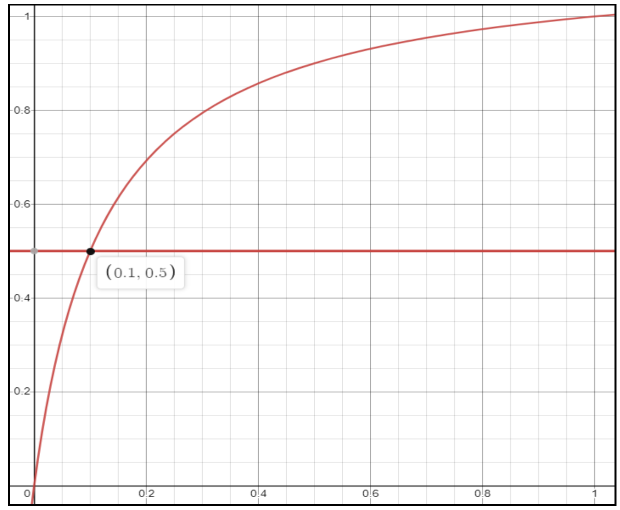
If we know the probability that a customer orders y given that they order x (P(y|x) and the probability that they customer order x (P(x)), then we can solve for the probability that they customer order both (P(y|x)*P(x)=P(x,y)). If we know the probability that a customer orders x given that they order y (P(x|y), and we know the probability that they order y (P(y)), then we can solve for the probability that a customer orders both (P(x,y) = P(x|y)*P(y)). Recall that if x=y and y=z, then x=z. Here, P(y|x)*P(x) = P(x,y), and P(x,y) = P(x|y)*P(y). Given the truth of these two equalities, P(x|y)*P(y) = P(y|x)*P(x). This is just an algebraic rearrangement of Bayes' Theorem, which states that P(A|B) = P(B|A)*P(A)/P(B). This is also equivalent to P(A,B) / P(B), and to P(B|A)*P(A) / P(B|A)*P(A)+P(B|¬A)*P(¬A) (Note that "P(¬x)" denotes the probability that x is not true, which is equal to 1 - P(x)).

Theories As Compositions Of Hypotheses:

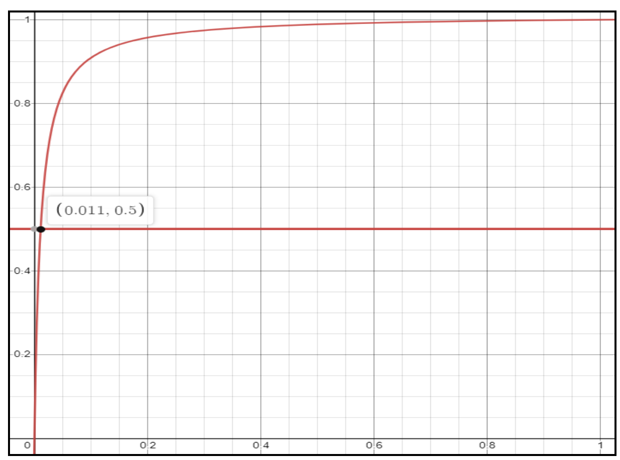
We can use Bayes' Theorem to decide how likely a hypothesis is to be true. Let "P(H)" denote the prior believed probability of a hypothesis, let "P(E)" denote the prior believed probability of some evidence, and let "P(E|H)" denote what —the probability of the evidence— would be in a reality where the hypothesis has a 100% chance of being true. P(E) is of course derivable by calculating P(E|H)*P(H)+P(E|¬H)*P(¬H). If something is to be thought of as being true by reason alone, and if there is not yet reason to think that P(H) is higher than P(¬H), then given that P(H) and P(¬H) are, by definition, mutually exclusive such that one of the two must be true and P(¬H) + P(H) = 1, we should think of the two possibilities as being equally plausible, meaning that P(H) = P(¬H) = 0.5.

Returning to the example by which we derived Bayes' Theorem, if our hypothesis (H) is that the next customer will order x, if the only alternative possibility is that the next customer will not order x, and if we have no reason to think that the alternative possibility is more likely to be true than the former possibility, then P(x) = P(¬x) = 0.5 with P(x) and P(¬x) being prior beliefs about x. Let's assert that P(y|x) = 0.9, and that P(y|¬x) = 0.1. Our hypothesis (H) is still that the next customer will order x, but we now have information about the relationship between the hypothesis and the evidence (E). Given these assertions, 0.9 = P(H|E) = (P(E|H) * P(H)) / P(E) = (P(y|x) * P(x)) / P(y) = (P(y|x)*P(x)) / P(y|x)*P(x)+P(y|¬x)*P(¬x) = 0.9*0.5 / (0.9*0.5 + 0.1*0.5) = 0.9.

Now, let's assert that 30% of customers order x. This new information denotes that P(x) is now 0.3, and P(¬x), by definition, is now 0.7. Reworking our calculations, P(H|E) (or P(x|y)) is now .7941176471, meaning that taking our assertions for granted, a customer who orders y would have a 79.41176471% chance of ordering x if 30% of all customers (y-ordering or otherwise) order x. Notice that given our assertions, if a function yielding P(H|E) in terms of P(H) were written such that a(b)=P(H|E) and b=P(H) (This function being 0.9*b/(0.9*b+0.1*(1-b)) = a(b)), different sets of hypotheses on the interval 0 ≥ b ≤ 1 (e.g. b1 & b2), holding b1 - b2 constant, would yield different a(b1) - a(b2) figures; in other words, the function is such that switching from one hypothesis to a second may not have the same effect on the posterior as switching from a third to a fourth, and this isn't exclusively a function of the difference between hypotheses. Indeed, we can calculate confidence regions such that we can calculate exactly how bigoted we would have to be in our prior beliefs in order to get the posterior to be outside of a certain range. Here is a graph of the function such that the vertical axis is P(H|E) and the horizontal axis is P(H):



Notice that, for instance, 90% of possible choices of prior belief yield a posterior probability larger than 50%. Also notice that if we arbitrarily make the P(E|H) and P(E|¬H) values more extreme, the bigotry in prior required to reach the same threshold of posterior has to also become more extreme:



In the latter graph, a change of P(E|¬H) from 0.1 to 0.01 resulted in a change in —≥50% threshold— from 90% of P(H) values to 98.9% of them.

For now, let's abandon the idea of thinking about an infinite number of hypotheses and just stick to P(x) = 30%. Given this hypothesis, P(x|y) = 79.41176471%. Now, let's consider the probability that a customer who orders item z also orders item x. Let's assert that a customer who orders item z has a 90% chance of ordering item x, that a customer who does not order item z has a 10% chance of ordering item x, and that the probability of a customer ordering item z is unrelated to the probability that said customer orders item y (Meaning that P(y|z) = P(y|¬z) P(y)). These are the same parameters we were working with in order to derive that P(x|y) = .7941176471, so P(x|z) should return the same value. However, let's instead consider P(x|y,z). If y is true, then the probability of x is not 30%, but 79.41176471%. Given the hypothesis that P(x) is 79.41176471% from the start, P(x|z) would not be 79.41176471%, but rather 97.2%. Thus, P(x|y,z) = 0.972. Given the evidence of y, our prior for P(x) becomes a posterior of 0.794. Treating the posterior as the prior when considering newer evidence z, we can come up with a theory-wide posterior of 0.972 for our theory that the next customer will order item x when starting with a theory-wide prior of 30%, 0.3 being P(x). With strong enough evidence or enough lines of evidence in favor of a theory, arbitrary choices of prior can have miniscule influence on the posterior.

Of course, parameters like P(x|y), P(x|¬y), P(x|z), P(x|¬z), P(y|¬z), and P(y|¬z) are almost never simply given, but must instead be derived by statistical inference where some data discordantly supports the various hypotheses about P(x|y) to different degrees and apparent patterns in the data always have some chance of being apparent due to mere random noise in the data. Rather than asserted values being plugged in, parameters would be substituted for entire probability density functions in order to get a theory-wide posterior distribution. How to do this is beyond current scope and can be read about either in source 1212 or in [chapter 1].

P-Values:

Academia is currently obsessed with obtaining results which pass a criteria known as "statistical significance" [more here & here]. Basically, there is a statistic called a p-value which can be computed to go along with any given effect size statistic. When scientists want to know whether or not a hypothesis predicts their data, they operationalize this by saying that it would predict a certain effect size statistic of a certain magnitude. A p-value, given figures for effect size and statistical power, tells us the chance that we would see an effect size at least as substantial as —what really appears in the data— if the hypothesis were false and the observed effect size were really just the result of random fluctuations in the data. In other words, p-values only tell us about P(E|H).

Ordinarily, there is a threshold of 5%, or p=0.05, where a result is arbitrarily declared to have met the criteria of "statistical significance". There is no objective reason to place the threshold at 5% vs 4% or 1%, it's just that 5% is commonly accepted as being subjectively low. Though arbitrary, this threshold is popular enough to matter such that authors are more likely to submit their significant results than their insignificant ones, such that their colleagues are more likely to cite their significant results than their insignificant onets, and such that prestigious journals being more likely to publish their significant results than their insignificant results [more here & here]; this leads to serious distortions of the research literature's view of what results' true effect sizes really are [more here].

Moreover, this likely leads to serious distortions of the research literature's view of what P(H|E) is, as P(H|E) is a function of more than just P(E|H); it is P(E|H) divided by P(E) rather than just P(E|H). A reason that elegant theories which offer simple explanations for known phenomena have a greater tendency to be correct than convoluted ones is that if, for example, a generic hypothesis has a random 50% chance of being true, then a theory which requires one hypothesis has a 50% chance of being true while a theory which requires two has a 25% chance of being true. Take for example the observation that you come home from work and your window is broken, your laptop is missing, and your front door is unlocked. A burglary, if it happened, would have a pretty high chance of producing this evidence, say 80%. Of course, there is a potential alternative hypothesis where you left your laptop at work, a neighborhood kid hit a baseball through your window, and you forgot to lock the front door. The alternative hypothesis, if true, has a higher likelihood of explaining the observations, but is it more likely to be true? Let's assert that 1/1,000,000 houses get burglarized per day, that 1/100,000 houses per day get a baseball accidentally sent through one of its windows, that you tend to accidentally leave your laptop at work once per every 100 days or so, that you forget to lock the front door once per every 100 days or so, and that the probabilities of these events are all independent of the probabilities of the others. The burglary theory has a prior of

1/1,000,000, but the accident theory has a prior of 1/100,000 times 1/100 times 1/100 = 1/1,000,000,000. The theory with a lower prior has a 100% chance of explaining the observation, but by contrast the burglary could've been done without breaking the window, giving the observations an 80% chance of occurring in the case of an average burglary. However, when combining explanatory power with priors, we see that the product of 1 and 0.000000001 is 1/800th the size of 0.8 times 0.000001, meaning the burglary, despite its lower ability to explain the data, is 800 times as likely to be true. Of course, if you get a call from your boss saying you left your laptop at work, and a visit from an angry parent making their children apologize for their reckless behavior, the new information should update our priors to 1 times 1 times 1/100, suddenly making the accident theory very likely to be true.

In sum, elegant theories tend to be more parsimonious than convoluted ones because elegant ones merely lack the ability to explain phenomena while convoluted ones predict the existence of suites of unverified phenomena. When a theory is contrived in the mere pursuit of the lowest p-values, this can easily come at the expense that the theory depends on the plausibility of a suite of potential phenomena whose implausibility is esoteric.

Experimentation:

Currently elegant theories aren't always necessarily correct. In a given paradigm, there are sometimes multiple elegant theories which are all able to explain the currently established phenomena. Progress is made by figuring out how to discriminate against the incorrect ones. While equally able to explain currently established phenomena, competing theories are often very different in character, and these differences in character are what makes it of interest to discriminate between them. Fortunately, the stark differences in character often mean that meaningfully distinct, competing theories often make starkly different predictions of the existence of various unidentified phenomena. To discriminate between currently elegant theories, we must design circumstances such that proper analyses of the data they generate are equipped to convincingly evidence the existence or non-existence of predicted phenomena. The act of doing this is called experimentation.

The key to experimentation is to design a set of circumstances under which —the resulting patterns which we detect with our operationalizations and statistics— can only be explained by one or more currently competing theories being incorrect.

It is extremely important to ask the right research questions and design experiments which are actually able to answer them. We can get our operationalizations, statistics, and biases nailed down very well, but if we design an experiment which is only equipped to assess whether or not, for example, a raw correlation exists, then the high clarity of the resulting statistical signal is often of meager usefulness and illumination. It is often better to get a rough, approximate answer to the right research question (such as whether or not a causal effect exists) than it is to get an extremely clear answer to the wrong research question (such whether or not a raw correlation exists).

A good statistician must be brought in to design experimental circumstances well before data is even collected so that it can be known ahead of time that useful conclusions can be taken from the analysis, whatever its results culminate in. When the statistician is brought in after an experiment, he often can only do a postmortem assessment where he uncovers what went awry and impaired its elucidative value.

Limitations:

There are often various reasons why the right experiment cannot be done or the right observation cannot be made. In physics, black holes have gravity which is too strong to let light escape, so we cannot measure what goes on inside the event horizon. In the social sciences, it is an unethical research practice to experimentally cut people's arms off in order to assess the impact of dexterity on quality of life. However, there are sometimes various 'natural experiments' where naturally occurring circumstances allow observation to be informative without much effort.

For example, it used to be unclear what impact raw wealth has on fertility. It is hard to experimentally manipulate wealth due to such experiments requiring large amounts of wealth from the experimenter. However, there have been natural experiments where fluctuation in home value, or local oil revenue, influence peoples' wealth for reasons unrelated to what ordinarily causes individual differences in wealth. In these circumstances, it has been convincingly demonstrated that gains in wealth actually cause increases in fertility despite the ordinary observation that the raw correlation between fertility and wealth is negative [1164 & 1165]. Theoretical progress has thus been made in this sliver of social science despite the fact that the obvious experiment is difficult to do.

Similarly informative analyses can also sometimes be done without access to such special conditions. For example, let's assert that we have variables A, B, C, and D. One theory posits that D causes C and A, and that C causes B. Another theory agrees that C causes B, but posits that D does not exist at all, and that by contrast, B causes A. D is unobservable, but if C is held constant, then D no longer has any bearing on how A and B covary. In this case, the effect of B on A can then be applied to the effect of C on B. Then, the potential influence of D is no longer a worry. There is at least one known case where this method was used in a dataset where the results of an actual randomized control trial were available, and the results closely mirrored that of the true causal effect [1214].

For more on statistical methods for causal inference, see source 1213 or [chapter 1].

On Peer Review:

The Incentives:

Many members of the general public have never been involved in the process of publishing a scientific paper, even many of those who are highly scientifically literate. Given this, they likely do not know how the peer review process actually operates at an experiential level. Given this, the following quote contains a quite scathing description of the process from a standard peer reviewed paper [4] by J. Scott Armstrong, a standard, public academic. Reference names are replaced with source numbers:

> *"Here is how the current quality control system works. Researchers, sometimes working in teams, spend hundreds of hours working on a specialized topic, often collecting empirical evidence and applying formal analytical techniques. They write papers and often benefit from pre-submission peer reviews. They strive to follow standards for scientific work and they sign their names to their work. Their futures depend to some extent on the quality of their paper. These papers are then reviewed by people who are working in related areas but generally not on that same problem. So the reviewers typically have less experience with the problem than do the authors. Of course, there may be aspects of the research, such as methodology, in which the reviewers have more expertise. Reviewers generally work without extrinsic rewards. Their names are not revealed, so their*

Continued:

> *reputations do not depend on their doing high quality reviews. Perhaps this leads them to spend little time on their reviews. In any event, on average, reviewers spend between two and six hours in reviewing a paper (49; 50; 51; 52), although they often wait for months before doing their reviews. They seldom use structured procedures. Rarely do they contribute new data or conduct analyses. Typically, they are not held accountable for following proper scientific procedures. They match their opinions against the scientific work by the authors… Reviewers appear to base their judgments on cues that have only a weak relation to quality. Such cues include (1) statistical significance, (2) large sample sizes, (3) complex procedures, and (4) obscure writing. Researchers might use these cues to gain acceptance of marginal papers (34, page 197).*
>
> *Although it typically has little relationship to whether the findings are important, correct, or useful, statistical significance plays a strong role in publication decisions as shown by studies in management, psychology, and medicine [Sources 35, 38, 39, 40, & 41]. The case against statistical significance is summarized for psychologists by [Source 42] and for economists by [Source 43]. If the purpose is to give readers an idea of the uncertainty associated with a finding, confidence intervals would be more appropriate than significance tests.*
>
> *[Source 44] conducted an experiment to determine whether reviewers place too much emphasis on statistical significance. They prepared three versions of a bogus manuscript where identical findings differed by the level of statistical significance. The reviewers recommended rejection of the paper with*

Continued:

nonsignificant findings three times as often as 7 the ones with significant findings. Interestingly, they based their decision to reject on the design of the study, but the design was the same for all versions.
Using significance tests in publication decisions will lead to a bias in what is published. As [Source 45] noted, when studies with nonsignificant results are not published, researchers may continue to study that issue until, by chance, a significant result occurs. This problem still exists [Source 41].
Large sample sizes are used inappropriately. Sometimes they are unnecessary. For example, reviewers often confuse expert opinion studies with surveys of attitudes and intentions. While attitudes and intentions surveys might require a sample of more than a thousand individuals, expert opinion studies, which ask how others would respond, require only 5 to 20 experts [Source 46, p. 96]. Even when sample size is relevant, it is likely to be given too much weight. For example, source 47, in a study of election polls for the U.S. presidency, concluded that the sample size of the surveys was loosely related to their accuracy.
Complex procedures serve as a favorable cue for reviewers. One wonders whether simpler procedures would suffice. For example, in the field of forecasting, where it is possible to assess the effectiveness of alternate methods, complex procedures seldom help and they sometimes harm accuracy [Source 46]. Nevertheless, papers with complex procedures dominate the forecasting literature. Obscure writing impresses academics. I asked professors to evaluate selections from conclusions from four published papers [Source 48]. For each paper, they were randomly assigned either a complex version

Continued:

(using big words and long sentences, but holding content constant), the original text, or a simpler version. The professors gave higher ratings to authors of the most obscure passages. Apparently, such writing, being difficult to understand, leads the reader to conclude that the writer must be very intelligent. Obscure writing also makes it difficult for 8 reviewers and readers to find errors and to assess importance. To advance their careers, then, researchers who do not have something important to say can obfuscate."

The wait for many authors to get a paper published can be even longer than discussed by Armstrong because a rejection doesn't mean that they have to delete their paper, it just means that if their heart is set on publishing, they just have to keep on going through journal after journal while never being allowed to be reviewed by multiple journals at the same time. 85% of the papers rejected by the Journal of Clinical Investigation were eventually published elsewhere, and the majority of these were either not changed or changed in only minor ways [53]. Source 4 reports that source 54 obtained similar results for papers rejected by the British Medical Journal, but I could not find the full text of source 54, just the citation. Source 55 reached a similar conclusion for papers in the social sciences. Source 56, in a study of papers rejected by the American Political Science

Review, concluded that of the 263 papers which were then submitted to another journal, 43% contained no revisions based upon the APSR reviews. It would seem that however much quality a journal's peer review actually demands, it doesn't actually guarantee improvements until standards are met because papers can just be endlessly reviewed until publication.

On the use of obscure language, scientists have created over 1,000,000 acronyms since the 1950's, the rate of creation has been accelerating, and almost 80% have been used fewer than 10 times [327].

Publish Or Perish:

At the time of its inception in 1955, Eugene Garfield, the creator of the impact factor, did not imagine that some day his tool would become a controversial and abusive measure. Originally it was just meant to be a tool to help librarians choose which material to order for their libraries in order to satisfy the most researchers by measuring the popularity of research [78], little did he imagine how much the scope of its use would expand. Focus groups of scientists report career pressures to publish high volumes of papers with positive results that confirm orthodoxy in high impact factor journals [74]. Universities want to be able to say that all of their professors publish in all of the 'best' journals. Many universities do not focus on teaching ability when they hire new faculty and simply look at the publications list [75]. Tragically, in some countries, the number of publications in journals with high impact factors condition the allocation of government funding for entire institutions [76].

For many, it is publish or perish.

Just as quantitative evidence repeatedly shows that financial interests can influence the outcome of biomedical research [79 & 80], publish or perish culture affects all manner of research behavior including salami slicing [81] to publish the shortest papers one can get away with. In 2006 alone, an estimated 1.3 million papers were published alongside a large rise in the number of available scientific journals from 16,000 in 2001 to 23,750 by 2006 [82]. The number of journal articles is estimated to have passed 50 million in 2009 [83].

Journal rank is most commonly assessed using Thomson Reuters' Impact Factor which has been shown to correspond well with subjective ratings of journal quality and rank [84, 85, 86 & 87]. However, despite the perceived prestige and the importance placed on the impact factor, all evidence seems to suggest that the perverse incentives actually causes papers published in high impact factor to be more unreliable on average than papers published in "worse" journals.

Journal rank is predictive of the incidence of fraud or misconduct being the reason for a paper's retraction as opposed to other reasons for retraction [133 & 89], and larger journals also have more total retractions [13]. The fraction of retractions made due to misconduct has risen more sharply than the overall retraction rate, with the majority of retractions now being due to misconduct [89 & 90]. This is consistent with focus groups which suggest that the need to compete in academia is a threat to scientific integrity [74], with the fact that those found to be guilty of scientific misconduct often invoke excessive pressures to produce as partial justification for their actions [91], and with surveys suggesting that competitive research environments decrease the likelihood that researchers follow scientific ideals [92] while increasing the likelihood to witness scientific misconduct [93].

Although 77% of variance in journal retraction rate is accounted for by journal rank [89], retracted papers are such a low percentage of papers that it is possible that the number of retractable papers is higher than the number of retractable papers which have actually been caught and retracted, or that detection problems partially contribute to the strength of this relationship making increased readership in high ranking journals more likely to detect errors.

It isn't possible to measure the contribution of such detection effects, so what can other measures of quality say about the effect of impact factor on the rest of publications?

When aiming to compare the quality of papers in larger journals to papers in smaller ones, some aspects of an article's quality can be rather subjective things to analyze. This is supposed to be judged by the peer review process itself, but peer review is the very thing under scrutiny. However, what we can do is look at traits like statistical power, and if one journal repeatedly has underpowered studies, we can take that as a proxy for other qualities.

Source 5 has many such proxies, statistical power being one of them. A sample of 650 neuroscience studies showed no relationship between statistical power and journal impact factor:

Source 5 - Figure 3, data from source 6:

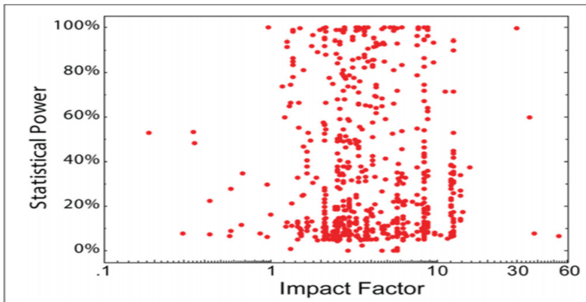


FIGURE 3 | No association between statistical power and journal IF. The statistical power of 650 eligible neuroscience studies plotted as a function of the IF of the publishing journal. Each red dot denotes a single study. Figure from Brembs et al. (2013).

Another indicator was crystallographic quality (the quality of computer models derived from crystallographic work) This lets us see how often journals deviate from known atomic distances, and what is found, is that higher impact journals have worse crystallographic work, meaning that their molecular models have more errors than the lower impact journals:

Source [5](#) - Figure 1, data from source [8](#):

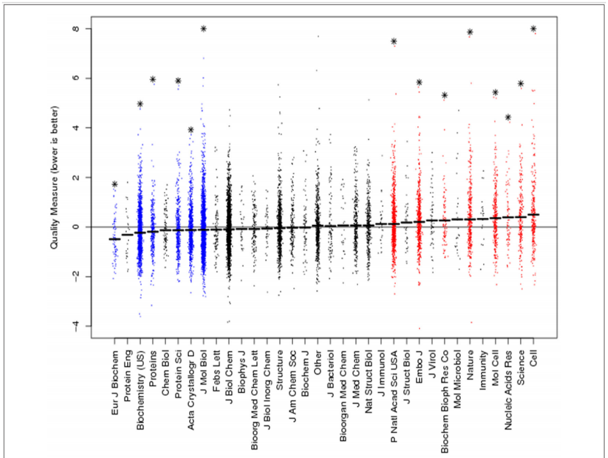


FIGURE 1 | Ranking journals according to crystallographic quality reveals high-ranking journals with the lowest quality work. The quality metric (*y*-axis) is computed as a deviation from perfect. Hence, lower values denote higher quality work. Each dot denotes a single structure. The quality metric was normalized to the sample average and journals ranked according to their mean quality. Asterisks denote significant difference from sample average. Figure courtesy of Dr. Ramaswamy, methods in Brown and Ramaswamy (2007).

We could say that this is a rather limited indicator of journal quality, fair enough, and to the extent that this is an indicator of other qualities is unknown, but it's another objective trait to add to the list.

Figure 5 looks at the rate in which papers from various journals get gene symbols of SNPs wrong. Taking nature as an example, a journal famous enough for me to know about it, about ⅓ of all genetics papers mislabel some bit of genetic data somewhere in the paper.

Source [5](#) - Figure 5, data from source [10](#):

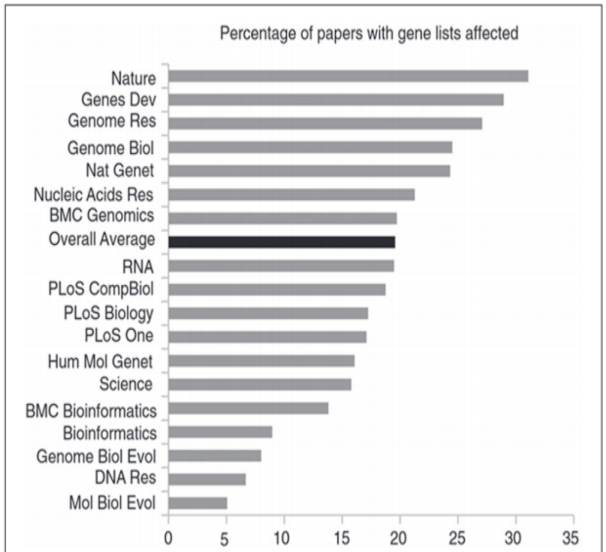


FIGURE 5 | Journals with above-average error-rate rank higher than journals with a lower error-rate. Shown is the prevalence of gene name errors in supplementary Excel files as the percentage of publications with supplementary gene lists in Excel files affected by gene name errors. Figure modified from Ziemann et al. (2016).

Not that a mislabeled piece of data here and there is the biggest deal ever, but it's another objective indicator of quality.

Figure 4 looks at how often studies have randomized control trials, and how many of them had double blind results in experiments on animals (Practices that exist to attempt to limit the influence of author bias on a study's results). What was found was that higher impact journals had roughly the same rate of blinding as lower impact journals, but less randomization than lower impact ones.

Source 5 - Figure 4, data from source 11:

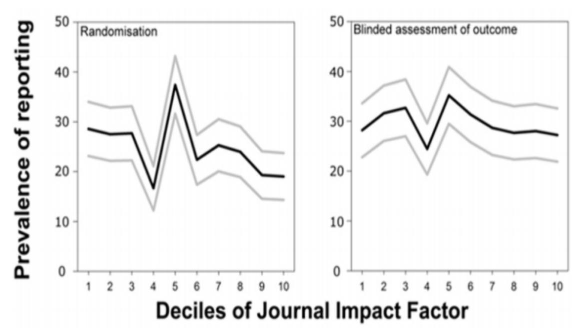


FIGURE 4 | High ranking journals do not have a higher tendency to report more randomization nor blinding in animal experiments. Prevalence of reporting of randomization and blinded assessment of outcome in 2671 publications describing the efficacy of interventions in animal models of eight different diseases identified in the context of systematic reviews. Figure modified from Macleod et al. (2015).

Figure 6 shows a correlation between journal impact factor and the miscalculation of p-values:

Source 5 - Figure 6, data from source 12:

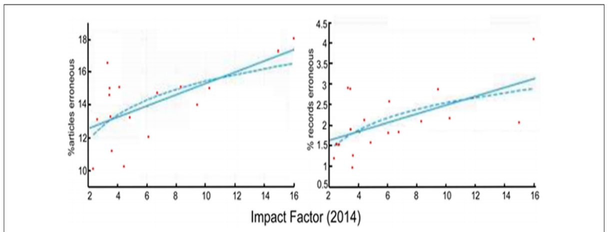


FIGURE 6 | *p*-value reporting errors correlate significantly with journal rank. The correlation of the median percentage of articles with erroneous articles (left; which can contain multiple erroneous records) or individual records (right) in a given journal and journal IFs. Both linear and logarithmic (log[journal IF]) trend lines are shown. Figure redrawn from Szucs and Ioannidis (2016).

The percentage of papers with at least 1 miscalculated p-value in the paper was around 18% in the highest impact journals and around 12% in the lowest impact ones. Higher impact journals had about 3% of p-values miscalculated while lower impact one had 1.5% of p-values miscalculated. Another objective sign that larger journals don't publish better papers.

Sidenote: P-values:

If a statistical signal should not exist for the full population, then there's a small chance that a random collection of data from a random sample from the population would appear in such a way as to make it look like there were genuine signal being detected. A p-value, using statistical power and effect size, calculates the chance that a result would look at least as extreme as it appears to be if the null hypothesis were actually true.

Caution: $P(X|Y) \neq P(Y|X)$:

Here is a pop quiz for the reader:

1% of women have breast cancer. 80% of women with breast cancer get positive mammograms. 9.6% of women without breast cancer also get positive mammograms. What is the probability that a woman with a positive mammography actually has breast cancer?

Let's say that we have 10,000 women. 1% of them have breast cancer, so 100 have breast cancer. Of the 100 women with breast cancer, 80% of them (so 80 women) get positive mammograms, the other 20 do not. Of the 9,900 who do not have breast cancer, 9.6% of them get positive mammograms (so 950 women). To recap, 80 women with positive mammograms have breast cancer, 950 do not. In total, 1,030 women have positive mammograms, and of these, 7.8% have breast cancer.

Remember what a p-value is: A p-value calculates the chance that a result would look at least as extreme as it appears to be if the null hypothesis were actually true. It does not calculate the probability that the null hypothesis would be true given a result which looks as it does. It can help inform us what such a probability is, but extra thought is required.

"P(x|y)" is the probability of x given y is true.

Figure 2 looks at the effect size in gene association studies divided by the pooled effect size estimate derived from a meta analysis. A higher number means that a study's effect size deviates from the results that most papers find than a study with a lower number, and the larger the circle, the larger the sample population that was used. What this shows, is that higher impact factor journals have smaller sample sizes, with bigger, flashier, more exciting results which aren't replicated:

Source 5 - Figure 2, data from source 9:

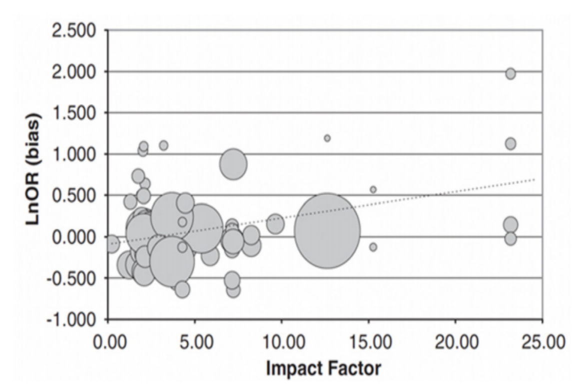


FIGURE 2 | Relationship between impact factor (IF) and extent to which an individual study overestimates the likely true effect. Data represent 81 candidate gene studies of various candidate genes with psychiatric traits. The bias score (*y*-axis) represents the effect size of the individual study divided by the pooled effect size estimated indicated by meta-analysis, on a log-scale. Therefore, a value greater than zero indicates that the study provided an overestimate of the likely true effect size. This is plotted against the IF of the journal the study was published in (*x*-axis). The size of the circles is proportional to the sample size of the individual study. Bias score is significantly positively correlated with IF, sample size significantly negatively. Figure from Munafò et al. (2009).

Also, the efficacy of high impact factor journals should not be a surprise given the substance of what impact factor actually is and how it is calculated [6].

Publication Bias:

Source 5 - Figure 2 is evidence that journal rank / publish-or-perish culture is tied to the decline effect of publication bias [6]. The decline effect is basically the phenomenon that the first paper which observes an effect has a large effect size, but subsequent papers that attempt to replicate the first either fail to replicate it or come up with much lower effect sizes. The usual pattern is of the initial study being published in a high impact journal followed by smaller journals showing that the effect fails replication. One particular case showcasing this pattern in the decline effect is source 94. Source 77 makes a good introduction to the evidence on publication bias, to quote from it, keeping the sources but replacing source numbers, see the following:

> *"In many fields of research, papers are more likely to be published [95, page.371; 96; 97; & 98], to be cited by colleagues [99, 101, & 102] and to be accepted by high-profile journals [103] if they report results that are "positive" – term which in this paper will indicate all results that support the experimental hypothesis against an alternative or a "null" hypothesis of no effect, using or not using tests of statistical significance. Words like "positive", "significant", "negative" or "null" are common scientific jargon, but are obviously misleading, because all results are equally relevant to science, as long as they have been produced by sound*

Continued:

logic and methods [104, & 105]. Yet, literature surveys and meta-analyses have extensively documented an excess of positive and/or statistically significant results in fields and subfields of, for example, biomedicine [106], biology [107], ecology and evolution [108], psychology [109], economics [110], sociology [112]. Many factors contribute to this publication bias against negative results, which is rooted in the psychology and sociology of science. Like all human beings, scientists are confirmation biased (i.e. tend to select information that supports their hypotheses about the world) [113, 114, & 115], and they are far from indifferent to the outcome of their own research: positive results make them happy and negative ones disappointed [116]. This bias is likely to be reinforced by a positive feedback loop from the scientific community. Since papers reporting positive results attract more interest and are cited more often, journal editors and peer reviewers might tend to favour them, which will further increase the desirability of a positive outcome to researchers, particularly if their careers are evaluated by counting the number of papers listed in their CVs and the impact factor of the journals they are published in. Confronted with a "negative" result, therefore, a scientist might be tempted to either not spend time publishing it (what is often called the "file-drawer effect", because negative papers are imagined to lie in scientists' drawers) or to turn it somehow into a positive result. This can be done by re-formulating the hypothesis (sometimes referred to as HARKing: Hypothesizing After the Results are Known [118]), by selecting the results to be published [119], by tweaking data or analyses to "improve" the outcome,

Continued:

or by willingly and consciously falsifying them [120]. Data PLoS ONE | www.plosone.org 1 April 2010 | Volume 5 | Issue 4 | e10271 fabrication and falsification are probably rare, but other questionable research practices might be relatively common [121]. Quantitative studies have repeatedly shown that financial interests can influence the outcome of biomedical research [79 & 80] but they appear to have neglected the much more widespread conflict of interest created by scientists' need to publish.

Source 77 also provides direct evidence that publish or perish culture is tied to publication bias. It looks at U.S. states by how many papers are published in each state and how often positive results are achieved. More 'productive' states have more publication bias. Controlling for per capita research expenditure and/or a few other variables strengthens the relationship.

Interrater Reliability:

Related to publication bias is inter-rater reliability. While low interrater reliability doesn't necessitate publication bias, low inter-rater reliability is evidence that the peer review process doesn't follow a consistent standard, and thus doesn't follow an objective one since disagreement means that at least one party is wrong. Further, low inter-rater reliability itself is evidence that the journal system is capable of contributing to publication bias rather than publication bias being entirely a function of self selection among the authors themselves. If reviewers were akin to two computers running the same objective algorithms on the same exact paper, you would expect them to come to pretty similar conclusions. If inter-rater reliability is low, that's a good sign of subjectivity which gives room for people to put their own bias into the process.

In 2000, the journal *Brain* (an Oxford publication) looked into reviewer agreement at other journals [25]. Unfortunately, those journals only agreed to this on the condition that they remain anonymous, so we're trusting Oxford that they picked a 'good' selection.

Journal A:

Acceptance agreement: 47% vs. 42.5% by chance alone Priority agreement: 35% vs. 42.5% by chance alone

Journal B:

Acceptance agreement: 61% vs. 45.74% by chance alone Priority agreement: 61% vs. 46.32% by chance alone.

By the way, I inferred the numbers for chance here by counting the pixels in the bar chart. Readers may find this silly and absurd, but this is something I find myself having to do quite often when looking at published peer-reviewed papers that don't have supplementary data posted. To be fair, anonymity was guaranteed, but detailed data could have been provided that just has names omitted.

Source 26 is a meta analysis going over 48 studies on inter-rater reliability, and they found that the average amount of agreement was about 0.34/1.00, 0 being the lowest possible among of agreement and 1 being the maximum. In addition, if you look at source 26 - Figure 1, you'll see that within journal agreement varies wildly and that agreement above 0.8 is never achieved:

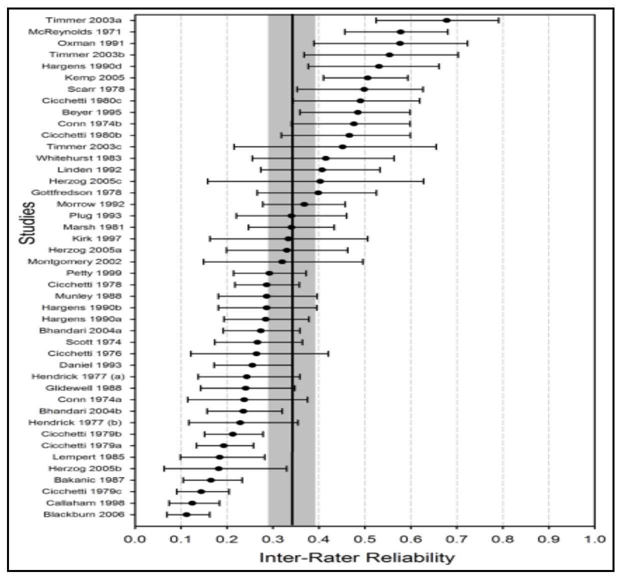


Figure 1. Forest plot of the predicted inter-rater reliability (Bayes estimate) for each study (random effects model without covariates) with 95% confidence interval (as bars) for each reliability coefficient (sorted in ascending order). The 95% confidence interval of the mean value (vertical line) is shaded grey. Predicted values for the same author and year but with different letters (e.g., Herzog 2005a and Herzog 2005b) belong to the same study. doi:10.1371/journal.pone.0014331.g001

Results on inter-rater reliability are yet again confirmed in Domenic Cicchetti's 1991 paper, "*The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation*" in Cambridge's Behavioral and Brain Sciences journal [27].

Results Bias:

As mentioned earlier [more here], there are more positive results than we would expect from random chance in various fields such as biomedicine [106], biology [107], ecology and evolution [108], psychology [109], economics [110], sociology [112], etc. Contrary to popular perception, this isn't, by itself, evidence of publication bias. The skew could be because of publication bias, or it could just be that hypotheses aren't randomly generated, and the hypotheses that scientists come up with are more likely to be true than a randomly formulated hypothesis. This stated, we know from experimental evidence that publication bias is a contributing factor.

In *Review Bias*, from Annals of Internal Medicine [15], the author sent out papers to various journals about transcutaneous electrical nerve stimulation (TENS), the paper they wrote was fake, and they wrote two identical versions of the paper, one with a positive result (a result which supported the hypothesis), and one with a negative result. The positive result was sent to 8 journals and the negative result was sent to 8 different ones. We can see from source 15 - Table 1, that in this sample at least, the results matter in terms of how reviewers judge the quality of study design, patient descriptions, statistical methods, end points, and linguistic quality.

Source 15 - Table 1:

Table 1. Absolute Numbers of Ratings Given by "Pro" and "Contra" TENS Reviewers*

Categories	1 (Unsatisfactory)		2 (Defective)		3 (Adequate)		4 (Good)		5 (Excellent)		*P* Value†
	Pro	Contra	Pro	Contra	Pro	Contra	Pro	Contra	Pro	Contra	
Study design	0	1	1	2	4	3	3	2	0	0	0.12
Patient description	1	3	0	3	3	1	4	1	0	0	0.04
Statistical methods	0	0	1	2	3	6	4	0	0	0	0.06
End points§	0	1	1	4	0	2	5	0	0	0	0.01
Linguistic quality	0	0	0	1	4	4	2	3	2	0	>0.2

* Numbers indicate absolute frequencies of ratings. TENS = transcutaneous electrical nerve stimulation.
† Using two-tailed probability, Wilcoxon rank test for independent samples.
§ No rating provided in two cases.

If reviewers disagreed with the result, they were more likely to say that the methodology is poor despite papers with different results having identical methodologies.

Another similar experimental manuscript sting operation [16] submitted 146 papers to various journals dealing with social work and what they classify as "allied disciplines". Negative papers were rated worse, and had more journals outright decline to review the paper at all.

In 2010, source 17 did a study on the efficacy of a randomized control trial on a form of joint surgery. There were 2 versions of the paper sent to 238 reviewers which were identical in everything except for the results. The two outcomes were a positive effect from the surgery, and no effect from the surgery. The reviewers were biased towards the versions with positive results. There were also 7 intentionally planted errors in both results of the paper. For the positive version, reviewers found an average of 0.41 errors; for the negative result, reviewers found an average of 0.85 errors. So on average, the reviewers found less than ⅐th of the intentionally planted errors. In terms of methods scores, positive results were rated better. In terms of acceptance of the manuscripts to even be reviewed, the positive version was accepted 97.3% of the time, and the negative version was accepted 80% of the time. Apart from confirming the previous findings on bias, what is particularly concerning here is how low the error detection rate is even when reviewers dislike the paper's results and obsessively put it under a microscope to look for errors in order to try to reject it.

Positive results have also been found to be more common in the soft sciences (e.g. social science like sociology) than in the hard sciences (e.g. natural science like chemical engineering) [136].

Something else of note is that systemic issues aren't the biases; publications with positive results are more likely to be cited by colleagues [99, 101, & 102]. Publications that fail replication are also more cited [1077], so this likely helps explain why.

Anti-Author Biases:

Source 21 looked at the acceptance rate of over 50,000 real papers based on author characteristics in American Heart Association Journals. The study also looked at a switch from open to double blind peer review where both the reviewers and the authors didn't know each other or anything about which institution the author came from. Prestigious institutions were 57.4% more likely to have their papers accepted in the open setting, but only 33.8% more likely to have their papers accepted in the closed setting. So whether you think that papers from prestigious institutions are better because they train their students better, or just because they select for better students during the admissions process, we see that on top of that quality advantage, they have an extra 23.6% premium, not for any tangible talent, but just for having the name of the prestigious institution printed next to their name. However, there were 3 studies with 5 experiments, and it was found that on average, when the review process was supposedly "double blind", the reviewers could still correctly guess who the author was 41% of the time. So, if we assume that we can linearly apply the observed blinding effect to the previous results in order to guess how well papers written by prestigious institutions actually do when the reviewers correctly guess who the author is 0% of the time, I would estimate papers from prestigious institutions to only be accepted only 13.7% more often, giving prestigious institutions a 43.7% premium that has nothing to do with the quality of their papers.

Source 22 managed to get the 2017 web search and data mining conference, which had a 15.6% acceptance rate, to have a singleblind review where the author doesn't know who the reviewer is, but also a double blind where the reviewer doesn't know who the author is either. They looked to see what the effect was from how famous the author was, whether or not the author came from a top company or university, and whether or not the author was female. Authors from a top company were 2.1 times as likely to get a paper accepted when the reviewer knows who the author is. Likewise, the premium for author fame was 1.63, the premium for university was 1.58, and the premium for being female was 0.78. If you were to attempt to apply the findings thus discussed and assume that double blind isn't really double blind, then the distances between all of these premium numbers and 1.0 would be more exaggerated than thus discussed. On a side note, I would like to point out that the last premium mentioned is something of a unicorn to me. This is an empirically validated instance of gender discrimination against women, but more importantly, this could realistically be interpreted as just being further prestige bias with women, on average, being less famous.

Effects On Quality:

Are peer-reviewed papers in general better than papers which straight up aren't reviewed at all? Yes, somewhat, but not impressively so. When simply surveying people, they certainly say that the effect of peer reviewing papers is to improve them. Over 70 out of the 96 responding authors in source 30 said that they found the reviewers' suggested revisions to be reasonable. The survey from source 31 of 361 statisticians and psychologists found that 72% thought that the net effect of refereeing upon the quality of the article was to improve it. However, the abstract, quoted in the right column, has important qualifications:

> *"76% encountered pressure to conform to the strictly subjective preferences of the reviewers, 73% encountered false criticisms (and 8% made changes in the article to conform to reviewers' comments they knew to be wrong), 67% encountered inferior expertise, 60% encountered concentration upon trivia, 43% encountered treatment by referees as inferior, and 40% encountered careless reading by referees. At some time in their general experience with the peer-review system, 66% believed that referees' comments were contrived to impress the editor, 63% felt that the editor regarded their knowledge and opinion about the reported research as less important than that of the referees, 44% felt they were being treated like a supplicant, and 47% accepted a referee's suggestion against their better judgment."*

So, researchers, despite their problems with the peer review process, seem to think that it is in some way beneficial. But what does the actual evidence say?

The evidence I know of is mixed on whether or not there is even any real benefit. The paper, *Effects of Editorial Peer Review* [28], notes that source 29 was "the only identified study addressing the effects of peer review validity." This should be a rather eye-popping statement, the only one they found? This is not the only one that exists, but this does characterize the general amount of evidence that exists on the topic. Given the gravity of importance science gives this topic, you would expect not just that there's more evidence in existence, but that they'd be able to bunch it up into categories, compare validity between different fields, do regressions, etc. You would be mistaken, a faith based process rather than an evidence based one lies at the heart of science.

So what does source 29 say? It compared studies published in peer-reviewed journals to papers published in review-deficit ones from a sample of 123 studies about road safety. The studies were compared with a point system on the basis of whether or not they specified any moderating variables, whether or not they controlled for confounding variables, their overall study design, whether or not they specified how severe accidents or injuries were, mean sampling size, total sampling size, and sampling technique. The results are that peer-reviewed papers had larger sample sizes, but that is it.

Contradictory to these findings, source 32 has a sample of 111 manuscripts submitted to Annals of Internal Medicine. They went through the peer review process, and judges who weren't told which version of the paper was peer-reviewed were given the same paper before and after the peer review process and were told to rate which version was superior to the other on a 1-5 scale for 34 different aspects of quality. Peer-reviewed versions of papers were rated to be better on 33 out of the 34 measured aspects of quality.

To re-summarize source 17, 2 versions of a paper on the efficacy of a randomized control trial on a form of joint surgery were sent to 238 reviewers which were identical in everything except for the results; positive effect of surgery versus no effect. The reviewers were biased towards the versions with positive results. There were 7 intentionally planted errors in both papers. For the positive version, reviewers found an average of 0.41 errors; for the negative result, reviewers found an average of 0.85 errors. So on average, the reviewers found less than 1/7th of the intentionally planted errors even when reviewers disliked the paper's results and obsessively put it under a microscope to look for errors in order to try to reject it. Apart from the low error detection rate, of importance is the fact of subjectivity in error detection probably means that journals aren't using objective methods which means that they are able to vary between journals in their detection rates which could explain differences in results between journals, in which case, source 29 had more manuscripts which it sent to more journals than source 32 and is thus more generalizable. Alternatively, perhaps source 32 is the better one for having a more detailed evaluation of quality. The problem is that we don't know because there is barely any research.

Is a 1/7th error detection rate especially low? In 2009 the British Medical Journal engaged in an internal sting operation [19]. Fiona Godlee & colleagues sent out a paper to over 600 reviewers working for the British Medical Journal. The paper had 9 intentionally placed major errors, and 5 intentionally placed minor errors. The study looked at the training level for 3 groups of researchers, a group that wasn't given any training, a group that was given a packet of materials and told to self teach, and a group that was given face to face training. The control group on average found

2.74 of the 9 intentionally planted major errors compared to the self taught group finding 3.01, and the face to face group finding 3.12. The average of all groups in a single group was 2.96. In terms of finding the minor errors, the control group actually outperformed both of the trained groups. While presented as a study testing the efficacy of a reviewer training program, it is a de facto sting against the British Medical Journal. In 2014, Journal Citation reports gave the British Medical Journal an impact factor of 16.378, putting it at 4th place among all general medical journals in the world. This is a higher error detection rate than source 17, and I think that the very fact that they had the humility and integrity to engage in this sting operation at all is evidence that the British Medical Journal is probably better than average. Other journals, which don't even bother with this kind of self-testing, are probably even worse.

Overall, we can say that things lean in the direction that the peer review process removes at least some errors, but the evidence used to back this definitely should not make advocates of the peer review process jump for joy. Moreover, the evidence on how small journals stack up to the more prestigious journals overwhelmingly shows that the smaller journals are better because they don't have to deal with the conflicts of interest to the same degree [more here]. If the conflicts of interest also apply to papers published in review-deficit journals, it could actually be that such deleterious effects overpower the meager positive effects of the peer review system. Overall. any belief that confidence in the peer review process is supported by some kind of large body of evidence is clearly not justified.

SCIgen:

In 2005, a sting operation was done by MIT graduate students Jeremy Stribling, Dan Aguayo, and Maxwell Krohn. They wrote a program called SCIgen which generates fake academic papers. It works through methods similar to some of the text spinning algorithms that hackers use to bypass spam filters. In their sting, they submitted a paper to the 2005 World Multiconference on Systemics, Cybernetics and Informatics. That paper [18] was titled *Rooter: A Methodology for the Typical Unification of Access Points and Redundancy*.

The three authors were invited to speak at the conference, where they exposed the hoax. The program SCIgen is available on the internet free to download and use by anyone. By 2014, at least 16 SCIgen generated papers had been discovered to have been floating around in Springer Journals [1].

A funnier sting involves the publication of a Feminist rewrite of Mein Kampf [420 & 427]. Reviewers at the grievance journal *Affilia* peer-reviewing Feminist *Mein Kampf* [420], 1939, colourized:

According to source 2, SCIgen papers had an acceptance rate of 13.3% at the ACM digital library, and 28% for Institute of Electrical and Electronics Engineers. Now certainly the ACM digital library and the IEEE are not the most prestigious journals, but 16 got into Springer. I don't know what percentage of the SCIgen papers which were submitted to Springer were successful, but at least 16 of them were.

If completely bogus, nonsense-jargon-filled papers can get in at least some of time, what about papers which aren't so transparently awful? What about papers whose authors are smarter liars than a text-spinning algorithm? What about accidentally bad papers? This is the point. Nobody would say that the prestigious journals are churning out thousands of SCIgen-tier papers, but the fact that SCIgen papers are sometimes accepted calls into question the seriousness of the peer review process.

Another similar sting operation was done by John Bohannon in his Sciencemag article: *Who's afraid of Peer Review?* [3]. Bohannon wrote 304 papers (which were slightly different, but essentially the same) about a fictional moss that supposedly inhibits cancer growth. Among the errors were descriptions of a correlation between moss exposure and cancer inhibition when his charts and data showed zero correlation. The 304 slightly different papers were sent to 304 Journals. Bohannon sent the paper to 167 Directory of Open Access Journals (DOAJ), 121 to Jeffrey Beall's list, and 16 on both Beall's list and the DOAJ. Beall's list is a list of Journals determined by Jeffrey Beall to be bogus. Here are the results of his submissions:

Attribute	DOAJ	Beall's List	Overlap
Total Submissions	167	121	16
Total Responses	144	97	14
Rejected without Peer Review	64	3	3
Rejected with Peer Review	16	10	2
Accepted without Peer Review	29	47	6
Accepted with peer Review	35	37	3

Again, that junk journals reject junk articles less often is not interesting. What is interesting is that the article got into *Drug Intervention Today*, published by *Elsevier*, and one of Wolter Kulwer's Journals: The *Journal of Natural Pharmaceuticals*. Both went into damage control mode with Elsevier stating that they didn't own the Journal whose content they were publishing, and with Kulwer simply deleting the Journal. That they even cared and responded to the sting operation is a sign of integrity and thus quality, and there's no reason to believe that these journals were somehow worse than any others.

As an aside, the acceptance of the correlation mistake establishes something important: Sometimes an author and his data disagree, so when citing a paper, cite the author's data and results rather than the author's hopes and dreams.

On Replication:

Another way to examine the efficacy of peer review is to look at replication. One good thing we can do to judge the veracity of a given result is to look at whether or not other authors can look at a paper and use it to do a separate study using identical or similar methods to see whether or not they can achieve similar results. However, while replication is a good tool to judge the quality of a result, it is not necessarily a good one for judging the quality of a journal, researcher, field, or institution. While replication rate may plausibly happen to be a proxy for journal quality, journals should not be judged from the replication rates of their studies. This is because scientists don't tend to just do boring replication studies over and over again on the topics where they know what the result is going to be, scientists rather tend to push the boundaries of their field by doing experiments that hack away at whatever people disagree about. We could get a replication rate of 100% by churning out thousands of papers that test whether or not 2+2=4.

This stated, it is obviously a sign of journal quality and possibly integrity whether or not their papers provide the resources which are required to test their studies for replication (However, in some cases with human subjects, authors may be able to argue that publishing these resources would mean disclosing private information that they don't have permission to expose. Still, we should go as far as possible with open data with privacy being the exception rather than the rule).

One problem in science which is fortunately being counteracted by regulation on the funding agency level, the journal level, and the government level, is that oftentimes, individual researchers are the only ones who have access to datasets. Since individual authors aren't good at keeping track of them, old datasets often become completely inaccessible; the odds of getting access decline by 17% per year 3 years after initial publication [131]. Perhaps this could actually be another source of publication bias aside from author-side factors and journal-side factors. If individuals have such control, then perhaps they could deny researchers access to data if they know what the results of various tests will be and they don't want them published. However, I know of only one example of a researcher claiming this to have happened to them [132]. An alternative explanation could be that perhaps those in control of datasets want to hoard scientific discoveries for themselves and will prevent results from being published by others for this more benign reason.

In 2013, Melissa Haendel et. al. looked at 238 biomedical papers from 84 journals [23]. Of all of the studies, she found that the percent of studies with the identifiable resources that are necessary for replication to be as follows: Antibody: 44%, Cell Lines: 43%, Constructs: 25%, Knockdown reagents: 83%, Organisms: 77%. Only 5 of the journals analyzed had, by her definition, "stringent" resource reporting guidelines. In source 24, from 2008 to 2012, 389 researchers were asked how willing they would be to share protocols and raw data (the bones). In 2008, 80% of the respondents would be willing to share additional protocols beyond what was gone over in the methods section, but only 60% would be willing to share raw data. In 2012, Only 60% of researchers said they were willing to provide additional protocols, and only 45% said they would be willing to share raw data. Keep in mind that this is just a survey; it could be that this overestimates how many would actually share this information should push come to shove. Even if low replication rates were reasonable, we would still expect replication rates to be higher if researchers were at least capable of testing for the presence or lack of replicability. Giving fellow researchers access to data is increasingly important for the research community [24], and open data can help to detect fraud [1167].

According to a Nature poll of 1,576 researchers, over 70% have tried and failed to reproduce another scientist's experiments, and more than half have failed to reproduce their own experiments [124]. Despite the overwhelming majority saying that there is a crisis in reproducibility, most still say that they trust the published literature. Source 125 tried to replicate 100 psychology experiments, and 47% of replications had the same findings as the original studies.

Some look on in horror at the roughly 1 in 2 chance that a novel finding is actually correct, however it is not immediately obvious that a

50 - 50 chance should be the benchmark coinflip of comparison because it is not immediately obvious that a randomly formulated hypothesis from a text spinning algorithm would have a 1 in 2 chance of being correct. There are a great many different kinds or results that can theoretically be obtained from an experiment, with studies reporting the results that they happen to find. The random result wouldn't be 1 out of 2, it would be 1 out of however many plausible results there are, and 1 in 2 are much better odds than that. However, these replication odds are indeed pathetic when put into the context of the extreme excess of positive results [more here]; if the extreme glut of positive results is due to researchers choosing the hypotheses which are likely to be correct, then it doesn't seem like replication rates should be this low.

My —incredulity at the question of why replicability is low— has actually been demonstrably unwarranted. There is actually good reason to believe that low replication rates are predominantly due to bad research practices rather than hypothesis selection. Maximum expected replicability (So-called "Maximum expected replicability" was not 100% replicability, but the ~86% replicability which should be predicted from statistical power and effect size) is achievable if the good research practices of high statistical power, preregistration, and full methodological transparency, are carried out [1166]. Moreover, the 2015 psychology replication study from earlier [125] 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, a 2016 paper on the replication rate of economics [126] found a replication rate of 88% for findings with an initial p-value of less than .001. Source 287 found that replication could be predicted by effect size and study design. Using p-values and other such similar clues, multiple papers have found that researchers are correctly able to predict which of a set of previous findings will successfully replicate the strong majority of the time [129 & 130]. Thus, if we consume research intelligently, we don't have to worry so much about buying into false-positive results.

Replicability By Field:

It's important to realize that these replication trends have nothing to do with psychology. Source 126 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, source 127 found that cancer research replicated only 11% of the time. Even worse, in Neuroscience, an attempt at replicating 17 brain imaging studies [128] replicated zero of them. Assuming a theoretical 18th would have replicated, this would seem to imply that at most, Neuroscience papers replicate 5.5% of the time. I am unaware of any attempts to replicate the physical sciences, but the Nature poll from earlier [124] broke down the survey's results by field. Just averaging the results within fields, in no field does the average researcher expect results to replicate more than 75% of the time. Below is a summary table for replication results by field; the physical sciences from the Nature poll are marked as estimated on the bottom half of the graph:

Discipline:	**Tested Replication Rate:**	**Citation:**
Differential Psychology	87%	Source 287
Experimental Philosophy	70%	Source 144
Economics	61%	Source 126
Cognitive Psychology	50%	Source 125
Social Psychology	25%	Source 125
Pharmacology	21%	Source 139
Oncology (cancer)	11%	Source 127
Neuroscience	No Successful Replications	Source 128
Discipline:	**Self Reported Expectations For Replication Rate:**	**Citation:**
Physics	73%	Source 124
Chemistry	65%	Source 124
Astronomy	65%	Source 124
Material Science	60%	Source 124
Biology	59%	Source 124
Earth and Environmental Science	58%	Source 124
Engineering	55%	Source 124
Medicine	55%	Source 124
Other	52%	Source 124

For comparison, here are figures for statistical power by field:

Discipline:	Mean / Median Statistical Power:	Citation:
Neuroscience	21%	Source 156
Brain Imaging	8%	Source 156
Social and Behavioral Sciences	24%	Source 155
Cognitive Neuroscience	14%	Source 154
Psychology	23%	Source 154
Medical Research	23%	Source 154
Breast Cancer	16%	Source 152
Glaucoma	11%	Source 152
Rheumatoid Arthritis	19%	Source 152
Alzheimer's Disease	9%	Source 152
Epilepsy	24%	Source 152
Multiple Sclerosis	24%	Source 152
Parkinson's Disease	27%	Source 152
Education	23%	Source 153
Intelligence	49%	Source 14
Intelligence - Group Differences	57%	Source 14

Notes on table creation: Source 14 is the 2018 preprint which is, frankly, superior to the published version. Power to detect median effect was used wherever possible. In some mega-analyses, power to detect median effect was not reported; in these, median effects were small, so power to detect small effects was used.

References:

No man is an island; researchers need to cite the work of other researchers. Human knowledge is not the result of the analyses of any single researcher or of any single research paper, but a result of the accumulation of knowledge throughout the existence of humanity. Researchers may use mathematical formulas or statistics which were formulated by others without themselves actually proving or understanding their veracity. It is also often unnecessary to design analyses which are sophisticated enough to relax commonly held assumptions if there is a wealth of external research literature demonstrating the veracity and robustness of said assumptions. It should be no surprise that scientific literature often contains an enormous number of references to other works. Presumably, a cited premise should be substantiated by the given reference(s). When a given citation fails to substantiate the claim for which it is marshalled, a 'quotation error' has occurred. Quotation errors are a threat to the progress of research because they can result in the propagation of unverified or incorrect information. While necessary to do, it is a time-consuming hassle to read so many works, and so researchers often just copy the reference information from a second-hand source. The problem comes when a long chain of researchers copy references from each other. Eventually, it turns into a game of telephone where misrepresentations creep in, even if every person in the chain was acting in good faith. Citation lineages can sometimes be measured objectively like copyright traps when a citation formatting error is made by an earlier author, and all following authors precisely copy the same formatting error, such as the reproduction of Gould's idiosyncratic reference error [150, p. 135].

How much of a problem are quotation errors? Source 1168 reviewed evidence from 23 previous papers on the topic, and although great heterogeneity in operationalizations was observed, it was concluded that regardless, "quotation errors were found in significant numbers" in "all previous studies surveyed". Its review included the fields of ecology, marine biology, physical geography, and various social sciences. The paper itself also examined 250 random citations, and found a misrepresentation rate of 25%. This doesn't even include the rate at which reference identification information is written currently, and it obviously does not mean that the 75% majority is accurate due to diligence rather than due to luck. Nevermind examination for evidence quality rather than paper opinion. In one of the papers cited [37], it was found that

3% of references were recorded so poorly that the original source could not even be located and inspected. Even in medical journals, which should presumably be dealing with one of the harder sciences, one paper found an error rate of 48% [33]. Interestingly, when multiple references are marshalled in support of the same statement, they are more likely to be represented accurately [1168]; "string citations" despite making up 63% of all citations, account for only 34% of errors.

Even reviewers' ability to detect transgressions as major as plagiarism seems weak. One paper [16] conducted an experiment where two intentionally methodologically flawed modifications were made to a previously published paper and sent to journals in sociology, psychology, counseling, medicine, and social work. Only two of the 110 journals to which it was sent said that the paper had already been published. This occurred despite the fact that the original paper had been cited frequently. Although the study's control group had been omitted from the original paper, few reviewers mentioned this as a problem. The paper concludes that only six of the 33 received reviews were competently done.

Interpretations of results are often also skewed by certain types of people. For example, primary study authors of significant studies are more likely than methodologists to believe that a strong association exists in a heterogeneous meta-analysis [1169].

On Academic Experts:

Given the general lack of transfer effects for the applicability of knowledge [see Chapter 3], the high rate at which students forget the material which they are taught [1189, p.40], and the general irrelevance of the material which people learn in school [1189, Ch.2], reasonable priors dictate that we should not have have high expectations for the quality of academic expertise. Moreover, if scientific progress is to be taken as a positive externality which should accelerate economic growth, this does not fit well with the broader picture showing national educational attainment to be unrelated to national wealth/growth [more here]. These things aside, there are two criteria with which we can judge the observed quality of expertise against the fundamental skills scientists are supposed to have in order to do competent research [more here]: statistical literacy, and predictive accuracy.

To recap, statistical literacy is essential for properly interpreting the patterns we observe, and even poor theorists who overfit at the expense of elegance should be able to make better predictions than laymen about their fields of expertise.

Unfortunately, experts are breathtakingly statistically illiterate [more here], and they make predictions that are about as good as the layman is often equipped to make [more here].

Statistical Literacy:

Numerous studies have shown that the vast majority of academics working in psychology, epidemiology, and even the hard sciences don't understand basic statistical concepts like p-values, confidence intervals, and t tests. In addition, they fail simple applied questions as well:

Source 288:

In this sample of 759 Professors and students, more than 85% of students and professors from the following fields endorsed at least one misinterpretation of p-values [more here] and/or confidence intervals: science, engineering, medicine, math/statistics, management, psychology, economics.

Source 289:

When given a quiz concerning common statistical issues dealt with in psychological research, a sample of 551 psychologists on average answered 55% of questions correctly.

Source 290:

At least one of six misinterpretations of confidence intervals were endorsed by 97% of a sample of 118 psychology researchers.

Source 291:

In a sample of 113 Psychology professors and students, at least 1/6 misinterpretations of a t-test were endorsed by 80% of psychologists teaching statistics (mean = 1.9), 89.7% of psychologists not teaching statistics (mean = 2.0), and 100% of psychology students (mean = 2.5).

Source 292:

When 261 Epidemiologists were told about an intervention in which the rate of disease recovery was higher of those taking Drug A than for those taking Drug B, 79% of epidemiologists denied that a person was probably more likely to recover if assigned drug A rather than drug B when the p-value of the difference between the recovery rates exceeded .05. A p-value says based on the effect size and the statistical power how likely the result would be to come about by random chance from sampling error if there were really no effect.

Common errors in the interpretation of statistical significance come from the name statistical significance. If an incredibly small sample doesn't even have the statistical power to detect a large effect, then the effect of an incredibly important variable would fail to achieve statistical significance. Similarly, if you have a sample of 5 million people but the effect size is so small that you just barely get the p-value below the 0.05 standard, such an effect may exist, but it's clearly a lot less important than the effect in the previous example.

Source 293:

When told about a cancer intervention in which group A lived longer than group B, roughly 50% of the sample of 117 Statisticians denied that, "speaking only of the subjects who took part in this particular study", participants in group A lived longer than participants in group B when the p-value of the difference exceeded .05. In a sample of 140 Statisticians, when told about an intervention in which the rate of disease recovery was higher of those taking Drug A than for those taking Drug B, 84% of them denied that a person was probably more likely to recover if assigned drug A rather than drug B when the p-value of the difference between the recovery rates exceeded .05. People don't think about what the statistics actually mean, they just think about the blunt name: "statistical significance".

Source 294:

This paper had a sample of 25 private sector statisticians and 20 psychologists. In a drug trial resulting in a large effect size but an insignificant p-value, 52% of statisticians and 65% of psychologists thought no conclusion could be drawn about the drug's efficacy, 36% of statisticians and 35% of psychologists thought the drug was ineffective, and 12% of statisticians and 0% of psychologists thought the drug was effective.

Similarly, reviews of papers published in medical journals typically find that the majority of papers commit statistical errors than render them methodologically unacceptable [295 - table 1]:

Source 295 - Table 1:

Table I. Summary of some reviews of the quality of statistics in medical journals, showing the percentage of 'acceptable' papers (of those using statistics)

Year published	First author	Number of papers	Number of Journals	% papers acceptable
1966	Schor[4]	295	10	28
1977	Gore[5]	77	1	48
1979	White[6]	139	1	55
1980	Glantz[7]	79	2	39
1982	Felson[8]	74	1	34
1982	MacArthur[9]	114	1	28
1983	Tyson[10]	86	4	10
1985	Avram[11]	243	2	15
1985	Thorn[12]	120	4	< 40
1988	Murray[13]	28	1	61
1988	Morris[14]	103	1	34
1995	McGuigan[15]	164	1	60
1996	Welch[16]	145	1	30

Predictive Accuracy:

The general research literature does not broadly paint the picture that academic expertise ensures an impressive degree of predictive accuracy:

Source 71:

This paper looked at 137 studies comparing clinical predictions to mechanical predictions. The norm is that statistical prediction rules outperform expert judgements just about everywhere that this comparison has been made [Source 71 - Table 1].

Source 71 - Table 1:

Table 1
Studies Included in Meta-Analysis

Citation	Predictand	Accuracy statistic	Accuracy: Clinical	Accuracy: Mechanical
Alexakos (1966)	college academic performance	HR	39	56
Armitage & Pearl (1957)	psychiatric diagnosis	HR	30	31
Ashton (1984)	magazine advertising sales	corr	0.63	0.88
Barron (1953)	psychotherapy outcome	HR	62	73
Blattberg & Hoch (1988)	catalog sales; coupon redemption	corr	0.52	0.66
Blenkner (1954)	case work outcome	corr	0.00	0.62
Bobbitt & Newman (1944)	success in military training	regression coefficient	0.93	0.87
Bolton et al. (1968)	vocational rehabilitation outcome	corr	0.30	0.40
Boom (1986)	diagnosis of jaundice	HR	85	90
Boom et al. (1988)	diagnosis of jaundice	HR	88	96
Boyle et al. (1966)	diagnosis of thyroid disorder	HR	77	85
Brodman et al. (1959)	general medical diagnosis	HR	43	48
Brown et al. (1989)	diagnosis of lateralized cerebral dysfunction	corr	0.43	0.64
Buss et al. (1955)	prediction of anxiety	corr	0.60	0.64
Caceres & Hochberg (1970)	diagnosis of heart disease	HR	74	84
Campbell et al. (1962)	job performance	corr	0.15	0.29
Cannon & Gardner (1980)	general medical diagnoses, optimality of treatment recommendations	HR	63	64
Cebul & Poses (1986)	presence of throat infection	HR	69	99
Clarke (1985)	surgery recommendation	HR	59	69
Cooke (1967)	psychological disturbance	HR	77	76
Cornelius & Lyness (1980)	job analysis	corr	0.73	0.76
Danet (1965)	future psychiatric illness	HR	65	70
Dannenberg et al. (1979)	prognosis of medical illness	accuracy coefficient	0.22	0.21
Dawes (1971)	success in graduate school	corr	0.10	0.51
De Dombal et al. (1974)	diagnosis of gastrointestinal disorders	HR	71	92
De Dombal et al. (1975)	diagnosis of gastrointestinal disorders	HR	83	85
De Dombal, Horrocks, et al. (1972)	diagnosis of gastrointestinal disorders	HR	50	97
De Dombal, Leaper, et al. (1972)	diagnosis of appendicitis	HR	83	92
Devries & Shneidman (1967)	course of psychiatric symptoms	HR	75	100
Dicken & Black (1965)	supervisory potential	corr	0.09	0.30
Dickerson (1958)	client compliance with counseling plan	HR	57	52
Dickson et al. (1985)	diagnosis of abdominal pain	HR	55	73
Dunham & Meltzer (1946)	length of psychiatric hospitalization	HR	34	70
Dunnette et al. (1960)	job turnover	HR	53	73
Durbridge (1984)	diagnosis of hepatic or biliary disorder	HR	62	74
Edwards & Berry (1974)	psychiatric diagnosis	HR	63	74
Enenkel & Spiel (1976)	diagnosis of myocardial infarction	HR	78	57
Evenson et al. (1973)	medication prescribed	HR	77	75
Evenson et al. (1975)	length of hospitalization	HR	76	71
Geddes et al. (1978)	degree of pulmonary obstruction	HR	96	95
Glaser & Hangren (1958)	probation success	HR	83	84
Glaser (1955)	criminal recidivism	mean cost rating	0.14	0.35
S. C. Goldberg & Mattsson (1967)	improvement of schizophrenia	significance test	8.15	10.78
L. R. Goldberg (1965)	psychiatric diagnosis	corr	0.28	0.38
L. R. Goldberg (1969)	psychiatric diagnosis	HR	62	69
L. R. Goldberg (1976)	business failure	corr	0.51	0.56
Goldman et al. (1981)	cardiac disease survival or remission	corr	−0.12	−0.11
Goldman et al. (1982)	diagnosis of acute chest pain	HR	79	73
Goldman et al. (1988)	prediction of myocardial infarction	HR	73	76
Goldstein et al. (1973)	cerebral impairment	HR	95	75
Gottesman (1963)	personality description	HR	62	53
Grebstein (1963)	prediction of IQ	corr	0.59	0.56
Gustafson et al. (1973)	diagnosis of thyroid disorder	HR	88	87
Gustafson et al. (1977)	suicide attempt	HR	63	81
Halbower (1955)	personality description	corr	0.42	0.64
Hall (1988)	criminal behavior	HR	54	83
Hall et al. (1971)	diagnosis of rheumatic heart disease	HR	62	73
Harris (1963)	game outcomes and point spread	HR	60	69
Hess & Brown (1977)	academic performance	HR	68	83
Holland et al. (1983)	criminal recidivism	corr	0.32	0.34
Hopkins et al. (1980)	surgical outcomes	HR	84	91
Hovey & Stauffacher (1953)	personality characteristics	HR	74	63
Ikonen et al. (1983)	diagnosis of abdominal pain	HR	67	59
Janzen & Coe (1973)	"diagnosis" of female homosexuality	HR	57	85
Jeans & Morris (1976)	diagnosis of small bowel disease	HR	83	83
Johnston & McNeal (1967)	length of psychiatric hospitalization	HR	72	75
Joswig et al. (1985)	diagnosis of recurrent chest pain	HR	69	86
Kahn et al. (1988)	detection of malingering	HR	21	25
Kaplan (1962)	psychotherapy outcome	HR	66	70
Kelly & Fiske (1950)	success on psychology internship	corr	0.32	0.41
Khan (1986)	business startup success	corr	−0.09	0.13
Klehr (1949)	psychiatric diagnosis	HR	67	64
Klein et al. (1973)	psychopharmacologic treatment outcome	corr	0.12	0.90
Kleinmuntz (1963)	maladjustment	HR	70	72
Kleinmuntz (1967)	maladjustment	HR	68	75
Klinger & Roth (1965)	diagnosis of schizophrenia	HR	77	43
Kunce & Cope (1971)	job success	HR	67	77
Lee et al. (1986)	death and myocardial infarction	corr	0.58	0.64
Leli & Filskov (1981)	presence, chronicity and lateralization of cerebral impairment	HR	79	79
Leli & Filskov (1984)	diagnosis of intellectual deterioration	HR	75	73
Lemerond (1977)	suicide	HR	50	50
Lewis & MacKinney (1961)	career satisfaction	corr	0.09	0.56
Libby (1976)	business failure	HR	74	72
Lindzey (1965)	"diagnosis" of homosexuality	HR	70	57
Lindzey et al. (1958)	"diagnosis" of homosexuality	HR	95	85
Lyle & Quast (1976)	diagnosis of Huntington disease	HR	61	68
Martin et al. (1960)	diagnosis of jaundice	HR	87	79
Mathew et al. (1988)	diagnosis of low back pain	HR	74	87
McClish & Powell (1989)	intensive care unit mortality	ROC	0.89	0.83
Miller et al. (1982)	general medical diagnosis	HR	53	40
Mitchell (1975)	managerial success	corr	0.19	0.46
Oddie et al. (1974)	diagnosis of thyroid disorder	HR	97	99
Orient et al. (1985)	diagnosis of abdominal pain	HR	64	63
Oskamp (1962)	presence of psychiatric symptoms	HR	70	71
Peck & Parsons (1956)	work productivity	corr	0.71	0.61
Pierson (1958)	college success	HR	43	49
Pipberger et al. (1975)	diagnosis of cardiac disease	HR	72	91
Plag & Weybreun (1968)	fitness for military service	corr	0.19	0.30
Popovics (1983)	cerebral dysfunction	corr	0.17	0.16
Poretsky et al. (1985)	diagnosis of myocardial infarction	HR	80	67
Reale et al. (1968)	diagnosis of congenital heart disease	HR	73	82
Reich et al. (1977)	diagnosis of hematologic disorders	HR	68	71
Reitan et al. (1964)	diagnosis of cerebral lesions	HR	75	73
Rosen & Van Horn (1961)	academic performance	HR	55	57
Royce & Weiss (1975)	marital satisfaction	corr	0.40	0.58
Sacks (1977)	criminal recidivism	HR	72	78
Sarbin (1942)	academic performance	corr	0.35	0.45
Schiedt (1936)	parole success or failure	HR	68	76
Schofield & Garrard (1975)	performance in medical school	HR	76	78
Schofield (1970)	performance in medical school	deviation score	0.07	−0.06
Schreck et al. (1986)	diagnosis of acid-base disorders	HR	55	100
Schwartz et al. (1976)	diagnosis of metabolic illnesses	HR	92	85
Shapiro (1977)	outcome of rheumatic illness	Q	0.20	0.15
Silverman & Silverman (1962)	diagnosis of schizophrenia	HR	55	64
Smith & Lanyon (1968)	juvenile criminal recidivism	HR	52	54
Speigelhalter & Knill-Jones (1984)	diagnosis of dyspepsia	ROC	0.85	0.83
Stephens (1970)	schizophrenia prognosis and course	corr	0.51	0.29
Storment & Finney (1953)	assaultive behavior	corr	0.00	0.57
Sutton (1989)	diagnosis of abdominal pain	HR	65	57
Szucko & Kleinmuntz (1981)	lie detection	corr	0.23	0.42
Taulbee & Sisson (1957)	psychiatric diagnosis	HR	63	63
Thompson (1952)	juvenile delinquency	HR	64	91
Truesdell & Bath (1957)	academic dropouts	HR	71	75
Ullman (1958)	course of group home placement	HR	59	78
Walters et al. (1988)	malingering	HR	56	93

Source 71 - Table 1 (continued):

Table 1 (*continued*)

Citation	Predictand	Accuracy statistic	Accuracy: Clinical	Accuracy: Mechanical
Warner (1964)	diagnosis of congenital heart disease	HR	66	66
Watley & Vance (1963)	college achievement and leadership	HR	59	72
Webb et al. (1975)	occupational choice	HR	35	55
Wedding (1983)	diagnosis of cerebral impairment	corr	0.74	0.84
Weinberg (1957)	personality characteristics	corr	0.41	0.65
Werner et al. (1984)	assault by psychiatric inpatients	corr	0.14	0.56
Wexler et al. (1975)	medical diagnosis	HR	65	85
Wiggins & Kohen (1971)	graduate school success	corr	0.33	0.58
Wilkinson & Markus (1989)	minor psychiatric morbidity	ROC	0.74	0.89
Wittman & Steinberg (1944)	psychiatric prognosis	HR	41	68
Wormith & Goldstone (1984)	criminal recidivism	corr	0.21	0.39
Yu et al. (1979)	optimality of treatment for meningitis	HR	30	65

Note. For Accuracy Statistic, HR = hit rate (nearest %), corr = correlation coefficient (generally Pearson), ROC = area under Receiver Operating Characteristic curve.

Activate Windows

The paper also found that:

> *"Similarly, training and experience (amount of training, general experience in the field, specific task-relevant experience) do not significantly predict the degree of superiority of mechanical over clinical prediction."*

and that:

> *"When results of an interview are used as predictive data, the ES favors the mechanical prediction more than when no interview is available [with interview, weighted M ± SD = 0.224 ± 5.06; without interview, 0.070 ± 2.29, ((134) = 5.02, p < .0001]."*

So when clinicians were given an interview of the subject, their predictions become worse because the interview is introduced to all sorts of extraneous clues which aren't statistically validated.

Another interesting result:

> *"Use of medical data (physical examination, laboratory tests) as predictors is associated with smaller differences [with medical data, weighted M ± SD = 0.083 ± 3.00; without medical data, 0.16 ± 3.61, ((134) = 2.66, p < .009]"*

So when experts were given medical data, their predictions improved, but the expert predictions when given medical data were still inferior to SPRs that did not have access to the same data.

There is also reason to believe that SPRs would beat experts even more severely in modern day than they did back then:

1. Increased computer hardware power
2. More refined statistical algorithms; more data is available, more algorithms have gone under more testing over time, and computers never forget unless somebody forgets to make a backup or something.

Source 1170:

This paper examined three experienced pathologists (and a fourth judge which was the average of the three) who assessed the severity of cancer in 193 patients based on 5 point scales of various symptoms that they deemed important. Severity, if accurately assessed, should significantly negatively correlate with survival time, but this was not true for any of the pathologists. In fact, the severity rating of the average of the three doctors (judge 4) had a non-significant and positive correlation with survival time:

Source 1170 - Table 1:

TABLE 1
Correlation of Global Judgment with Survival Time

Judge	Global-survival time: r	Global-survival time: r^2	Global-log survival time: r	Global-log survival time: r^2
1	−.002	.000	−.038	.002
2	.116	.012	.098	.010
3	−.139	.019	−.127	.016
4	.143	.020	.072	.005

Note: All r's based on n = 193; r needed for significance at $p < .01$ is .179.

Moreover, using the same symptom ratings as the doctors, a computer algorithm was able to significantly predict survival rates. This implies that the doctors had useful information available to them, but combined and weighed that information in such a way that they failed to utilize any of its predictive validity:

Source 1170 - Table 3:

TABLE 3
CROSS-VALIDATION FOR THREE MODELS FOR PREDICTING SURVIVAL TIME USING COMPONENTS WITH AND WITHOUT THE GLOBAL JUDGMENT

	Components alone						Components + global judgment					
	Linear		Conj.		Disj.		Linear		Conj.		Disj.	
Judges	R	R^2	R	R^2	R	R^2	R	R^2	R	R^2	R	R^2
1	.258*	.066	.320*	.103	.380*	.145	.150	.023	.292*	.086	.359*	.129
2	.220	.048	.171	.029	.293*	.086	.331*	.111	.250	.063	.340*	.116
3	.201	.040	.195	.038	.229	.052	.149	.022	.135	.018	.191	.036
4	.249	.062	.273*	.074	.350*	.122	.295*	.087	.278*	.077	.377*	.142

Note: The original sample was $n = 100$ and the cross-validated sample was $n = 93$.
* $p < .01$.

Even when using all the judges ratings at once, they added little to the predictive validity of the model:

Source 1170 - Table 5:

TABLE 5
RESULTS FOR MULTIPLE JUDGES ON INITIAL FIT AND CROSS-VALIDATION

	Components						Components + global judgment					
	Linear		Conj.		Disj.		Linear		Conj.		Disj.	
	R	R^2	R	R^2	R	R^2	R	R^2	R	R^2	R	R^2
Initial fit	.453	.205	.510	.260	.549	.301	.521	.271	.510	.260	.560	.314
Cross-validation	.202	.041	.210	.044	.363*	.132	.287	.083	.180	.032	.396*	.157

Note: Initial fit on $n = 193$. Cross-validation done on $n = 100$ for initial fit and $n = 93$ for cross-validated sample.
* $p < .01$ for cross-validated sample.

Source 1171:

This paper had 9 physicians estimate the probability of pneumonia developing in 1,531 patients. The main result suggests that the doctors were only marginally more accurate than guessing at random would have been:

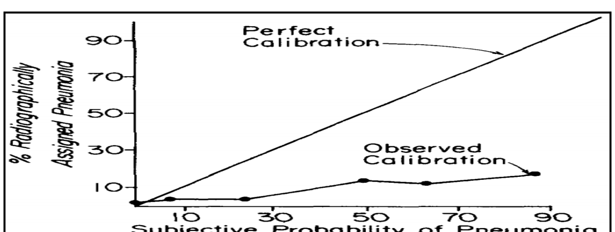


Figure 1. Relationship between physicians' subjective probability of pneumonia and the actual probability of pneumonia.

Source 1172:

This paper had doctors predict the probability that patients with heart disease would survive over the next one and three year periods. Doctors assigned a roughly equal probability to patients who ended up living and those who ended up dying:

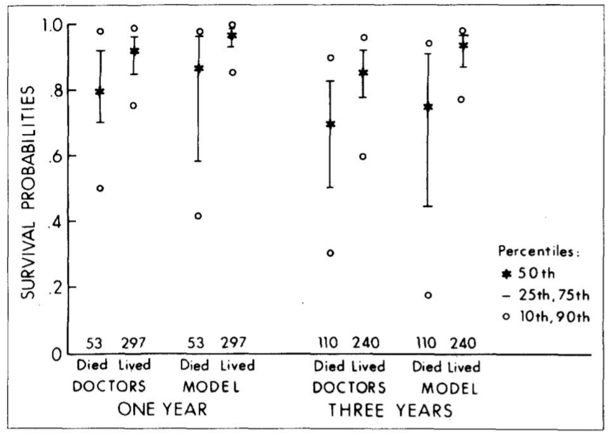


Figure 1. Selected percentiles of the distributions of one- and three-year survival probabilities predicted by the doctors and the model. At one year, 53 of the 350 patients (15 percent) had actually died, and at three years, 110 of the 350 patients (31 percent) had died.

Thus, Doctors seem pretty bad at predicting things like whether you have a disease, how severe your disease is, and whether you will live for the next few years given your disease.

Source 1173:

Turning to economics and finance, this paper analyzed the returns to stocks after sorting them by long term growth forecasts given by financial analysts. The highest returning stocks were those in the bottom 10% of projected growth while the weakest returns were seen among stocks in the top 10% of expected growth, suggesting that one could make significant gains by treating financial experts as sort of ‘anti-experts’:

Source 1173 - Figure 1:

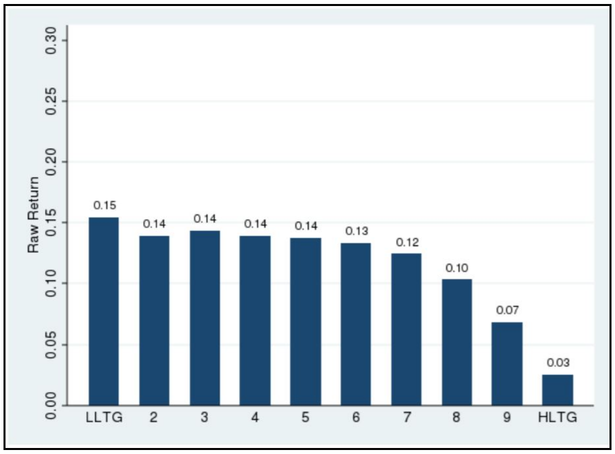
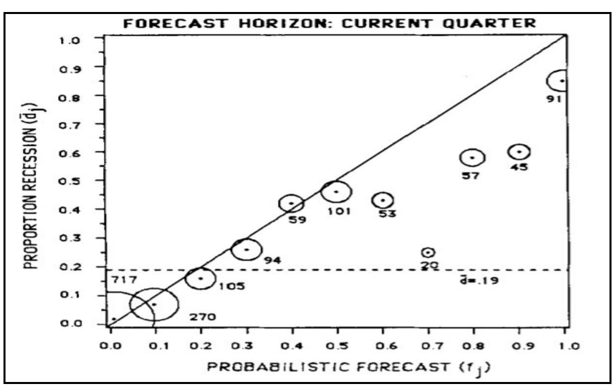


Figure 1. Annual Returns for Portfolios Formed on LTG. In December of each year between 1981 and 2015, we form decile portfolios based on ranked analysts' expected growth in earnings per share and report the geometric average one-year return over the subsequent calendar year for equally-weighted portfolios with monthly rebalancing.

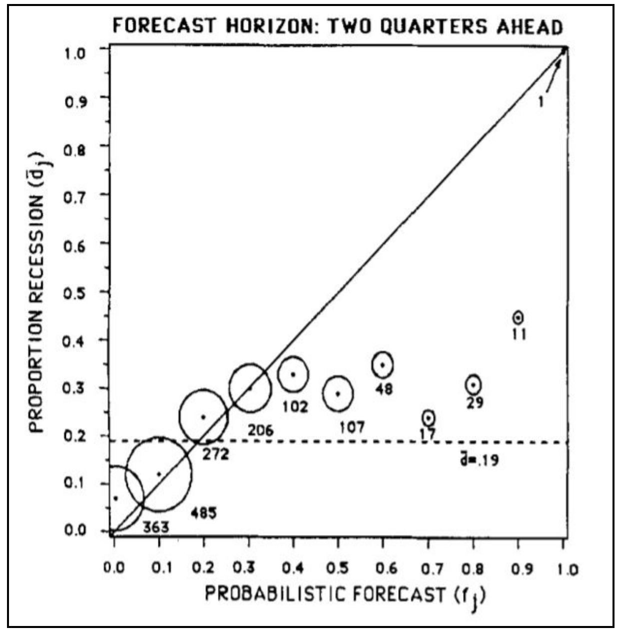
Source 1174:

This paper reported on 40 professional economic forecasters who were surveyed yearly from 1968 to 1988. They could sort of predict recessions that were just about to happen, but if the period was more than a couple months, their predictive accuracy quickly fell to something similar to what we'd expect if they were guessing randomly:

Source 1174 - Exhibit 1:



Source 1174 - Exhibit 1 (Continued):



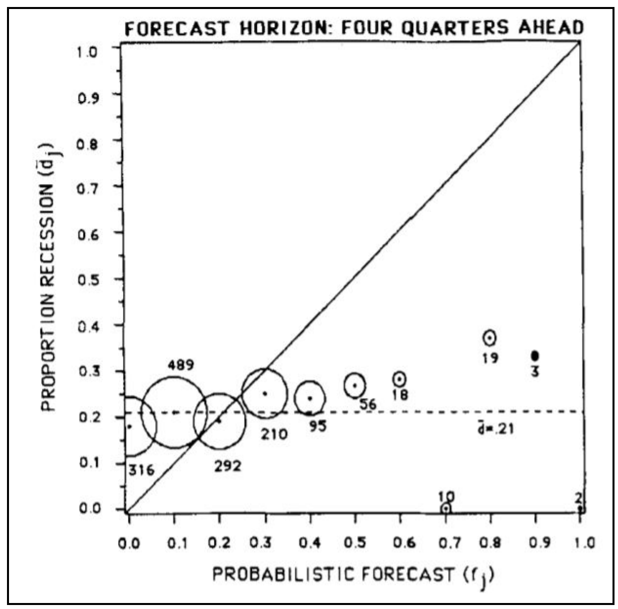


Exhibit 1. Calibration plots of the pooled forecast data for each forecast horizon (Q0 to Q4). The numbers inside the plots represent the frequencies of the forecast categories. Also, the sizes of the bubbles are proportional to these frequencies. The horizontal line in each frame represents the base rate (*d*) of recession; it equals the mean of the outcome variable *d* for the data in each forecast horizon.

Source 1175:

This paper compared the ability of experts (behavioral economists and relevant psychologists) and non-experts to predict the results of behavioral experiments aimed at changing the degree of effort people put into various tasks:

Source [1175](#) - Table 3:

Table 3. Accuracy of Forecasts by Group of Forecasters versus Random Guesses

	Average Accuracy (and s.d.) of *Individual* Forecasts	Accuracy of *Mean* Forecast (Wisdom of Crowds)	% Forecasters Doing Better Than Mean Forecast	Wisdom of Crowds: Accuracy Using Average of Simulated Group of Forecasters, Mean (and s.d.)	
				Group of 5	Group of 20
	(1)	(2)	(3)	(4)	(5)
Panel A. Mean Absolute Error					
Groups					
Academic Experts (N=208)	169.42 (56.11)	93.48	4.33	113.98 (23.15)	98.80 (11.68)
PhD Students (N=147)	171.42 (76.05)	91.65	8.16	117.99 (31.07)	97.78 (14.43)
Undergraduates (N=158)	187.84 (85.97)	87.86	3.16	115.46 (35.30)	94.80 (17.80)
MBA Students (N=160)	198.17 (86.04)	100.72	8.11	129.31 (34.34)	110.65 (17.05)
Mturk Workers (N=762)	271.57 (144.81)	146.93	17.85	173.01 (68.21)	150.93 (39.57)
Benchmark for Comparison					
Random Guess in 1000-2500	415.99				
Random Guess in 1500-2200	224.63				
Panel B. Mean Squared Error					
Groups					
Academic Experts (N=208)	49822 (34087)	12606	2.88	20046 (7894)	14438 (3234)
PhD Students (N=147)	53081 (50081)	11980	6.12	21365 (11268)	13895 (4142)
Undergraduates (N=158)	60271 (61112)	9769	2.53	19883 (12267)	12336 (4645)
MBA Students (N=160)	69855 (63213)	13334	3.90	24676 (12661)	16156 (4781)
Mturk Workers (N=762)	128801 (130473)	23660	9.71	44747 (32929)	28931 (13868)
Benchmark for Comparison					
Random Guess in 1000-2500	249534				
Random Guess in 1500-2200	75423				
Panel C. Rank-Order Correlation Between Actual Effort and Forecasts					
Groups					
Academic Experts (N=208)	0.42 (0.32)	0.83	4.81	0.65 (0.18)	0.76 (0.09)
PhD Students (N=147)	0.48 (0.30)	0.86	6.80	0.70 (0.18)	0.80 (0.09)
Undergraduates (N=158)	0.45 (0.31)	0.87	5.06	0.69 (0.17)	0.80 (0.09)
MBA Students (N=160)	0.37 (0.33)	0.71	18.52	0.56 (0.21)	0.67 (0.11)
Mturk Workers (N=762)	0.42 (0.35)	0.95	0.26	0.69 (0.20)	0.87 (0.07)
Benchmark for Comparison					
Random Guess in 1000-2500	0.00				
Random Guess in 1500-2200	0.00				
Panel D. Correlation Between Actual Effort and Forecasts					
Groups					
Academic Experts (N=208)	0.45 (0.29)	0.77	9.41	0.64 (0.16)	0.73 (0.09)
PhD Students (N=147)	0.51 (0.28)	0.86	4.86	0.72 (0.15)	0.82 (0.07)
Undergraduates (N=158)	0.49 (0.30)	0.89	3.90	0.72 (0.16)	0.84 (0.07)
MBA Students (N=160)	0.42 (0.32)	0.77	15.11	0.62 (0.19)	0.72 (0.09)
Mturk Workers (N=762)	0.43 (0.35)	0.95	0.00	0.70 (0.19)	0.88 (0.06)
Benchmark for Comparison					
Random Guess in 1000-2500	0.00				
Random Guess in 1500-2200	0.00				

When accuracy was measured as mean absolute error, the ranking of accuracy was experts > phd students > undergrad students > MBA students > MTurk Workers when considering individual forecasts. The differences between experts and students was small. When considering group forecasts, the ranking of accuracy was undergrad students > phd students > academic experts > MBA students > Mturk workers. When accuracy was measured as the correlation between predicted and observed effort rather than mean absolute error, the ranking was phd students > undergrads > experts > Mturk workers > MBA students when considering individual forecasts and Mturk workers > undergrads > phd students > experts = MBA students.

In no case was the rank order of prediction what we would predict if we assumed academia teaches people knowledge that increases their understanding of the real world.

Source [1176](#):

Turning to lawyers, this paper found that a sample of legal experts was only able to predict the results of supreme court cases at a rate modestly better than chance. Computer models were far more accurate:

Source [1176](#) - Table 1:

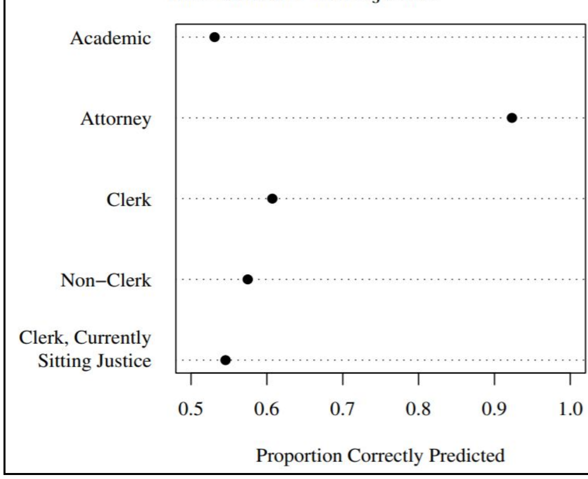
Table 1: Machine and expert forecasts of case outcomes for decided cases (N=68). Row percentages are in parentheses. The estimated (conditional maximum likelihood) odds ratio is 2.073 (p=0.025, Fisher's Exact Test).

	Case Outcome Forecast		
	Correct	Incorrect	Total
Machine	51 (75.0%)	17 (25.0%)	68 (100.0%)
Experts	101 (59.1%)	70 (40.9%)	171 (100.0%)

Moreover, the accuracy of these legal experts was largely driven by private attorneys. Academics only had an accuracy rate of 53%, scarcely better than random chance:

Source [1176](#) - Table 1:

Figure 5: Proportion correct expert forecasts of case outcomes by expert background. The figure is based on the following forecasts: 145 forecasts by academics; 26 by practicing attorneys; 84 by experts who clerked for the Supreme Court; 87 by non-Supreme Court clerks; and 34 forecasts by experts who clerked for a currently sitting Justice.



Source [1177](#):

With respect to psychologists, it’s been shown that only 16% of developmental psychologists were able to correctly predict that self control had increased among children over the last 50 years:

Source [1177](#) - Figure 2:

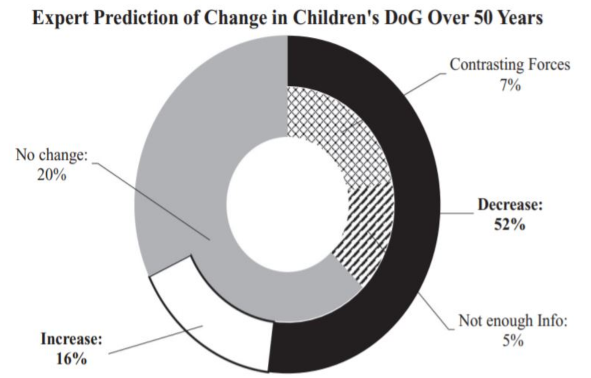


Fig. 2. N = 260. All participants who selected ‘no change’ (32%) were subsequently asked if this was because there was no change in DoG time, contrasting forces pushing ability up and down, or if there was not enough information to tell.

Source [1178](#):

This paper asked a range of social scientists to predict how the COVID-19 pandemic was going to impact things social scientists study (e.g. depression rates, political polarization, etc.). Said social scientists (n=717) were no more accurate than lay people (n=394) in their predictions.

Source [1179](#):

Contrary to these findings, this paper looked at the accuracy of 1,514 strategic intelligence forecasts. The average deviation between the objective and predicted probability of events was 13%:

Source [1179](#) - Figure 4:

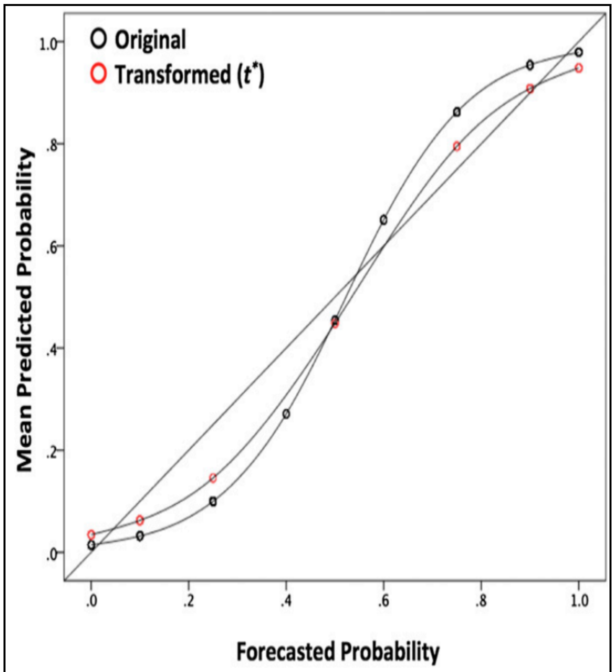


Fig. 4. Calibration curves before and after recalibration to t^*.

This degree of calibration is higher than what we’ve seen in other work. Unfortunately, there was no non-expert control group, so it is hard to judge how impressive this result really is. It should also be noted that these were short term predictions (59% under 6 months and 96% under one year) which probably increases accuracy [[1180](#)].

Source [1181](#):

This is a good overall book on the subject. There is a lot in it to unpack, but it is noteworthy that there is an inverse relationship between the qualities associated with good judgement, and the qualities valued in Media pundits.

This research literature is imperfect. The samples are limited and we might like to test other sorts of predictions that have not been studied. But the totality of available evidence suggests that academic experts in fields like finance, economics, psychology, law, and medicine, either can't predict reality well at all, or can't predict reality significantly better than interested non-experts. Overall, the evidence on predictive accuracy is another arrow pointing in the direction of our reasonable priors [more here].

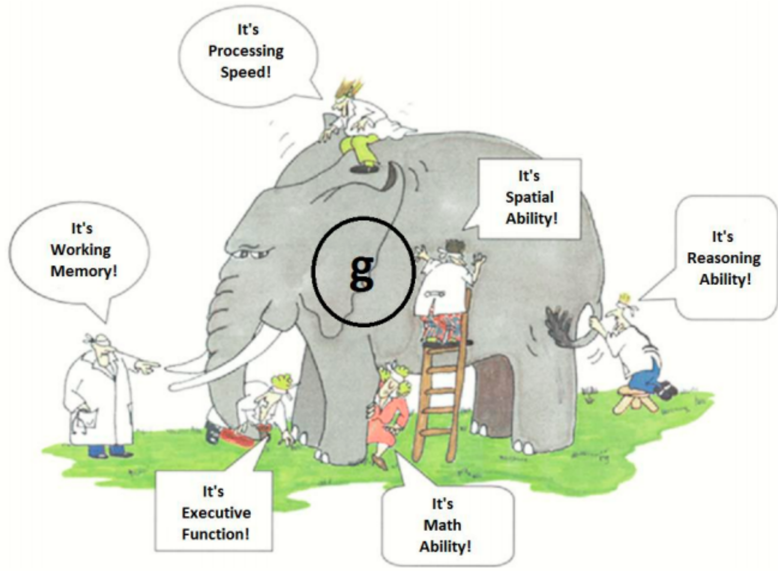
3. The g Factor

Navigation:

Navigation Continued:

Previous Chapter | Table Of Contents | Next Chapter

Blind Men and g Elephant



Summary:

Intelligence is important, so important that we call ourselves Homo-Sapiens, which is Latin for “wise man” [999]. So what is intelligence? Is it processing speed? Reaction time? Working memory? Verbal ability? Spatial ability? Humor ability? Rationality? Street smarts? Emotional intelligence? Video game abilities? Nobody has ever been able to come up with an assessment for any sort of cognitive ability which does not correlate with the rest of them [more here]. The intercorrelations are caused by a general underlying factor of intelligence [more here], and it consistently explains 30-50% of variance in a battery of cognitive ability tests [140]. The general factor also appears to be a genuine human trait rather than things like socioeconomics, education, culture, etc being general variables which affect many initially independent intelligences thereby causing them to all correlate with each other [more here]. Given this, we are statistically forced to accept the general factor of intelligence as measuring intelligence, at least to some degree, regardless of which cognitive ability we insist upon defining as intelligence.

Intelligence is a substantially heritable, substantially polygenic trait, with millions of genetic variants contributing to variance in intelligence [more here], with ~50% of variance in an IQ battery during childhood being caused by variance in genetics, ~80% of variance being due to genetics in adulthood, and heritability being ~90% for the general factor [more here]. The classical twin method is generally valid [more here], and our heritability figures apply to nationally representative samples [more here]. Many neurological influences on intelligence have been discovered [more here], and individual genetic variants appear to be tiny general factors, each explaining a small amount of variance in all tests [more here].

As we should predict from the trait’s generality, intelligence is probably the best predictor of life success [more here], influencing everything from educational and occupational success, to self control, to financial decision making, to longevity, to criminal behavior and beyond. This stated, the general intelligence factor is by no means the only important influence [more here]. High intelligence doesn’t guarantee correctness; although it increases the likelihood of rational thinking, it doesn’t matter how smart you are if you don’t stop to think.

Statistical Validity:

Imagine going to a local gym with a clipboard to record how much weight everybody can lift across a diverse series of different exercises, (lift 1, lift 2, and so on) and then testing for all of the correlation coefficients (r) between performance in every single exercise and performance in every single other exercise. This is done, and it produces the following correlation matrix (fictional example):

Lift	1	2	3	4	5	6	7	8	9	10	11	1st PC
1	1.0											.83
2	.67	1.0										.80
3	.72	.59	1.0									.80
4	.70	.58	.59	1.0								.75
5	.51	.53	.50	.42	1.0							.70
6	.45	.46	.45	.39	.43	1.0						.70
7	.48	.43	.55	.45	.41	.44	1.0					.68
8	.49	.52	.52	.46	.48	.45	.30	1.0				.68
9	.46	.40	.36	.36	.31	.32	.47	.23	1.0			.56
10	.32	.40	.32	.29	.36	.58	.33	.41	.14	1.0		.56
11	.32	.33	.26	.30	.28	.36	.28	.26	.27	.25	1.0	.48

Note that every single correlation in the matrix is positive, meaning that high performance on any given lift is associated with high performance on any other given lift, with higher correlations meaning a stronger association. In a sense, every single lift variable is a general factor which measures every single other variable to some degree. Lift 1 explains 100% of the variance in lift 1 ($r^2 = 1$), it explains 44.89% of the variance in lift 2 ($r^2 = .4489$), 51.84% of the variance in lift 3 ($r^2 = .5184$), and so on. Add the r^2 statistics together, and we get 3.8068. Divide by the number of variables in the matrix, 11, and lift 1 explains ~34.61% of variance in the lift correlation matrix. If we do the same for lift 2, we don't quite get the same result. Lift 2 tends to correlate with all the other variables less strongly than does lift 1, r^2 statistics added together equal 3.5101, and lift 2 explains 31.91% of variance in the dataset.

Zach is able to curl 1 gram more than Evan. Given this information, would we predict Zach to bench more and squat more weight, or would we predict Evan to do so? If forced to pick one or the other, we would choose Zach, but we wouldn't be very confident in our prediction. If on the other hand, 10,000 Zachs could, on average, bench press 50 kilograms more than 10,000 Evans can on average, and if we observed that the more that a lift predicts other lifts, the larger the Zach-Evan strength gap in terms of said lift is, then we would be very confident in saying that the group of Zachs is stronger than the group of Evans.

Factor Analysis:

How might we explain the pattern of intercorrelations? A statistical tool called Factor Analysis was developed by Charles Spearman to help answer such a question. Essentially, factor analysis is applied to a correlation matrix, post-hoc, to posit imaginary mediating variables to account for the variance in the correlation matrix with a smaller number of variables than exists in the raw correlation matrix. Here is a simpler matrix to consider:

Variable:	1:	2:	3:
1:	1.0	-	-
2:	1.0	1.0	-
3:	1.0	1.0	1.0

With the three variables all correlating perfectly, many would say that the three shouldn't even be considered to be separate variables. Given this, an obvious option that we have is to posit a single general variable (which we will abbreviate as "g") which perfectly correlates with all three variables:

Variable:	1:	2:	3:	g:
1:	1.0	-	-	1.0
2:	1.0	1.0	-	1.0
3:	1.0	1.0	1.0	1.0

In factor analysis, "g" would be referred to as a latent variable or latent factor. Latent variables are defined by the regression equations which are applied to raw measured variables in order to "predict" the latent variable. In other words, latent variables are defined by the statistical weights of measured variables, meaning that if all measured variables in the regression equation are standardized (expressed in z-scores), a latent variable is defined by the degree to which it correlates with the raw measured variables. In factor analysis, the degree to which a latent variable correlates with a measured variable is referred to as the degree to which said measured variable "loads" on said latent variable. In our example, variable 2 loads 1.0 on g. 1.0 is the "g-loading" of variable 2. 1.0 is also the g-loading of variables 1 and 3. A single general variable isn't our only explanatory option. If we wanted to, we could actually further complicate the raw correlation matrix. In our example table, we could posit a latent variable (g1) which correlates at 0.5 with all of the measured variables, meaning that it explains 25% of variance in every individual variable, and 25% of variance in the entire dataset. We could also posit a second latent variable (g2), which correlates at 0.0

with g1, but which also correlates at 0.5 with every single measured variable. With the two latent variables put together, we can explain 50% of variance in the dataset. With four such uncorrelated latent variables which load on every observed variable at 0.5 (g1, g2, g3, & g4), we could explain 100% of variance:

Variable:	1:	2:	3:
1:	1.0	-	-
2:	1.0	1.0	-
3:	1.0	1.0	1.0
g1	0.5	0.5	0.5
g2	0.5	0.5	0.5
g3	0.5	0.5	0.5
g4	0.5	0.5	0.5

We can also relax the requirement that latent variables be uncorrelated with (orthogonal to) each other, and posit latent variables which are exclusively defined by their loadings upon other latent variables, leaving us with oblique factors rather than orthogonal factors. Say for example that g1 correlated with g2, g3, and g4, each at 0.1; this common variance could be posited to be a third-order latent variable, with g1, g2, g3, and g4 being second-order variables, and the measured variables being first-order variables. Given the factor loadings remaining as previously defined, such multiple collinearity would require more latent variable(s) to be posited if we are to explain 100% of variance with latent variables.

We could also keep the requirement of orthogonality and simply say that a third order general factor is a sort of meta-property of the correlation matrix, that it explains 100% of variance in the measured variables, and loads at 0.5 on all of the second-order latent variables despite all of the second-order latent variables loading at 0.0 on each other.

This is the basic goal of factor analysis, to posit explanatory latent variables. A lot of the details of the technique have to do with the decision sequence (factor count, extraction method, rotation method, etc) determining what rules that factors are to follow before variables are actually posited. This is done in an attempt to make sure that factors are interpretable or sensible. For guides to factor analysis, see sources 175 (cited 4443 times!) and/or 176 (cited 14796 times!).

The Positive Manifold:

The first example correlation matrix is actually real intelligence test data from source 174:

Source 174 - Table 19.1:

(a)	V	S	I	C	PA	BD	A	PC	DSp	OA	1st PC
Vocabulary											.83
Similarities	.67										.80
Information	.72	.59									.80
Comprehension	.70	.58	.59								.75
Picture arrangement	.51	.53	.50	.42							.70
Block design	.45	.46	.45	.39	.43						.70
Arithmetic	.48	.43	.55	.45	.41	.44					.68
Picture completion	.49	.52	.52	.46	.48	.45	.30				.68
Digit span	.46	.40	.36	.36	.31	.32	.47	.23			.56
Object assembly	.32	.40	.32	.29	.36	.58	.33	.41	.14		.56
Digit symbol	.32	.33	.26	.30	.28	.36	.28	.26	.27	.25	.48

In this table, the first principal component ("1st PC") is basically a general latent variable which is common to all intelligence tests assessed in the sample. In this example, "1st PC" explains 48% of all variance. This finding, that scores on every single intelligence test ever created correlate with scores on every single other intelligence test ever created, is referred to as the positive manifold, and is the most well replicated finding in all of psychology. Source [140] reviews the correlation matrices of over 450 factor analytic studies and finds a general factor of intelligence to be a universal, finding consistently that the general factor of intelligence ("g factor" or "g") consistently explains 30-50% of variance in any given test battery. This is a more impressive proportion of variance to explain than many initially think because about 30% of variance is explained by test specificity, and about 10% of variance is

Principal Components Analysis:

"1st PC" in source 174 - Table 19.1 means first principal component. Principal components analysis finds the mathematically largest possible amount of variance which is common among all variables in a dataset, and posits it to be a latent variable: the first principal component. After the first principal component is extracted, principal component analysis creates a new correlation matrix showing what all of the intercorrelations would look like if the first principal component were held constant. The mathematically largest possible amount of common variance in the new matrix is then posited to be the second principal component, a third matrix is created, and so on until enough principal components have been extracted that no associations between any of the measured variables remain when all principal components are controlled for.
There is controversy over the use of principal components analysis because principal components are almost certainly overfitted to whichever dataset they were extracted from because they find the mathematically highest possible amount of common variance that each principal component can explain in a dataset, and the concept of statistical error applies to factor analysis too. The loadings of the measured variables on each of the principal components, according to principal components analysis, are almost certainly larger than they "really" should be. For more discussion of why, see the section on confirmatory factor analysis [more here].

explained by measurement reliability. The specificity of any given test question (or test item) is basically the degree to which performance on a question gives researchers absolutely no clue as to how somebody will perform on any other question. Measurement reliability is basically the degree to which participants will randomly give different answers when they take a test once, and then take the same test again.

The positive manifold is not merely a western phenomenon, it has been observed around the globe [181] and even in other species [182].

Various intellectuals have taken issue with the idea of a general factor of intelligence and have attempted to falsify the idea of it by explicitly setting out to create batteries of tests which do not produce uniformly positive correlations when tested. Despite the best attempts of psychologists for over a century, the g-factors of sufficiently large and diverse test batteries are highly correlated, pointed in roughly the same direction. The most straightforward was to test this is to employ latent variable modeling (SEM/CFA) and correlate the general factors from different IQ batteries. However, there is one study which does something perhaps more illustrative:

Source 238:

In this paper, Thorndike conducted a study which was explicitly designed to test the stability of a test's g-loading in multiple batteries (i.e. if we put the same test in two

Test Specificity:

High test specificity may arise, for example, if an incorrigible idiot is obsessed with horses, knows a lot about them, and answers questions about them correctly despite being relatively ungifted in actual cognitive abilities.

It doesn't matter how smart somebody is in the sense that they may be wrong about many things if they never stop to think about them.

This also invites an interesting consideration: one occurrence of potential possibility may be that certain people have a greater tendency to stop and think about things, and may in turn tend to score better on tests of knowledge because of this even beyond the degree of educational opportunity that such people experience. Such behavior would turn this kind of test specificity into common factor variance, and this is indeed something that people do to different degrees. Source 350 for example puts the heritability of independent reading at 62% for 10 year olds and 55% for 11 year olds. In his book [140], John Carroll argues for a three stratum hierarchical theory/model of cognitive abilities, with first-order measured tests at the bottom, second-order oblique factors in the middle, and the third order general intelligence factor (g) at the top. The most widely accepted model of intelligence, the Cattell-Horn-Carroll model of intelligence now includes both several fluid (low information load) and crystallized (high information load) abilities [259].

different non-overlapping test batteries, and extracted that test's g-loading from both batteries, how similar will the g-loadings be?). Thorndike started with 65 highly diverse tests used by the U.S. air force, he took a random 48 of them, and he randomly divided the 48 of them into 6 test batteries, with 8 tests in each, and with none of the 48 tests in more than one battery. Then, with the 17 tests not in any battery, they were inserted one at a time into all 6 batteries. The average correlation between g-loadings for all 17 tests was .85. From eyeballing the g-loadings in source [238](#) Table 2, it also seems like the most g-loaded tests were the ones whose g-loadings were most stable across batteries. If a g factor extracted from one of the batteries was itself treated as a probe test to be inserted into the other 5 batteries, the stability of its g-loading would likely be much higher.

Source [238](#) - Table 2:

Table 2. Factor loadings of 17 classification tests when inserted in six different matrices

Test	Matrix 1	2	3	4	5	6
1. Spatial orientation II	63	65	63	58	51	62
2. Reading comprehension	62	47	54	53	52	68
3. Instrument comprehension	48	56	63	51	49	58
4. Mechanical principles	43	61	59	47	33	57
5. Speed of identification	52	48	48	51	59	53
6. Numerical operations I	48	26	40	40	50	50
7. Numerical operations II	52	32	46	46	53	55
8. Mechanical information	20	30	26	18	08	49
9. General information	30	39	35	27	18	48
10. Judgment	43	35	39	37	39	51
11. Arithmetic reasoning	61	48	56	53	51	62
12. Rotary pursuit	21	30	33	24	24	28
13. Rudder control	12	28	28	15	09	28
14. Finger dexterity	34	25	38	35	33	37
15. Complex coordination	46	53	57	51	48	54
16. Two-hand coordination	25	35	37	35	33	39
17. Discrimination reaction time	52	55	61	59	60	61

This is strong, clear evidence that the g-loading of a subtest is not dependent on the test battery context in which its g-loading is derived, and this result has been replicated at least twice over [[1210](#) & [1211](#)].

[Thurstone:](#)

In a famous study published in 1938 [[504](#)], Thurstone claimed to have developed a test of seven independent mental abilities, these being verbal comprehension, word fluency, number facility, spatial visualization, associative memory, perceptual speed, and reasoning. However, the "g men" quickly responded, with Charles Spearman and Hans Eysenck publishing papers [[505](#) & [506](#)] showing that Thurstone's independent abilities were not independent, indicating that his data were compatible with Spearman's g model.

[Guilford:](#)

The idea of non-correlated abilities was taken to its extreme by J.P. Guilford who postulated as many as 160 different cognitive abilities. This made him very popular among educationalists because his theory suggested that everybody could be intelligent in some way. Guilford's belief in a highly multidimensional intelligence was influenced by his large-scale studies of Southern California university students whose abilities were indeed not always correlated. In 1964, he reported [[507](#)] that his research showed that up to a fourth of correlations between diverse intelligence tests were statistically insignificant. However, this conclusion was based on bad psychometrics. Source [508](#) reanalyzed Guilford's data and showed that after correction for statistical artifacts such as range restriction (the subjects were generally university students), the reported correlations are uniformly positive.

British Ability Scales:

The British Ability Scales were carefully developed in the 1970s and 1980s to measure a wide variety of cognitive abilities, but when the published test data was analyzed [509], the results were disappointing:

> *"the solutions have yielded perhaps a surprisingly small number of common factors. As would be expected from any cognitive test battery, there is a substantial general factor. After that, there does not seem to be much common variance left"*

This is despite the scales deliberately including tests with 'purely verbal' and 'purely visual tasks', tests of 'fluid' and 'crystallized' mental abilities, tests of scholastic attainment, tests of complex mental functioning such as in the reasoning scales and tests of lower order abilities as in the Recall of Digits scale.

CAS:

The Cognitive Assessment System (CAS) battery is based on PASS theory, which draws heavily on the ideas of Soviet psychologist A.R. Luria. It disavows g, asserting that intelligence consists of four processes called Planning, Attention-Arousal, Simultaneous, and Successive. The CAS was designed to assess these four processes.

Source 510 did a joint confirmatory factor analysis of the CAS together with the WJ-III battery, concluding that notwithstanding the test makers' aversion to g, the g factor derived from the CAS is large and statistically indistinguishable from the g factor of the WJ-III. The CAS therefore appears to be the opposite of what it was supposed to be: an excellent test of the "non-existent" g and a poor test of the supposedly real non-g abilities it was painstakingly designed to measure.

Independently, source 242 tested the CAS and the Woodcock-Johnson on 155 students between 8 and 11 years of age with joint confirmatory factor analysis, and the correlation between g factors was .98.

Triarchic Intelligence:

Robert Sternberg introduced his "triarchic" theory of intelligence in the 1980s and has tirelessly promoted it ever since while at every turn denigrating the proponents of g as troglodytes. He claims that g represents a rather narrow domain of analytic or academic intelligence which is more or less uncorrelated with the often much more important creative and practical forms of intelligence. He created a test battery to test these different intellectual domains. It turned out that the three "independent" abilities were highly intercorrelated, which Sternberg absurdly put down to common-method variance. A reanalysis of Sternberg's data by Nathan Brody [511] showed that not only were the three abilities highly correlated with each other and with Raven's IQ test, but also that the abilities did not exhibit the postulated differential validities (e.g., measures of creative ability and analytical ability were equally good predictors of measures of creativity, and analytic ability was a better predictor of practical outcomes than practical ability), and in general, the test had little predictive validity independently of g.

Piagetian Tasks:

The Swiss developmental psychologist Jean Piaget devised a number of cognitive tasks in order to investigate the developmental stages of children. He was not interested in individual differences (a common failing among developmental psychologists) but rather wanted to understand universal human developmental patterns. He never created standardized batteries for his tasks. Source 512 studied a battery of 27 Piagetian tasks which were completed by a sample of 150 children. Factor analysis of the Piagetian battery yielded a strong general factor underlies the tasks, with g-loadings ranging from 0.32 to 0.80:

Source 512 - Table 1:

Table 1
Correlations With Piagetian Composites of the Individual Tasks and Comparison With the General and Group Factor Loadings

Item	27 tasks	22 tasks	13 tasks	General	Factor 2	Factor 3	Factor 4
Conservation of substance	826	845	865	76	44	07	01
One-for-one exchange	697	722	734	66	42	−04	−05
Dissolution (weight)	592	578		51	24	09	21
Dissolution (substance)	701	716	699	65	31	13	03
Dissolution (volume)	369			36	02	−12	18
Conservation of weight	736	757	788	66	50	09	−03
Term-to-term correspondence	744	761	764	67	46	05	00
Class inclusion (animals 3)	554	547		52	19	50	−10
Class inclusion (animals 4)	685	668		66	13	52	−03
Class inclusion (animals 5a)	305			32	−14	39	03
Class inclusion (animals 5b)	408			42	−06	43	06
Conservation of volume 1	716	720	736	64	40	−06	15
Conservation of volume 2	454			43	−11	06	40
Rotation of beads	649	662		63	24	−05	06
Conservation of length	714	732	734	66	34	08	04
Conservation of length (rods)	775	798	838	67	57	05	01
Changing criterion	700	697	693	66	24	03	08
Conservation of liquid	759	783	812	68	54	−02	−01
Class inclusion (beads)	735	733	722	70	24	11	05
Disassociation (weight & volume)	746	708	658	70	04	08	34
Intersection of classes	405			40	−02	09	12
Rotation of squares (1)	630	639		60	14	−09	25
Rotation of squares (2)	506	504		48	10	−06	22
Two-three dimensions	608	614		56	24	−09	15
Changing perspectives (mobile)	688	662		69	−06	11	21
Changing perspectives (fixed)	827	816	771	80	12	08	18
Chemistry	671	660		66	08	05	16

Note. Decimals have been omitted.

Is the Piagetian general factor the same as the regular one? The same sample also took Wechsler's test. Scores were highly correlated, clearly indicating that they measured the same general factor. A small caveat is that the study included an oversample of mildly mentally retarded children in addition to normal children. Such range enhancement tends to inflate correlations between tests, so in a more adequate sample the correlations and gloadings would be somewhat lower. On the other hand, the data have not been corrected or measurement error which reduces correlations. Here are the correlations:

Source 512 - Table 2:

Intercorrelations of Three Piagetian Tests, Wechsler Verbal and Performance IQs, and the Academic Achievement Composite

Test	2	3	4	5	6
1. Piaget–27 items	—[a]	—[a]	800	825	754
2. Piaget–22 items		—[a]	795	825	739
3. Piaget–13 items			763	798	719
4. Verbal IQ				805[b]	840
5. Performance IQ					792
6. Achievement composite					—

When this correlation matrix of four different measures of general ability is factor analyzed, it can be seen that all of them load very strongly (~0.9) on a single factor:

Source 512 - Table 3:

Table 3
Unrotated and Rotated Factor Loadings From the Intercorrelations in Table 2

	Unrotated			Rotated		
Test	1	2	h^2	General	1	2
Piaget (27-item)	894	−192	836	873	273	005
Verbal IQ	913	126	849	890	054	230
Performance IQ	906	−122	836	884	226	056
Academic Achievement	896	209	846	874	007	285

Note. Decimals have been omitted.

It can be said that a battery of Piagetian tasks is about as good a measure of g as Wechsler's test. It does not matter at all that Piagetian and psychometric ideas of intelligence are very different and that the research traditions in which IQ tests and Piagetian tasks were conceived have nothing to do with each other; the positive manifold emerges without regard to the type of cognitive abilities called for by a test.

Video Games:

For the first time ever, a team of researchers measured videogame scores and also gave the participants standard IQ tests [241]. It was epic. The latent factors extracted from the video game score data shared a high percentage of common factor variance (81%) leading to a general video game factor (VG). The g factor extracted from classical IQ testing highly correlated with general gamer epicness (VG) at .93. The high correlations are all in spite of the restriction of range from participants all being university undergraduates.

Woodcock-Johnson:

The Woodcock-Johnson is another such test that was originally developed without regards to the g factor. It was originally developed for the Cattell-Horn theory where intelligence is posited to be best explained by fluid intelligence, which is supposed to be pure reasoning ability, and crystalized intelligence, which is supposed to be how much information somebody has memorized, and a multitude of fluid and crystallized latent oblique variables without a third-order g factor on top. See source 515 for descriptions of the tests. The 29 subtests of the revised 1989 edition of the Woodcock-Johnson IQ test are all correlated [516].

Source 516 - Table 1.4:

Table 1.4: Pearsonian intercorrelation matrix, combined kindergarten to adult sample (decimals omitted). 29 variables from the Woodcock-Johnson psycho-educational battery — revised, $N = 1425$ (correlations corrected for age).

Variable:		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
Memory for Names	1	1000																												
Memory for Sentences	2	279	1000																											
Visual Matching	3	213	254	1000																										
Incomplete Words	4	167	255	191	1000																									
Visual Closure	5	148	103	178	176	1000																								
Picture Vocabulary	6	404	403	202	267	229	1000																							
Analysis-Synthesis	7	275	324	280	205	161	323	1000																						
Visual-Auditory Learning	8	542	343	267	192	205	382	376	1000																					
Memory for Words	9	208	559	221	245	046	225	215	246	1000																				
Cross Out	10	170	241	621	168	241	242	291	265	203	1000																			
Sound Blending	11	245	323	245	367	133	323	265	332	335	246	1000																		
Picture Recognition	12	293	216	212	123	234	256	233	299	155	257	212	1000																	
Oral Vocabulary	13	388	534	310	319	234	632	419	405	364	315	389	304	1000																
Concept Formation	14	306	382	306	236	206	325	484	376	227	305	275	269	458	1000															
Memory for Names (Delayed Recall)	15	721	236	155	168	129	383	269	460	173	123	242	236	359	284	1000														
Visual-Auditory Learning (Delayed Recall)	16	345	164	162	120	192	255	269	460	110	168	192	275	271	323	446	1000													
Numbers Reversed	17	259	416	384	227	129	255	368	321	401	309	316	206	396	354	225	182	1000												
Sound Patterns	18	233	257	204	221	109	269	271	259	243	229	294	168	331	299	222	214	282	1000											
Spatial Relations	19	280	266	278	158	265	317	389	369	189	343	225	288	388	404	240	289	311	294	1000										
Listening Comprehension	20	331	469	266	334	204	576	349	344	279	263	351	256	642	375	294	221	308	274	320	1000									
Verbal Analogies	21	379	454	334	228	242	522	455	445	310	344	355	322	639	496	377	330	403	304	465	526	1000								
Calculation	22	256	331	435	142	132	299	423	347	252	358	293	208	471	401	249	242	413	257	376	374	483	1000							
Applied Problems	23	337	416	419	206	175	439	470	388	312	388	360	273	603	489	313	268	438	315	486	524	631	655	1000						
Science	24	380	437	260	285	233	633	368	364	246	280	323	246	658	389	362	270	336	260	385	619	544	440	570	1000					
Social Studies	25	371	477	298	262	200	626	386	374	270	278	323	255	693	411	348	245	332	256	344	638	595	508	617	702	1000				
Humanities	26	390	447	308	281	252	622	343	414	297	284	355	283	665	359	368	283	326	282	340	572	598	427	536	633	672	1000			
Word Attack	27	281	370	356	263	119	316	303	366	322	255	484	202	468	329	269	228	398	316	312	326	415	422	450	346	354	398	1000		
Quantitative Concepts	28	342	427	408	205	162	497	437	416	309	361	320	244	615	413	337	280	433	299	440	513	624	656	728	602	637	576	471	1000	
Writing Fluency	29	225	350	494	193	123	260	309	347	266	410	358	196	398	335	197	194	365	229	276	285	394	420	426	293	336	409	488	434	1000

Note: Reprinted with permission: McGrew, K. S., Werder, J. K. & Woodcock, R. W. (1991). *WJ-R Technical Manual* (p. 345). Allen, TX: DLM.

John B. Carroll did confirmatory factor analysis on the WJ-R matrix presented above to successfully fit a ten-factor model (g and nine narrower factors) to the data. Loadings on the g factor ranged from a low of 0.279 (Visual Closure) to a high of 0.783 (Applied Problems). The g factor accounted for 59% of common factor variance:

Source 516 - Table 1.5:

Table 1.5: LISREL estimates of orthogonal factor loadings for 29 variables on 10 factors (decimals omitted).

Stratum:	3	2	2	2	2	2	2	2	2	2	
Factor:	*g*	Glr	Gsm	Gs	Ga	Gv	Gc	Gf	Gq	Lang	h^2
Factor No.:	1	2	3	4	5	6	7	8	9	10	
01 MEMNAM	**478**	**695**	—	—	—	—	—	—	—	—	712
02 MEMSEN	**587**	—	**396**	—	—	—	—	—	—	—	501
03 VISMAT	**499**	—	—	**709**	—	—	—	—	—	—	752
04 INCWDS	**340**	—	—	—	**308**	—	—	—	—	—	210
05 VISCLO	**279**	—	—	—	—	**472**	—	—	—	—	301
06 PICVOC	**566**	—	—	—	—	—	**531**	—	—	—	602
07 ANLSYN	**591**	—	—	—	—	—	—	**213**	—	—	395
08 VISAUD	**579**	**343**	—	—	—	—	—	—	—	—	453
09 MEMWDS	**424**	—	**782**	—	—	—	—	—	—	—	791
10 CRSOUT	**478**	—	—	**539**	—	—	—	—	—	—	519
11 SNDBND	**490**	—	—	—	**642**	—	—	—	—	—	652
12 PICREC	**398**	—	—	—	—	**260**	—	—	—	—	226
13 ORALVO	**749**	—	—	—	—	—	**377**	—	—	—	703
14 CNCPTF	**623**	—	—	—	—	—	—	**543**	—	—	683
15 MMNADR	**439**	**729**	—	—	—	—	—	—	—	—	724
16 VSAUDR	**404**	**320**	—	—	—	—	—	—	—	—	266
17 NMRVRS	**571**	—	**203**	—	—	—	—	—	—	—	367
18 SNDPAT	**436**	—	—	—	**144**	—	—	—	—	—	211
19 SPAREL	**580**	—	—	—	—	**219**	—	—	—	—	384
20 LISCMP	**619**	—	—	—	—	—	**424**	—	—	—	563
21 VBLANL	**761**	—	—	—	—	—	**162**	**052**	—	—	608
22 CALCUL	**652**	—	—	—	—	—	—	—	**432**	—	612
23 APLPRB	**783**	—	—	—	—	—	—	—	**335**	—	725
24 SCIENC	**651**	—	—	—	—	—	**491**	—	—	—	665
25 SOCSTU	**686**	—	—	—	—	—	**488**	—	—	—	709
26 HUMANI	**661**	—	—	—	—	—	**448**	—	—	**107**	649
27 WDATCK	**587**	—	—	—	**273**	—	—	—	—	**197**	458
28 QUANCN	**743**	—	—	—	—	—	**177**	—	**400**	—	743
29 WRIFLU	**549**	—	—	**286**	—	—	—	—	—	**685**	852
SMSQ	9515	1235	810	875	602	338	1341	343	459	519	16037
%CCV	59.33	7.70	5.05	5.45	3.75	2.10	8.36	2.13	2.23	3.23	100.00

Measures of goodness of fit for the whole model:
CHI-square with 343 degrees of freedom = 1488.60 ($p = 0.000$)
Goodness of fit index = 0.931; Adjusted goodness of fit index = 0.912
Root mean square residual = 0.039

Note: Analysis of the correlation matrix of Table 1.4, which see for full names of variables. Factor Names (as given by McGrew *et al.*, 1991): *g*: General Intellectual Ability; Glr: Long-Term Retrieval; Gsm: Short-Term Memory; Gs: Processing Speed; Ga: Auditory Processing; Gv: Visual-Spatial Thinking; Gc: Comprehension-Knowledge; Gf: Fluid Reasoning: Gq: Mathematics; Lang: Language. h^2: Communality or Squared Multiple Correlation; SMSQ: Sums of Squares; %CCV: Percentages of Common Factor Covariance.

This finding, that the g factor accounts for more variance than all other factors put together, again, is routine [140].

Eventually, the g factor was accepted and incorporated into the Cattel-Horn-Carroll theory of abilities [259], by now the dominant, unifying paradigm. The WJ-III now also features a g factor on top of the hierarchy. Source 243 tested the Delis-Kaplan Executive Function System and the WJ-III Tests of Cognitive Abilities on 100 children and adolescents recruited from general school classrooms. The correlations between latent g's were .99 and 1.00. The g factor from the Woodcock-Johnson also correlates with the CAS g factor at .98 [242].

Gardner's “Multiple Intelligences”:

It seems that the only way to come up with an intelligence which isn't g-loaded is to redefine physical prowess, or various personality variables, as “intelligences”. In 1983, Howard Gardner published his book, *Frames Of Mind* [517] which outlined his theory of “multiple intelligences“ which included 7 “intelligence modalities” – musical, visual, verbal, logical, bodily-kinesthetic, interpersonal and intrapersonal (self-reflective). In 1995, he added “naturalistic intelligence”, and in 1999 he added “spiritual / existential intelligence”.

In a Q&A [519], Gardner describes his theory as follows:

> *"The theory is a critique of the standard psychological view of intellect: that there is a single intelligence, adequately measured by IQ or other short answer tests. Instead, on the basis of evidence from disparate sources, I claim that human beings have a number of relatively discrete intellectual capacities. IQ tests assess linguistic and logical-mathematical intelligence, and sometimes spatial intelligence; and they are a reasonably good predictor of who will do well in a 20th (note: NOT 21st) century secular school." … "Belief in multiple intelligences theory implies that human beings possess several relatively independent computers; strength in one computer does not predict strength (or weakness) with other computers. Put concretely, one might have high (or low) spatial intelligence and yet that does not predict whether one will have high (or low) musical or interpersonal intelligence."*

Gardner incorrectly describes the standard view. G-theorists do not say that the g factor is the only latent variable, just that a general factor exists, and is hugely important in that all mental tests substantially load on it. Gardner is also incorrect in claiming that IQ tests stopped being able to predict school grades in the year 2000 [518].

Those two falsehoods aside, this throws down his disagreements. Gardner basically denies any general intelligence factor whereas mainstream intelligence researchers contend that intelligence is both general and specialized. However, this may not even characterize Gardner, as Visser [521] notes that Gardner has diluted MI theory somewhat by incorporating the existence of g and suggesting that the intelligences might not be entirely independent.

One of the major difficulties in assessing Gardner's "multiple intelligences" theory is that Gardner is opposed to psychometric testing [520], so we have no way to measure "multiple intelligences", and he provides no testable hypotheses that would support his theory if confirmed and which would disqualify his theory if nullified.

Following source 520, there was a back and forth between Lynn Waterhouse and Gardner where Lynn argues that Multiple Intelligences, the Mozart Effect, and Emotional Intelligence should be discarded because they are have no supporting evidence and are contrary to established findings [522, 523, 524, 525, & 526].

Despite Gardner's aversion to science, in 2006, Visser attempted to put the theory to the test anyways [521]. g-loadings ranged from 0.03 to 0.75 as seen below:

Source 521 - Table 3:

Table 3
g loadings of tests and correlations of tests with Wonderlic Personnel Test (WPT)

Ability Domain	Test	*g*-loading	$r_{(WPT)}$
Linguistic	Opposites	0.50 (0.61)	0.41** (0.56)
	Vocabulary	0.54 (0.66)	0.47** (0.64)
Spatial	Map Planning	0.55 (0.61)	0.48** (0.60)
	Paper Folding	0.50 (0.57)	0.48** (0.62)
Logical/ Mathematical	Subtraction and Multiplication	0.24 (0.25)	0.36** (0.42)
	Necessary Arithmetic Operations	0.70 (0.78)	0.67** (0.83)
Interpersonal	Cartoon Predictions	0.37 (0.55)	0.23** (0.38)
	Social Translations	0.53 (0.56)	0.38** (0.45)
Intrapersonal	Accuracy	0.16 (N/A)	0.11 (N/A)
	Consistency	0.27 (0.37)	0.27** (0.41)
Naturalistic	Diagramming Relationships	0.75 (0.83)	0.59** (0.73)
	Making Groups	0.57 (0.64)	0.38** (0.48)
Bodily-Kinesthetic	Stork Stand	0.03 (0.03)	−0.04 (−0.05)
	Mark Making	0.06 (0.06)	0.03 (0.03)
Musical	Rhythm	0.18 (0.34)	0.08 (0.17)
	Tonal	0.10 (0.24)	0.07 (0.19)

Values in parentheses are corrected for unreliability in the individual ability tests only (for the *g*-loadings) or in both the individual ability tests and the WPT (for the WPT correlations).
N=200. *p<0.05. **p<0.01, two-tailed.

Why the near zero loading of Bodily-Kinesthetic? The description of the ability, and even its very name, should inspire skepticism. To quote from the paper below:

> *"Gardner (1999) described this intelligence as the potential of using the whole body or parts of the body in problem-solving or the creation of products. Gardner identified not only dancers, actors, and athletes as those who excel in Bodily-Kinesthetic intelligence, but also craftspeople, surgeons, mechanics, and other technicians."*

So strength and dexterity are apparently now redefined as "intelligences".

Gardner has however dismissed Visser as "failing to grasp the core of MI theory" [527], to which Visser has responded in source 528. Visser concludes with the following:

> *"it remains unclear to us what it is that MI theory can explain about intelligence, above and beyond what has already long been known. Gardner could clarify this "core" for us, by providing falsifiable, testable, MI-based hypotheses that would predict results different from those predicted by existing models of the structure of mental abilities."*

Emotional Intelligence:

"Emotional Intelligence" is mostly just a combination of intelligence and personality measures [529], though it does have some validity beyond the two and may be another g-loaded factor like spatial ability, verbal, etc. In the paper [529], the correlation between IQ and their operationalization of emotional intelligence was .454. Combining IQ with the personality trait of agreeableness from the big 5 test, and with whether or not an individual is female in a regression model created a correlation of .617. However, psychometric tests generally don't have perfect reliability. Say you measure the height of a bookshelf once, and then do so a second time, a bunch of people do so and the correlation between time 1 and time 2 is .95 instead of a perfect 1.0; that's measurement unreliability. Correcting for measurement realibility in the IQ+Agreeableness+Sex composite brings the correlation R^2 for emotional intelligence up to .806. Further, a meta-analysis [530] looked at prediction of job performance from EQ, and its independent effect was smaller than that of IQ+personality.

Humor Ability:

In this paper [494], a sample of 270 young adults completed a battery of humor production tasks and three of the second-order abilities in the Woodcock-Johnson. The paper found that the general intelligence factor correlated with the paper's operationalization of humor ability at .51.

Street Smarts:

In a meta-analysis on the subject, source 377 found a .46 correlation between performance on situational judgement tests (SJTs) of real world problem solving and performance on standard IQ tests.

The Rationality Quotient:

Intelligence is related to rationality and skepticism towards unfounded beliefs [286]. In 2016, Stanovich, West, and Toplac came up with a formal test of rationality (the CART) in their book [376], which was supposed to be an attack on IQ tests for not being the same thing as rationality. However, their own data (table 13.11) shows performance on their "Comprehensive Assessment of Rational Thinking" test to correlate with IQ at .695. So with respect to critical thinking, IQ is strongly correlated with formal tests of rationality that gauge people's propensity to incorrectly use mental heuristics or think in biased ways:

Source 376 - Table 13.11:

Table 13.11
Correlation comparisons between the full-form CART (20 subtests), the short-form CART (11 subtests), and the residual CART (9 subtests) in RT60

	Full-Form CART	Short-Form CART	Residual CART
Cognitive Ability Composite3—Turk	.695	.671	.620
Cognitive Ability Composite3—Lab	.567	.546	.474
SAT Total—Turk	.313	.319	.253
SAT Total—Lab	.495	.489	.384
Cognitive Ability Composite4—Turk	.713	.699	.638
Cognitive Ability Composite4—Lab	.614	.595	.506
Sample (Turk = 1; Lab = 2)	–.283	–.260	–.280
Sex (Male = 1; Female = 2)	–.322	–.320	–.265
Actively Open-Minded Thinking scale—Turk	.628	.631	.508
Actively Open-Minded Thinking scale—Lab	.554	.568	.387
Deliberative Thinking scale—Turk	.267	.281	.191
Deliberative Thinking scale—Lab	.472	.470	.360
Future Orientation scale—Turk	.311	.296	.286
Future Orientation scale—Lab	.297	.278	.267

For Cognitive Ability Composite3 ($N = 747$)
Correlations > .075 significant at the .05 level, two-tailed
Correlations > .126 significant at the .001 level, two-tailed
For Cognitive Ability Composite4 and SAT ($N = 538$)
Correlations > .086 significant at the .05 level, two-tailed
Correlations > .141 significant at the .001 level, two-tailed

One formal logical fallacy is the appeal to authority fallacy ("the government says it therefore it's true!"). Source 378 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.

Standardized School Tests:

Standardized tests like the SAT, ACT, and GCSE used for measuring performance in schools are not designed to be diverse test batteries that yield high quality g-factors, but they also highly correlate with classical IQ tests nevertheless (see next page). Similarly, a meta-analysis [245] going over more than 200 samples totaling 105,185 students shows that IQ tests strongly predict grades at .54. The difference between standardized tests and grades are that grades are more subject to reference group effects where an A means higher performance than the local peer group but not necessarily higher performance than nationally representative samples (i.e. an A from one school may be equivalent to a C from another school). This is one of the reasons that equal predictive validity for two groups can sometimes be sometimes evidence of test bias against one of them. Strenze [253] also did a large review of longitudinal studies and found that IQ is actually slightly better at predicting educational attainment than are grades.

Source #	Test:	Correlation with IQ:	Sample Size:
246	SAT-Verbal	0.80	339
246	SAT-Math	0.70	339
246	ACT	0.87	339
247	SAT	0.86	917
247	SAT	0.72	104
248	SAT	0.58	97
249	GCSE	0.81	70,000

Others:

Source 239 tested 3 test batteries comprising 42 different cognitive tests as part of the Minnesota Study of Twins Reared Apart. The correlations between g factors were .99, .99, and 1.00. The three tests were the Comprehensive Ability Battery, the Hawaii Battery, and the Wechsler Adult Intelligence Scale. Each test battery utilizes many, highly diverse operationalizations of intelligence (see the report [239] for descriptions of the tests). All 861 correlations between subtests, regardless of test battery, were positive.

Source 240 tested 5 batteries on 500 Dutch seamen. With the exception of the Cattell Culture Fair Test, all of the correlations between g factors were at least .95. The lowest correlation between g factors, coming from the Cattell Culture Fair, was .77. The reason for the results from the Cattell Culture Fair is that it tests a very non-diverse set of 4 reasoning tasks each of which were very similar tasks, so it was more like a single g-loaded subtest than an entire battery being tested. These high correlations are in spite of the range restriction.

Source 244 tested six batteries on five samples of children and adolescents with sample sizes ranging from 83 to 200. Three correlations between g factors exceeded .95, but two were relatively lower at .89 and .93. The lower results may be due to sampling error and temporal changes related to growth.

Alternatives To g-Theory:

Given the evidence [more here], the existence of the positive manifold (the finding that intelligence tests all intercorrelate) can be appropriately characterized as scientific law in psychology (a scientific law being a repeatedly upheld observation, and a scientific theory being a well supported narrative that attempts to account explain the existence of many laws parsimoniously). A general intelligence factor could be posited as helping to explain the pattern of observed intercorrelations, but as we have noted, the mere finding of a positive manifold, on its own, is not necessarily enough to make a general factor of intelligence a theoretical necessity [more here]. A general intelligence factor is not necessitated by the positive manifold alone because there are alternative theories that, if true, could also explain the positive manifold. These alternative theories are known as "Mutualism" and "Sampling Theory".

Mutualism:

The first alternative theory, known as "Mutualism", posits that many intelligences exist in humans which are initially uncorrelated at birth, but which all assist each other's performance, causally affecting levels of the other intelligences, and thereby making all of the intelligences become correlated with each other when they initially were not.

The most obvious prediction which is made by Mutualism Theory, that intelligence tests will gradually become more correlated from birth until death, is not observed [1149]:

Source 1149 - Figure 2:

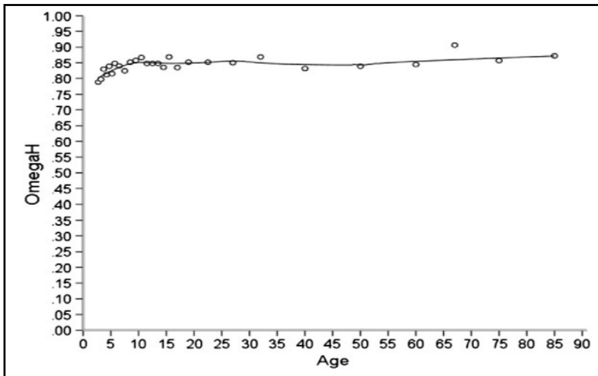


Fig. 2. Scatter plot depicting the association between age and the strength of the *g* factor (age range: 0 to 90 years).

Source 1149 - Figure 4:

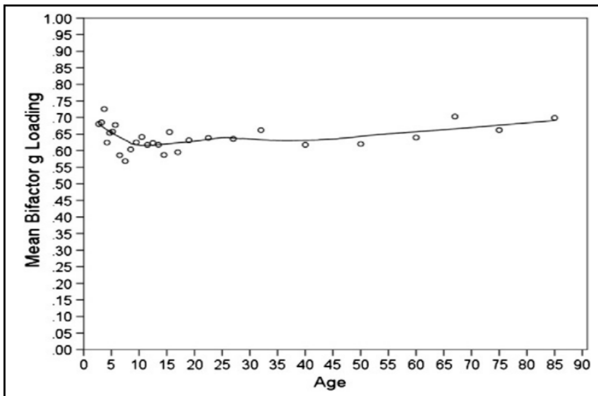


Fig. 4. Scatter plot depicting the association between age and mean bifactor *g* loading (age range 0 to 90 years).

Another problem for Mutualism which is worth mentioning is that many experimental interventions which aim to increase IQ affect the more specific variance in a test battery rather than the more general Variance. This has been observed with adoption [306], the Flynn Effect [274, more here], head start programs [142], cognitive training [276], retesting [275], deafness/blindness [952], and education [more here]. Additionally, when individuals are taught to perform better on tests or test items, this decreases test / item g-loadings rather than increasing peoples' general intelligence factor scores [275 & 416].

Sampling Theory:

The second theory to explain the positive manifold is that there are many intelligences, which may even be completely uncorrelated, but that the positive manifold is an artifact of test construction, meaning that performance on any given intelligence test is dependent on many independent abilities. Sampling theory states that the intelligence tests are correlated because they test performance on common abilities rather than the abilities themselves being correlated. Here is an illustration [7]:

Source 7 - Figure 5.2:

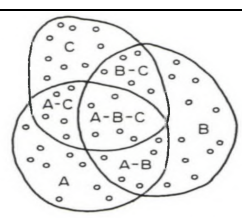


Figure 5.2. Illustration of the sampling theory of ability factors, in which the small circles represent neural elements or bonds and the large circles represent tests that sample different sets of elements (labeled A, B, and C). Correlation between tests is due to the number of elements they sample in common, represented by the areas of overlap. The overlap of A-B-C is the general factor, while the overlaps of A-B, A-C, and B-C are group factors. The non-overlapping areas are the tests' specificities. *Source: Bias in mental testing* by Arthur R. Jensen, Fig. 6.13, p. 238. Copyright © 1980 by Arthur R. Jensen. Reprinted with permission of the Free Press, a Division of Simon & Schuster, and Routledge Ltd.

One sort of version of sampling theory could be consistent with a completely biological intelligence: if some people are smarter than others because their neuron cells produce protein A and/or protein B, then while the ability to produce protein A may be a separate ability from producing protein B, the intellectual abilities that the proteins support may require (sample) the production of both proteins. This sort of a sampling theory is less falsifiable and may not even conflict with a unidimensionality of intelligence in a broader task-oriented sense that the layman may conceptualize the topic.

The first thing which should be mentioned is that if it is the case the sampling theory is true in a broad task-oriented sense, then we know that this phenomenon is certainly unintentional because various researchers have taken issue with g-theory, explicitly set out to create intelligence tests which are uncorrelated, and failed to accomplish this [more here]. There are also three more findings which likely falsify sampling theory, intentional or unintentional, in the task-oriented sense.

The first of them is that if sampling theory, in the task-oriented sense, is true, it would have to explain why performance on incredibly basic abilities, such as reaction time or sensory perception, have positive g-loadings. Reaction time, for instance, has a negative correlation of -.18 to -.28 with g, meaning that smarter people react faster on elementary cognitive tasks [1150].

The second of them is that tests which are seemingly highly dissimilar in the task sense are empirically highly correlated with each other, as sampling theory in the task-oriented sense should not predict [7 - pages 120-121].

Finally, the third and possibly most convincing is that the g-loading of a given subtest is mostly invariant with regards to which test battery the subtest's g-loading is calculated from [238, 1210, & 1211]. This, as well as the consistency of g factors derived from different test batteries, are clear demonstrations that the properties of g are largely invariant with regards to test content.

What Is Intelligence?

Given the findings thus discussed, explanations of the positive manifold alternative to g theory fail. Intelligence is thus a highly unidimensional trait, at least in the broad task-oriented sense. Thus, this unidimensionality should be represented as a single variable, g, via factor analysis. Given this, it doesn't matter how we choose to define intelligence. We could define intelligence as school achievement, rationality, street smarts, humor ability, emotional intelligence, working memory, reaction time, video game scores, etc, and it wouldn't matter. Regardless of our definition(s) of intelligence(s), theoretical background(s), or operationalization(s) of intelligence, the reality of g theory statistically forces us to accept the general factor of intelligence as measuring "intelligence", at least to some degree. So do IQ tests test intelligence? Sort of, IQ test batteries are just a collection of tests with the highest g-loadings.

Confirmatory Factor Analysis:

Factor analysis, as thus discussed [more here], has actually mostly been discussed in reference to a specific type of factor analysis called exploratory factor analysis. There is another kind of factor analysis called confirmatory factor analysis which aims to test models of latent variables against each other in a pre-hoc manner rather than a post-hoc manner by utilizing fit statistics of explained variance, or significance. Essentially, in confirmatory factor analysis, researchers specify models of intelligence beforehand (what all of the latent variables are and how much they should correlate with each other and with all of the raw measured variables), and then use confirmatory factor analysis to assess what the probability is that the various models of intelligence could generate the observed test data.

g-Theory performs well in confirmatory factor analysis [514; 513, pp.125-156; 1151; 1152; more covered here], with the Cattel-Horn-Carroll hierarchical model explaining a substantial portion of variance. In a sense, Carroll's 1993 book [140] which used exploratory factor analysis was also illustrative of this because the book showed that the same patterns emerged in each of the 450+ datasets in which it employed its EFA techniques.

However, it should be noted that factor analysis (both exploratory and confirmatory) is just a correlational statistical tool in the general linear model [175 & 176], and correlation is not causation. Confirmatory factor analysis, like exploratory factor analysis, is not equipped to favor certain models of intelligence over another; it is largely just a game of which theory's theorists are better at making models. Confirmatory analysis is equipped to show that a model with both a g factor and oblique second-order factors fits test data better than a model with only a g factor, but so is exploratory factor analysis. However, neither are equipped, based on the correlational structure of test data alone, to test g-Theory, Mutualism, sampling theory, etc against each other; external evidence is required. Both are also unequipped, based on correlational structure alone, to determine whether a model with a general factor and with correlated second-order factors is more theoretically parsimonious than a model with only correlated latent primary abilities at one level; both theories can have a model made for them which explains just as much test data as a model from another. In fact, there are an infinite possible number of equivalent solutions to factor analysis.

Infinite Solutions:

> Some deride factor analysis as being useless because there are an infinite possible number of equivalent solutions to the factor analysis of a dataset. However, what is missed by this thinking is that there are also infinite solutions (and a larger infinity) which factor analysis is equipped to say are not possible. Moreover, the impossible solutions are qualitatively different from the possible solutions, so it is useful and theoretically important to eliminate them.

Despite equivalent mutualist and general hierarchical model solutions to a given dataset being possible, a theory which just posits that the raw correlation matrix of measured variables is the true structure of intelligence will probably be advantaged in that it doesn't actually have to do any theorizing, and it automatically explains 100% of original variance without any effort on its part. One paper which does exactly this [1153], unsurprisingly, finds their mutualist model to account for test data better than their chosen hierarchical model. Not only was the mutualist model advantaged as thus described, the mutualist model was also clearly overfitted because it was derived from an exploratory factor analysis on the dataset which was used to do the comparison while only the g model was duly specified pre-hoc. These problems aside, comparison of model fit statistics is not equipped to decide upon one theory or another.

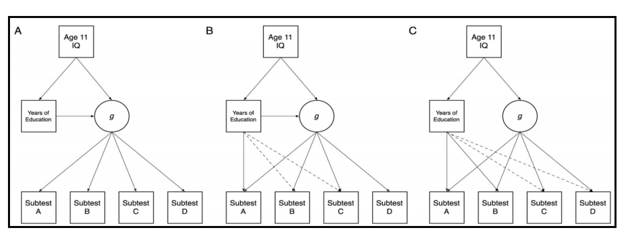
Is g A Trait?

So g exists [more here], and intelligence is substantially unidimensional. But what *is* g? One proposition is that g isn't an actual intellectual ability, but just a person's socioeconomic class. Worth noting is that even if it were shown that socioeconomic class causally affects the general factor of intelligence, which is a tall order on its face because causality is difficult to show, it could be the case that despite such a finding, g really is an intellectual ability, but socioeconomics just influences it. The influence of socioeconomics on g wouldn't necessarily prove that socioeconomics affects all of the specific abilities thereby causing them to correlate and explaining the positive manifold.

Education Duration:

The most recent meta-analysis on the effect of an extra year of education on IQ [630], a great, large, well done meta-analysis, finds an increase of at most 5 IQ points. It doesn't merely look at the correlation between IQ and grades or years of education, but rather it analyses three different types of quasi-experimental studies to see what effect schooling has on the IQ scores of individual people. No substantial publication bias was discovered in the meta-analysis. The fadeout effect [305] of IQ gains from **early intervention / Head Start** programs was also replicated in the new meta-analysis [630]; the effect size for the smallest age gap between retesting was a gain of ~2.4 IQ points while by contrast, the effect size for the largest age gap between retesting was a gain of ~0.3 IQ points. One thing the meta-analysis does not assess however is the effect of education on the general intelligence factor (g). Source [536] used structural equation modeling on an extremely longitudinal sample (~60 year gap) to see if the effects of education on IQ are actually on g. The first model tested was that extra education was purely associated with increases in g. The second model was that extra education was associated with increases in g as well as other, more specific abilities. The third model was that extra education was only associated with IQ through specific abilities rather than g. The authors found the last model was the best fit. They also ran other analyses to confirm these results; no matter what, the third model of education having no impact on g was the best fit:

Source 536 - Figure 1:



Similar results were shown by source [631](#). The authors in this study took longitudinal data on education and IQ and tested if the gains were associated with increase in various reaction time tests. This is mainly important because reaction times generally tell us about processing speed and reasoning ability in the brain. They found that the effects of education were not on reaction times after controlling for a number of variables. While the authors argue this does not tell us if the education gains are on g or not [536]. However, the effect of education on reaction times after controlling for other variables was larger on simple reaction times than on choice reaction times, which is the more g-loaded test [632]. Similarly, we can test this by seeing whether or not fluid intelligence is increased by education. Fluid intelligence has to do with reasoning abilities whereas crystallized intelligence is the accumulation of knowledge and skills over time. One study of about 1,367 eighth graders in Boston public schools found that while schools were able to increase the achievement test scores in the schools, the programs for the former were not able to increase fluid intelligence skills like working memory capacity and info processing [633].

Other longitudinal models show g variation causes educational achievement differences rather than the other way around. These are pretty straight-forward studies. Basically, they take data on IQ and abilities at two points and do a cross-lagged panel analysis. They take a cross-lagged path from g at time 1 and educational achievement at time 2 and another path from educational achievement at time 1 and g at time 2. They compare these and make a causal inference based on which is stronger. Both of the studies done on this show the path of g to educational achievement is stronger than the latter and that the other is statistically insignificant [634 & 635].

Finally, a Nijenhuis meta-analysis does not show much of a Jensen effect [697].

Given the evidence, educational duration affects specific abilities rather than g, so we don't even have to ask the question of whether or not education is g or is merely an influence on g.

Educational Quality:

But perhaps educational quality is what matters rather than the raw number of years of schooling. Probably not, voucher studies where a random selection of poor kids are sent to prestigious schools to be compared to poor kids who happened to not receive a voucher, which is thus an apples to apples comparison, find that school quality has barely any effect on school test scores:

The Cleveland Voucher Program [730]:

Grade:	Voucher:	No Voucher	Non-Applicant
1	555	546	548
2	587	577	580
3	615	605	607
4	632	620	624
5	643	636	636
6	654	639	638

The Milwaukee Voucher Program [731]:

Grade/Subject:	G1 - 2006:	G2 - 2006:	G1 - 2010:	G2 - 2010:
7 - R	432.2	435.3	492.2	485.4
8 - R	446.5	436.9	505.1	486.1
9 - R	458.0	472/9	593.5	492.0
7 - M	388.2	395.7	501.6	500.0
8 - M	426.3	424.4	504.2	493.3
9 - M	462.9	478.7	515.5	524.2

G1: Received Voucher; G2: Denied Voucher; M = Math; R = Reading.

The Washington DC Voucher Program [732]:

Group:	Math:	Reading:
Voucher:	541.00	645.92
Applicant:	543.36	645.24

Voucher given at the beginning of high school, test scores from the end of high school.

Income:

School test scores and grades, a proxy for IQ tests [more here], are not affected by guaranteed income experiments. Given this, we don't even need to test the effect on g, or if income is g.

Source 696:

This analysis of 16 experiments of randomly assigned welfare found that increased income improved teachers' ratings of student performance, but had no effect on test scores.

Source 698:

This guaranteed income experiment on children in North Carolina and Iowa produced no effect on GPA in Iowa and a 6.2% increase in GPA in North Carolina for young children. No effect was found in either state for high schoolers.

Source 699:

Differences in family income didn't predict sibling differences in most cognitive abilities with one exception: a $10,000 increase in income did predict a 0.22 SD increase in reading ability.

Source 700:

This guaranteed income experiment on poor Black children increased reading scores by .23 SD and had no effect on GPA for grades 4-6. It had no effect on reading scores and a negative effect on GPA (-.18SD) for grades 7 – 10.

Is g-Loading Cultural Loading?

Sources 656 and 657 claim to show that test heritabilities, g-loadings, and group differences are all larger on the more culture-loaded tests. The devil is in the operationalization of the culture-loading of a test, though the operationalization which is employed is very intuitive to the layman. Kan defines the cultural-loading of a subtest as the percent of content for a WISC subtest which is changed when the test is translated into a different language for a different country and/or the degree to which test content is crystallized. The eye catching results are that more heritable, g-loaded tests with larger group differences are the ones with more cultural loading. The degree to which test content is changed for international translations is likely exclusively determined by the degree to which test content is crystallized, having to do with information. It could just be that this sort of finding is just a peculiarity of the WISC, as the opposite has been shown when tested in other test batteries [658].

We may expect that since adoption transplants people from one socioeconomic culture to another, we may take adoption as a more objective cultural load variable. Given Kan's results, we may naïvely expect that IQ gains from adoption are to be stronger on the more g-loaded tests, but this is not the case [306]. Similarly, some other variables we might accept as more objective cultural load variables such as the degree to which test performance is impacted by adoption [306], head start programs [142], retesting [275], the Flynn Effect [274, more here], cognitive training [276], education [more here], and deafness/blindness [952] also show that the g-loaded tests are the 'culture reduced' ones.

Using multiple different procedures for classifying the culture loadings/biases of tests (e.g. expert opinion of the magnitude of content bias, group differences in the rank order of item difficulty, and more formal psychometric measures of group differences in how certain items are related to other items), Jensen and McGurk [658, p. 298] showed that holding constant item difficulty, by all measures, Black-White differences on culture-reduced items are larger than or equal to Black-White differences on cultural items [see also 659, pp. 56-62; 660, pp. 178-179; 661, pp. 426-427; & 662, pp. 210-213]. Given the extensive literature on this subject reviewed by [663 - ch. 4, 12, & 17; 184 - ch. 10, 11 & 12; & 7 - ch. 11], and given the evidence thus discussed, it must be recognized that group differences are larger on culture-reduced tests.

While Jensen has argued [184, p. 133] that culture-loaded tests are not necessarily culture-biased, he has made it clear that a culture-influenced test should be manifested through group differences in the **meanings** of the tests/items. What remains to be seen from Kan's results is whether or not these culture-loaded tests/items really behave differently across groups. By all evidence regarding racial bias in IQ tests, this is not the case [see more]. For alternative interpretations of Kan's results, see source 664.

-Note on the Method Of Correlated Vectors:

One sign that an environmental variable only affects specific abilities rather than the g factor would be if it affects less g-loaded tests more than it affects less g-loaded tests. This is the case for the effects of retesting [275], head start programs [142], deafness/blindness [952], the Flynn Effect [274; more here], and cognitive training [276].

The act of running the correlation between subtest g-loading and other subtest characteristics is called Jensen's method of correlated vectors (MCV), as devised by Arthur Jensen [7]. A correlation between subtest g-loading and other subtest characteristics is called a Jensen effect. Some cite sources 601, 602, 603, & 604 as proof that the MCV is a generally invalid method, but this is not their correct interpretation; these criticisms only apply to the results of item-level MCV results rather than test/subtest-level results. This is also understood by users of the MCV such that most tests avoid using CTT item-level statistics. Evading this issue, source 605 shows how Schmidt & Hunter's method for dealing with dichotomous variables can be used for the purposes of translating CTT item-level data into IRT, keeping MCV valid.

Conclusion:

Since socioeconomics, culture, education, head start programs, the Flynn Effect, retesting, cognitive training, education, and deafness/blindness do not affect the common factor variance, they cannot explain the existence of the positive manifold, g seems to be a genuine trait rather than just a genuine latent variable.

The Flynn Effect:

Many laymen know of the phenomenon dubbed “The Flynn Effect”; average “IQ scores” have been rising over time for quite some time. James Flynn wasn't the first to observe this phenomenon, but he popularized it and did a gargantuan amount of work demonstrating its occurence. Unfortunately, the Flynn Effect is beginning to stop in more developed countries, and in some countries, it is now reversing [262]. Moreover, in at least the Netherlands, the anti-Flynn Effect is g-loaded [263]. The normal Flynn Effect however, has negative MCV results [274]; source 274 meta-analyzed 11 data points from 5 papers (total N= 16,663), and found a -.38 correlation between Flynn Effect score gains and test g-loading. More experimentally a psychometric meta-analysis of 64 test-retest studies [275] yields the maximally negative -1.0 correlation between g-loadings and score gains from retesting. There is also evidence that score gains on IQ subtests cause decreases in the g-loadings of the subtests to which the gains apply [275 & 416].

-Types Of Measurement Invariance (IRT):

Statisticians can test for something known as measurement invariance, usually as a test for whether or not a test is biased against one group or another. The purpose is basically to test for whether or not a construct has the same properties in two different groups, and so is useful in discussion of the Flynn effect because it could be the case that score changes are a result of test properties changing with time rather than genuine increases in g. According to the book on Confirmatory Factor Analysis referenced earlier [176], a few different types of measurement invariance can be distinguished in the common factor model for continuous outcomes:

1. **Equal Form:** The number of factors and the pattern of factor-indicator relationships are identical across groups.
2. **Equal Loadings:** Factor loadings are equal across groups.
3. **Equal Intercepts:** When observed scores are regressed on each factor, the intercepts are equal across groups (When intercepts are unequal, individuals from two groups matched in latent abilities will have different mean scores on a subtest. Differences in intercepts means a systematic advantage for one group over another).
4. **Equal Residual Variances:** The residual variances of the observed scores not accounted for by the latent factors (item-specific variances) are equal across groups.

When types 1 & 2 are shown to hold, this is known as metric invariance. When type 3 also holds, this is known as strong/scalar invariance. When all four conditions are met, this is known as strict invariance.

Source 264:

This study was probably the first to assess measurement invariance across time. Wicherts and his colleagues used data from a variety of sources and measurement invariance was violated across every single one of them. This study provided very strong evidence that the Flynn Effect might not represent a genuine increase in any of the latent factors and much of it might just be changing psychometric properties. Wicherts and his colleagues warned that more data, especially IRT analysis, needs to be used. Did anyone apart from a handful of people actually listen? Of course not.

Source 277:

Pooling six articles with comparable cohorts separated by about 50 years or so, consistent violations of measurement invariance across cohorts who had taken Raven's Progressive Matrices were found. This is a good counter-counterpoint to people who say that g has changed because RPM is supposed to be an almost pure measure of g; it is nowhere near pure g, see source 278.

Source 265:

Alexander Beaujean's PHD dissertation; this was rather easy to find for a dissertation. The first half uses simulations to demonstrate that Item Response Theory is much more suitable than Classical Test Theory at distinguishing between genuine cognitive gains and psychometric artifacts. The second half of the dissertation used data from the mathematics section of the College Basic Academic Subjects Examination to examine the Flynn Effect. Using CTT, there was a retrograde of the Flynn Effect in the mathematics test of -.178 standard deviations per year. IRT analysis revealed a higher reverse Flynn Effect of -.222 sd units per year so CTT was masking the magnitude of the decline.

Source 266:

This one used Item Response Theory to examine the Flynn Effect in the NLSY. When controlled for differential item functioning, there was no Flynn Effect in the PPVT-R and a much more negligible Flynn Effect in the PIAT-M data. To quote the authors:

> *"Thus, for the data used in this study, the Flynn Effect appears to be largely the result of changing item properties instead of changes in cognitive ability."*

Estonian Data:

There's a lot of studies pertaining to the Estonian data and the situation is complex and somewhat contradictory. Source 267 along with source 264 analyzed the Estonian data and found that measurement invariance was violated. Shiu et al. 2013 [268] conducted an IRT-analysis and found evidence of a genuine increase in all but one subtest with substantial heterogeneity. Must & Must 2013 [269] (followed exactly after Shiu et al. 2013 in the volume and issue) found that much of the Flynn Effect in Estonia was explained by changes in test-taking behavior. On a related note, source 270 also analyzed the Estonian data and found evidence that it was due to increased guessing (Brand's hypothesis) and that controlling for guessing also increased the negative relationship between g-loadings and Flynn Effect score gains. Must & Must 2018 [271] found that the number of invariant indicators was only 23% between the 1933/36 and the 2006 cohort. Using only invariant items, there was no clear evidence of a long-term rise. However, they were able to conclude that the younger cohort was faster and there was a -0.89 correlation between test-taking speed and scores on non-invariant items.

Source 272:

This study used the GSS wordsum and found that using IRT score, there was no statistically significant change in any era for wordsum scores. MI was tenable across time, but IRT scores were used as they're better than sum-scores for a variety of reasons such as handling floor and ceiling effects.

Source 273:

This study used an extremely large (1.7 million) dataset of SAT, ACT, and EXPLORE test-takers. Factorial-invariance was violated across time. The study found evidence that the Flynn Effect functioned the same in the top 5% as it did for the rest of the curve.

Source 279:

This is an interesting one. Using confirmatory factor analysis to test for measurement invariance, partial-intercept invariance was the preferred model. Using IRT, the Flynn Effect was reduced. There was evidence that the Flynn Effect was partially driven by a decrease in the variability of test takers (Rodgers' hypothesis). While it did find evidence of differential item functioning, this wasn't necessarily due to guessing, the title pretty much says it all.

Source 280:

This study examined the Flynn Effect in series completion tests which show very large Flynn Effect gains. In cohorts separated by just 20 years, measurement invariance violations were observed. Bias in intercepts favored more recent cohorts.

Source 281:

Using the three Weschler scales of WISC, WAIS, and WPPSI, this study was able to separate latent vs observed gains in all three. Latent and observed gains had no systematic pattern of which was larger than the other. The amount of invariant indicators varied substantially with the 55% being the highest amount and 10% the lowest. The authors warn against naively assuming that raw-scores are directly comparable. There is evidence of legitimate gains here, but given the very small amount of invariant indicators, the latent factor(s) used in this study are very noisy and generally poor indicators of g (see source [282](#)). Source [281](#) also notes that:

> *"While the amount of invariance did not have an appreciable influence on the score differences in the current study, this is likely because of the simultaneous estimation of parameters for a given age group (Kolen & Brennan, 2004)."*

Kolen & Brennan 2004 is saved as [[283](#)].

Source [284](#):

Scores were compared with the Flynn Effect in the second, third, and fourth editions in the WAIS to be on the same scale across instruments. Measurement invariance was untenable in comparisons of the second and third versions. However, strict MI was tenable comparing the third and fourth versions. Between the third and fourth editions, there was no change in domain-specific factors. There was a change in g of the magnitude of .373 SD units. Presenting evidence of some legitimate gains, the authors still warn against the unwarranted assumption that observed scores are directly comparable.

Source [285](#):

An interesting recent one. A fairly large meta-analysis which showed IRT score declines for spatial-perception in German-speaking countries. The relationship was u-shaped which indicated an initial increase followed by a decline. The decline was even stronger when controlling for publication year and sample type with students obviously showing higher scores. This would indicate that some of the decline was masked by more educated people taking the test.

[The Malleability Of Intelligences:](#)

Also worth mentioning is the malleability of cognitive abilities in general. There is a phenomenon called the "Fadeout Effect" where the small, non-g IQ gains from head start programs fade over time [[305](#)]:

Source [305](#) - Figure 4:

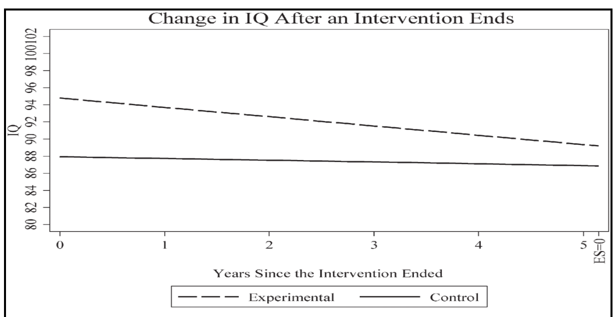
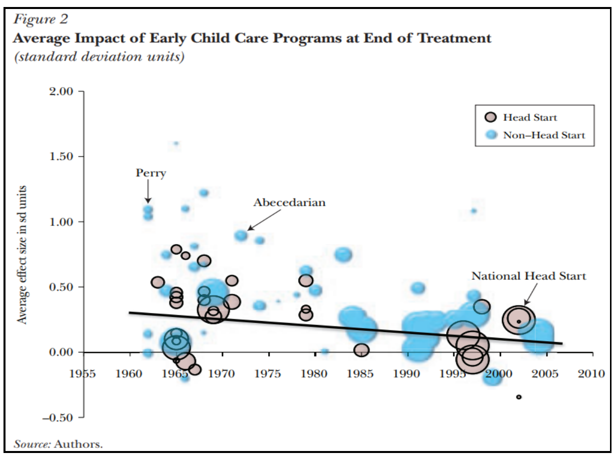


Fig. 4. IQ scores decline after an intervention ends.

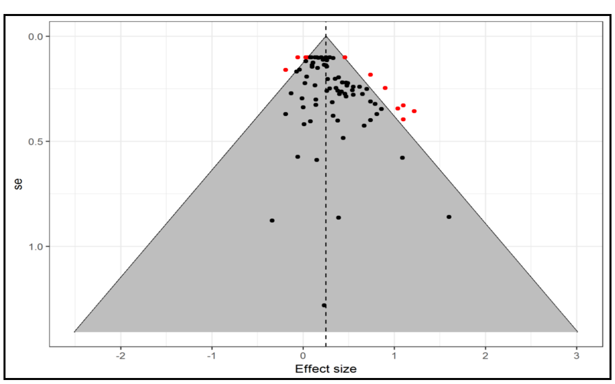
A meta-analysis on the effect of shared book reading on language development also finds the same thing [[694](#)]. As mentioned, the most recent meta-analysis on the effect of schooling duration on intelligence found the gains to fade somewhat with age [[630](#)].

Also worth mentioning is that the effect sizes of the educational intervention programs are inflated by publication bias. A meta-analysis on the impact of early intervention programs on IQ [137] puts the meta-analytic effect at less than half of a standard deviation increase in IQ. From its own report, we see in figure 2 that early intervention programs suffer from the decline effect where the first studies published about a topic with many citations in high impact factor journals are p-hacked, have lower statistical power, and publication bias pushes things towards the desirable results. See source 6 for more on the decline effect.

Source 137 - Figure 2:



A third party [138] put the data into an actual funnel plot and we can see that publication bias definitely inflates the meta-analytic effect size:



Given this, the early intervention literature would likely show the programs to have no effect on IQ with publication bias accounted for, not necessarily that they wouldn't have non-cognitive benefits.

On Heritability And Malleability:

The heritability of the general factor of intelligence is 91% [more here]. Many object to the importance of heritability estimates due to the fact that the heritability of a trait (the proportion of variance in a trait which is caused by variance in genetics) is not necessarily the same thing as the malleability of that trait. In a technical sense, this is true; even if the heritability of IQ were 100%, it could still be possible to raise or lower IQ by exposing the population to environments that no members were previously exposed to.

This being stated, heritability puts a constraint on malleability for the population in question. A heritability of 99% means that 99% of variance would be eliminated if everybody were turned into genetically identical clones. Similarly, a heritability of 99% would mean that 1% of the variance would be eliminated if the environment were equalized. This however does not mean that only 1% of variance can be eliminated by manipulating the distribution of environmental quality. If for example, it were the case that a 99% heritability is what it is because 1% of people are blind, then one may be able to get rid of 50% of the variance in IQ if they remove the eyes of all smart people, that is, if one were to deliberately try to distribute environment unequally, Harrison Bergeron style, in order to fight against the genetic advantages that certain people have.

The "heritability is not necessarily malleability" statement is often stated in ignorance of this. Therefore, many beliefs which are based on it are fallacious.

One other statement which is technically correct, but often used incorrectly, is the statement that it is nonsense to say that somebody's height is x% genetic. This is true, such a statement is nonsense. However, if we had two people with different IQ scores, causal hypotheses about the reasons for the difference are a reasonable thread of inquiry. Moreover, IQ is a particularly dumb topic in which to bring this point up; IQ scores, by design, tell us how people rank in terms of IQ. IQ is standardized such that the population mean is set to 100, and the standard deviation is set to 15. Bob having an IQ of 115 means that Bob is 1 standard deviation above the mean in IQ. In other words, he is smarter than about 84% of people. To merely state Bob's IQ score is to state his rank order in terms of the standardization sample that the test was standardized on. Thus, to ask what percentage of Bob's IQ is genetic is a reasonable question because by test construction, the question is to ask why Bob's rank is what it is.

The Biology Of Intelligence:

The Heritability Of Intelligence:

Large scale reviews of hundreds of twin studies looking at the simple overall population heritability for full scale IQ scores in Western samples show most studies putting the heritability at about .5 (50%) for children [111, & 308]. The following data is from source 111:

IQ Similarity Of Relatives Who Grew Up In The Same Home:

Relationship:	IQ Correlation:
Identical Twins	.86
Non-Identical (Fraternal) Twins	.55
Normal Siblings	.47
Parent-Offspring	.42

IQ Similarity Of Relatives Who Grew Up In Different Homes:

Relationship:	IQ Correlation:
Identical Twins	.76
Non-Identical (Fraternal) Twins	.35
Normal Siblings	.24
Parent-Offspring	.24

However, the heritability of IQ is a moving target. It rises with age up to about 80% in adulthood [more here]. Different IQ subtests are also more heritable than others, with IQ subtest heritabilities being highly correlated with g-loadings [355, 356, 357, 358, & 359]; this is also the case in chimpanzees [183]. The heritability of g in particular is .86 [493], and is .91 after correction for measurement reliability [843, more here]. Heritability is the percent of variance in phenotype between individuals which is caused by variance in genotype [more here], and our heritability estimates are calculated upon nationally representative samples [more here].

Generalist Genes:

Everything discussed on the validity of heritability [more here] is applicable to a statistic called the genetic correlation. Basically, to calculate a genetic correlation is to answer the question of the extent to which the genotype involved in phenotype 1 correlates with the genotype which is involved in phenotype 2.

Say for the sake of argument that one twin's IQ can be used to predict the second twin's income. If this prediction is more successful in MZ twins than it is in DZ twins, and the EEA is true, then it is known that the genotype involved in IQ is correlated with the genotype involved in income. Alternatively to the twin method, molecular genetic studies can test the degree to which genotypes which are associated with IQ are also associated with income. The genetic contribution to the raw phenotypic correlation can be derived as the product of the genetic correlation and the square roots of the heritabilities of the two phenotypes.

Are the genotypes which influence performance on one IQ subtest the same genotypes which influence performance on the rest? We can answer this question with genetic correlations.

This research consistently shows that the phenotypic correlations between cognitive abilities are mediated significantly and substantially by genetic factors called generalist genes [609, 345, 346, 347, 492, 493, & 951]. For example, a multivariate genetic analysis of general intelligence, reading, math, and language in a sample of over 5,000 pairs of 12-year-old twins [346] showed that genetic factors consistently accounted for more than half of the phenotypic correlations, ranging from 53% to 65%, with a mean of 61% and a mean 95% confidence interval of between 53% and 67%. The genetic correlations between the general factor and the specific abilities are also larger than the genetic correlations between the specific abilities and all the other specific abilities:

Source 346 - Figure 2:

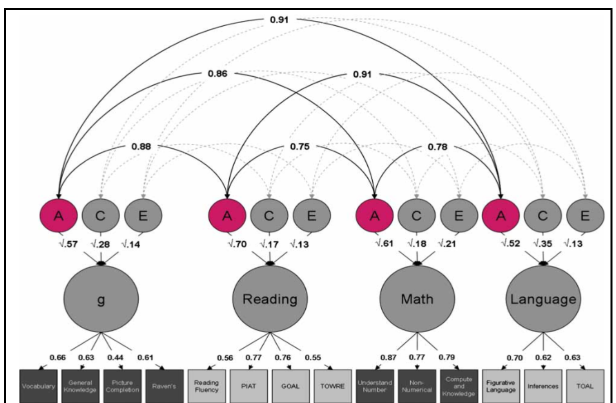


Figure 2. Cross-trait genetic effects. Estimates of cross-trait additive genetic effects (A) are highlighted.

The finding of generalist genes is also supported by evidence from multivariate GCTA [347]. One implication of these findings is that the phenotypic structure of these domains is similar to their genetic structure, as has been shown for example, for the domains of intelligence [348], and personality [349].

This is all of course consistent with the finding that the most heritable subtests are the most g-loaded [355, 356, 357, 358, & 359].

Interestingly, the same thing can be done for environmental effects, and to the extent that shared environmental effects influence intelligence, intelligence being influenced by generalist environmental factors is also supported:

Source 346 - Figure 3:

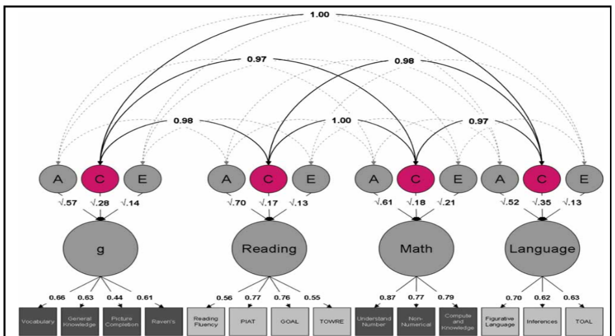


Figure 3. Cross-trait shared environmental effects. Estimates of cross-trait shared environmental effects (C) are highlighted.

For non-shared environment effects, as is predictable, things look more random:

Source 346 - Figure 4:

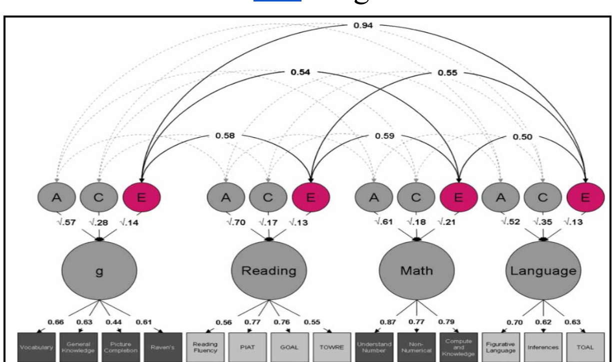


Figure 4. Cross-trait nonshared environmental effects. Estimates of cross-trait nonshared environmental effects (E) are highlighted.

Think back to our example table from earlier:

Variable:	1:	2:	3:
1:	1.0	-	-
2:	1.0	1.0	-
3:	1.0	1.0	1.0
g1	0.5	0.5	0.5
g2	0.5	0.5	0.5
g3	0.5	0.5	0.5
g4	0.5	0.5	0.5

In this sense, complicating things beyond the raw correlation matrix of measured tests in the way previously discussed [more here] is the empirically correct factor analytic solution.

IQ is a highly polygenic trait [more here], meaning that the independent contribution of any single SNP to intelligence test variance is incredibly small. Intelligence is thus mostly explained by millions of tiny general factors, or generalist genes.

The Neuroscience of g:

It should be noted that the field of Neuroscience is still in its early development. Replication is low [156 & 154], lower than many other fields [more here], statistical power, while more relatively acceptable, is still low, and there doesn't seem to be much multivariate research. For example, Haier's book, *The Neuroscience of Intelligence* [172], notes on page 146 that source 173 "is the only imaging study of intelligence to date that investigated both resting-state and task activation conditions in the same subjects".

The attitude of Neuroscientists in general seems to be to ignore individual differences and seems to be that individual differences are just random meaningless noise in the data. They may note that on average, brains light up more in xyz areas when performing some task, but they won't investigate if high IQ means different patterns in activation. They may even say that differences in patterns of activation is evidence that general intelligence is inconsistent and thus not real.

A meta-analysis of 90 functional MRI experiments [360] found test-retest reliability was found to be low (ICC = .397). So the results of an fMRI analysis often do not agree with the results of the same analysis done a second time meaning that the field lacks the statistical reliability needed to map brain activity to behavior.

Neuroscientists also have a high degree of researcher freedom [598]. This is bad for replication and scientific rigor [594]. To expose the degree of freedom, one can give many teams the same dataset and same research questions and tell them to analyze the data how they see fit, as previously done for football racism [597]. Source 598 analyzed the impact of flexibility on fMRI results by giving 70 research teams the same 9 hypotheses to test. There was only one hypothesis with mostly consistent support. For it, 84% of teams found a p value below .05.

Despite all of this, Neuroscience enjoys public perception of higher scientific rigor than psychology [599].

This being stated, there are some replicable neural correlates of g.

Brain Size:

One idea that the likes of Stephen J. Gould heavily ridiculed was the idea of brain size being related to intelligence [257]. He attacks the early work as being unobjective for using flawed methods like measuring skull volume by filling with lead shot pellets where the experimenter can fit more or less into a skull depending on how much force they apply when pushing it in. However, data trumps eloquent writing, and in modern day, brain size can be accurately measured with structural MRI. There are meta-analyses covering dozens of studies about the relationship between intelligence and brain size measured via MRI, and a relationship is consistently found [361 & 362]. Source 362 was the better, larger, more recent meta-analysis which checked for publication bias and it found a smaller relationship than source 361 did, a correlation of .24 rather than one of .63. Though source 362 is the better review on most things, source 361 was able to show that the general factor was most associated with brain size while source 362 did not test for this. Source 362 does however seem to vindicate the result by showing that the better indicator of general intelligence was more associated with brain size. Accordingly, corrections to source 362's dataset yields a correlation of .4 [654]. All in all, brain size seems to be able to explain 6% of variance in intelligence [362].

This relationship is causal. Within family differences in IQ are also related to within family differences in brain size [361]; this finding is a control for shared environment. Moreover, multiple studies have shown a genetic correlation between brain size and intelligence [363, 364, 683, & 954] meaning that the same genotype which explains brain size largely explains intelligence. The heritability of brain size is also 87% [851]. Furthermore, brain size and intelligence both follow the same pattern of increasing until the mid 20s and then declining in old age [361]. This is also consistent with evolutionary evidence of brain size increasing as hominids got closer to being modern humans [366].

Connectivity & Folding / Gyrification:

Both gray matter volume and white matter volume are related to intelligence [370]; gray matter slightly more so. Gray matter is located towards the surface area of the brain while white matter fills the interior. White matter connects gray matter together and transfers information. Perhaps folding (gyrification) in the brain allows more gray matter to be connected by less white matter. It has been suggested that folding could be related to intelligence [370]. Source 371 found a relationship between IQ and gyrification, that the associated areas are consistent with Haier's P-FIT, that the associated areas are highly consistent across samples, that gyrification can account for 11.5% of variance in the adult sample (N=440), and 5.2% of variance in the child sample (N=662). Source 392 looked at individual relationships at thousands of different points in the brain with 2,882 people. It calls the relationship minimal since the average independent effect of each point was a correlation of .05 in one sample and .1 in another. All effects of gyrification add up to explaining 11% of variance which it also calls minimal. Source 392 also showed that that the relationship between intelligence and gyrification was genetically mediated, and that this finding was statistically significant even for all of the small points of gyrification. Source 372 found that white matter tract integrity explained 10% of variance in intelligence.

Grey & White Matter Density:

In addition to the association with pure volume, gray matter density, white matter density, and neuron count are associated with higher IQ [862 & 665], and the associations are genetically mediated [665].

Plasticity:

Higher intelligence is related to higher brain plasticity and the relationship is genetically mediated [373].

Cellular differences:

Other proposed biological mechanisms for intelligence include differences in various cellular level qualities such as mitochondrial efficiency or pH level [865, 863, 864, & 367].

Neural Efficiency:

The neural efficiency hypothesis postulates that smarter people display less cognitive activation, as measured by glucose metabolism [374]. It's thought that smarter people can do more mental work with less energy, thus being more efficient. Source 375 extensively reviewed 27 studies confirming this finding using methods such as PET scans, EEG, and fMRI. However, fMRI and EEG studies reveal that task difficulty is an important factor affecting neural efficiency; smarter people display neural efficiency only when faced with tasks of subjectively easy to moderate difficulty, but no neural efficiency can be found during difficult tasks. In fact, smarter people seem to invest more cortical resources in tasks of high difficulty. Source 1154 was also able to account for 20% of variance in IQ with resting state fMRI data.

Multiple Traits:

A popular attitude among Neuroscientists seems to be that because this, that, or the other neural variable, by itself, only explains a small portion of variance in intelligence, that not very much of the variance in intelligence can be accounted for with neural variables. This is obviously fallacious because, like the genes, there are many neural variables which are associated with intelligence, so this fact inherently limits the amount of variance that each individual neural variable can account for. While many variables are subadditive, a handful of papers have been able to predict 20% of variance in intelligence with brain variables [369, 1154, 1155, 1156, 1157], and there is diversity among which measures used, so a fictional paper utilizing every known neural variable would likely be able to account for more variance.

P-FIT:

Overall, if you want more depth on neuroscience findings, read Richard Haier's book [172]. One of the main things Haier argues for is his parieto-frontal integration theory (P-FIT) of intelligence, the first evidence for which came from his review of 37 neuroimaging studies [368]. The finding is basically that a distributed network throughout the brain, and mainly in the parietal and frontal lobes are consistently involved in intelligence and perhaps that the connectivity within it is associated with intelligence.

Neuroscience & Sampling Theory:

Sometimes people reference a paper called *Fractionating Human Intelligence* [595] as proof of sampling theory explaining the positive manifold. Aside from the problems with the paper that Haier points out [596], it's worth pointing out what the paper actually does without the gish gallop.

The authors take a small sample, IQ test them, varimax the data into two highly correlated intelligence factors (let's call them i1 and i2, the real names were longer), and get two brain factors from the brain data which are somewhat negatively correlated (let's call them b1 and b2). The authors show that i1 and b1 correlate at ~.7, that i2 and b2 correlate at ~.7, and that the two brain factors are slightly negatively correlated. A theoretical simulation of sampling theory is shown, and it is shown that the "two" varimaxed intelligence factors both correlate with all of the first order tests. It is said that this sort of looks like sampling theory explaining the results.

The implication seems to be that the correlations between the two brain factors and the two intelligences are reason to interpret the data as support for sampling theory. The problem is that they never show the correlations between i1 and b2, or between i2 and b1.

They also found a g factor before rotation, but didn't show the associations between it and the two brain factors. What could easily be happening is that both brain factors affect all aspects of intelligence generally. It makes as much sense to lump them both into a single g factor as it does to lump brain size, brain folding, white matter efficiency, etc, into a single variable and call that general intelligence. Maybe the sampling theory advocates would take this as vindication that multiple brain variables explain the g factor and that the g factor isn't a single brain variable, but the thing is that in general, all the brain variables, though themselves independent of each other, all affect intelligence in a generalized way.

The Validity Of Heritability:

Let's bake a cake. What percentage of the cake's traits are caused by the ingredients? What percentage of the cake's traits are caused by the mixing, baking, etc? These are nonsense questions. Some better questions would be to bake two cakes and compare their reasons for turning out differently. Was cake 1 baked longer and at a lower temperature than cake 2? Or does cake 1 have the ingredients of a chocolate mousse cake as opposed to cake 2 which has the ingredients of a carrot cake?

This brings us to heritability; the questions we ask should be the same. Heritability figures tell us the proportion of phenotypic variance in a trait (such as intelligence) which is caused by variance in genotype. A useful way to think of the heritability of a trait is that it tells us the percentage of a trait's variance that would go away if everybody were born as genetically identical clones of each other.

Conflict With Common Sense?

Critics of heritability sometimes say that the correlation between phenotype and genotype is blindly assumed to be genotype causing phenotype even if environment is what causes genotype to correspond to phenotype, thus redefining the term environment, which is traditionally considered to be a very broad array of effects, as being something very different from what common sense would define "environment" to be. For example, the passage below, characteristic of source 480, gives the classic analogy of redhead oppression:

Source 480, Pages 66-67:

> *"If, for example, a nation refuses to send children with red hair to school, the genes that cause red hair can be said to lower reading scores... Attributing redheads' illiteracy to their genes would probably strike most readers as absurd under these circumstances. Yet that is precisely what traditional methods of estimating heritability do. If an individual's genotype affects his environment, for whatever rational or irrational reason, and if this in turn affects his cognitive development, conventional methods of estimating heritability attribute the entire effect to genes and none to environment."*

This conceptual criticism of heritability is fair as far as it goes conceptually, but this is a serious distortion of the way twin studies are used to estimate heritability and is thus completely divorced from the methodological reality of the field of quantitative genetics.

These sorts of gene-environment interaction effects have been tested for with foolproof methods, and they do not occur [more here]. To understand the evidence for this claim, it must first be understood what the twin methods themselves conceptually aim to do and how.

There are two twin study methods, twins reared together (also known as the classical twin method), and twins reared apart. The method of twins reared apart is what most people think of when they hear the term “twin study”. In it, one raises identical twins in different environments, measures the similarity in environment that the twins experience which the general population does not experience, and subtracts that from the correlation between identical twins to get the heritability estimate. Subtract the heritability estimate from 1, and one is left with the contribution of environmental effects.

The method of twins reared together, a frankly better method, exploits the difference in correlations between identical twins (referred to as monozygotic, or MZ twins) and non-identical fraternal (referred to as dizygotic, or DZ) twins. An assumed difference between the MZ twin class and the DZ twin class is that the MZ twin class has a kinship coefficient of 1 while the DZ twin class has a kinship coefficient of 0.5, meaning that MZ twins are 50% more genetically similar to each other than DZ twins are. So, to estimate heritability, one takes the difference in correlations between the two twin classes and divides the result by the difference in kinship to get a heritability figure. For the sake of argument, say that the height of MZ twins raised in the same environment correlates at 0.8, and the height of DZ twins raised in the same environment correlates at 0.4. The difference in correlations is 0.4, and the difference in kinship is 0.5. 0.4 divided by 0.5 equals 0.8, so in this case, the heritability of height taken from the twins reared together method is 80%. The reason for the division is that given the difference in kinship, the difference in correlation is assumed to extrapolate to mean that a difference in kinship of 1.0 rather than 0.5 would produce an increase in correlation of 0.8 instead of 0.4. In other words, it is assumed that if the difference in kinship is doubled, then the difference in correlation is doubled. The twins reared together method is better because non-adopted twins are much more common and representative of normal people than adopted twins; this makes the twins reared together method cheaper to do because of the larger supply of twins, and also more representative of the general population because the twins reared together method does not have to wrangle with adoption agencies and ethical research practices which cause range restriction of the environments that their heritability figures apply to. The twins reared together method can also differentiate between two types of environmental effects: shared and nonshared environment; the names are self-explanatory. One can take the correlation between MZ twins raised in the same family, subtract the genetic component, and the resulting portion of the MZ correlation which is not explained by genes is referred to as the contribution of shared environmental effects. The extent to which MZ twins reared together do not correlate with each other at all is called the unshared environment. A is short for genetic, C is short for shared environment, and E is short for nonshared environment. Usefully, twins reared together studies and twins reared apart studies, by design, always explain 100% of phenotypic variance within the population being studied; A + C + E = 1.0.

Method Assumptions:

By now it should already be clear why the two twin methods are much more sophisticated than simply calling the correlation between children and their parents a genetic effect by redefining certain environmental effects as genetic effects since they correlate with genotype; the classical twin method, at bare minimum, performs a sibling fixed effects control.

This being stated, environmentally driven gene-environment correspondence effects are not yet completely conceptually off of the hook. For example, in the classic redhead oppression example, both twins in an MZ pair are either both redheads, or are both non-redheads. The increase in kinship increases the chance that both will experience the exact same amount of oppression, and thus causes a difference in the phenotypic correlation even though that is not a genetically caused effect. The same applies to the method of twins reared apart. The same also applies to any molecular genetic evidence which looks at how actual, observed genotypes (genes, SNPs, copy-number variants, etc) differ among people and is measurably correlated with phenotype among random, unrelated individuals from completely different families.

However, pointing out this conceptual possibility, and taking it, by itself, as justification to ignore all heritability findings, is not justified. Yes, the similarity of monozygotic twins reared apart (MZA) is indeed taken to be a direct measure of heritability, but only to the extent that causally relevant environments of these twins are uncorrelated ("relevant environments" being defined as environmental variables that some people are appreciably exposed to in real life and which causally correlate with phenotype without genetic confounding). As has routinely been emphasized in the literature, the inference of heritability from MZA is considered legitimate only to the extent that there are no common environmental influences that could explain the concordance between the MZA twins reared apart, and to the extent that any common influences which do exist are accounted for.

The same applies to the method of twins reared together. Some of the phenotypic correlation between identical twins raised in the same homes may be accounted for by SES, or whatever environmental variable, but the twins reared apart method is concerned with the difference in correlations rather than the raw correlations.

So, to affect heritability figures, the effect that environment has on the MZ correlation must not be the same effect that it has on the DZ correlation. In other words, if net environmental influences which affect MZ twins are stronger than the environmental influences which affect DZ twins, then the difference in correlations will be larger than a genetic effect which would artificially inflate heritability figures. However, this criticism boomerangs onto twin method critics because it is also conceptually possible that environmental effects which affect MZ twins could be weaker than the environmental effects which affect DZ twins, which would mean that the difference in correlations would be smaller than the what genetic effects "want it to be", and that heritability estimates would be biased downwards. The assumption that environmental effects have the same magnitude of causal contribution to phenotypic correlations for both MZ and DZ twins is called the equal environments assumption (EEA), an assumption which is well supported [see more].

It is also possible that the equal environments assumption is a completely true assumption for normal variation, but that for specific group differences like the redhead example, there are specialized equal environments assumption violations that don't apply to the general population or to the within-group heritabilities, and have to be investigated separately. For the question of the between-group heritability of the Black-White difference in g, these specialized violations are known as x-factor hypotheses; the redhead oppression example is generally brought up by those concerned with the Black-White differences. This is not relevant to the overall national heritability figures, so evidence pertaining to it won't be discussed in this chapter, but evidence pertaining to it will be discussed in [chapter 7].

The Sociologist's Fallacy:

Sometimes it is asserted that MZ twins have more similar environments than DZ twins by various metrics, thus calling the equal environments assumption into question. The thing to remember about the equal environments assumption is that it is concerned with causality. So, if an environmental variable isn't even correlated with the phenotypic variable at all, then the greater similarity of MZ twins in terms of that environmental variable is obviously etiologically irrelevant. Second of all, if correlated with phenotype and genotype, the increased environmental similarity has to **cause** genotype to correlate with phenotype rather than the other way around. Say for the sake of argument that genotype causes intelligence and that intelligence causes educational attainment: Is "Environment" correlated with phenotype and genotype? Absolutely. Does "Environment" cause the correlation between phenotype and genotype? Not so fast. Is education environment or phenotype? Is it both? When looking at the heritability of intelligence after accounting for differential correlations with education, it could very well be that all that the results are saying is "When the effects of genotype on phenotype are controlled for, genotype has no effect on phenotype!" The sociologist's fallacy is committed when the raw correlational requirements are met, but the causality of the differential correlation is claimed to be entirely from environment to phenotype without evidence. Causality must be tested to confirm an EEA violation.

If correlational requirements are met, the direction of causality can be tested in the old fashioned ways: testing phenotypic responses to experimental manipulation of the environmental variable, longitudinal cross-lagged path models, etc.

One thing to consider is that if a purely environmental variable is found that causally, differentially amplifies correlations between the twin classes, it could very well be that other purely environmental variables also exist which drive heritability in the opposite direction. Such opposing effects should be assumed to cancel each other out in lack of evidence that effects in one direction are more important than effects which go in the other direction.

Gene-Environment Interaction:

Sometimes, some of the variance in a trait, such as good/bad behavior in children [870], can be apportioned to neither genetic nor environmental effects, but to a complex interaction of the two. This happens when phenotype and environment have bidirectional causality. Let's say for the sake of argument that MZ twins correlate at .8 in disruptive behavior, and that DZ twins correlate at .6 in behavior. Let's also say that 50% (.1) of the difference in correlations (.2) is mediated by differential similarity in parenting style. Some of the difference in correlation is still unmediated by parenting style, so is a pure genetic effect. But MZ twins are treated more similarly in parenting style than DZ twins; why? Well, causality must be tested. If causality between phenotypic similarity and environmental similarity is bidirectional, then this is a gene-environment interaction (GxE) effect. This may happen if poor behavior causes parenting to become harsher and harsher parenting causes behavior to become poorer in a feedback loop. Again, just like with the sociologist's fallacy where an effect cannot be assumed to be a purely environmental effect without evidence, an effect also cannot be assumed to be a GxE effect without evidence. If the EEA is tenable, that is, if causality is squarely from phenotype to environment with environment having no causal effect, then a GxE effect does not exist. The EEA is indeed generally tenable, and most GxE effects do not replicate [more here].

Another class of gene-environment interactions certainly happens everywhere, but does not cause variance between individuals: Imagine that all oxygen is removed, leaving us with only hydrogen, nitrogen, etc. Suddenly, everybody would die, nobody would be able to answer questions anymore, and strength would drop to zero. Though existing variance was somewhat genetic prior to removal, the variance in strength between oxygen and no oxygen is entirely environmental. As a more interesting example, a contrarian person in Maoist China may spite the Chinese government and become a Christian. However, the same person in Medieval Europe may spite the Catholic Church by becoming a Satanist or an Atheist. This change in religious belief is environmental, but individual variance in contrarianism may not be so environmental.

Assortative Mating:

Another potentially biasing assumption of the twins reared together method is the assumption of the magnitude of the difference in kinship. That DZ twins have a kinship of 0.5, is based on the random mating assumption. It could be that marital partners seek out people who are similar to one's self while dating. If this means that marital partners have more genetic similarity to each other than two random individuals from the population will have on average, this is known as assortative mating and it means that on average, any children they have will have a kinship greater than 0.5. Assortative mating would mean that the DZ kinship coefficient is larger than 0.5, which would mean that the difference in kinship is smaller than 0.5, which would mean that heritability figures were underestimated. The evidence pertaining to assortative mating does indeed show that this happens [more here].

“Identical” Twins:

Identical twins aren’t necessarily 100% genetically identical (b/c e.g. mutations), and to the extent that these genetic discrepancies affect IQ, they are usually erroneously treated as the nonshared environment [more here].

Heritability Between Who?

It is important to make sure that we measure the heritability of differences between the right people. This isn’t an issue with accurately measuring the heritability for a sample, but we must get the sample right if we are to generalize a heritability figure to the general population. So do we measure the right people? Yes, nationally representative samples, such as ones that straightforwardly use national militaries or school systems, come up with the same heritability figures as the rest of the literature. Additionally, between-poor heritability is the same as between-rich heritability. See evidence on sampling [here].

Heritability Of What?

This isn’t an issue with accurately measuring the heritability of whatever measure, but of making sure that we are choosing the right things to measure the heritability of. IQ tests aren’t 100% reliable; taking a test twice will result in two slightly different scores. The measurement error (unreliable variance) is solely caused by nonshared environmental effects, and the reliable variance of IQ is more heritable than the unreliable measurement error (g is also more heritable than the specific abilities) [more here].

Twins Reared Apart:

So the twins reared together method is vindicated by the assumption tests, but what about the twins reared apart method? Does society treat twins similarly, regardless of whether or not the twins know each other, because the twins look similar? Some evidence from twins reared together is relevant here; one good operationalization of this is physical attractiveness since attractive people are generally treated better, but attractiveness is uncorrelated with IQ [more here]. The similarity of identical twins reared apart also cannot be explained by non-total separation of the twins [more here].

“Find The Genes!”:

The same assumption violations (environment causing genotype to correlate with phenotype) are also just as conceptually possible for any attempts to calculate the heritability of a trait using molecular genetic methods that look at actual SNPs, copy-number variants, genes, etc, and how actually observed genetic variation is measurably correlated with phenotypic variation for people from different families. Without even taking into account the types of genetic effects which the twin studies can measure but the molecular genetic ones can't (e.g. rare variants, exotic variants, non-additive effects, etc) [more here], the twin studies are actually better than the molecular genetic evidence for assessing causality because in the twin studies, all of the assumptions can be tested, and if they aren't true, any violations of assumptions can be precisely corrected for in the calculation of heritability figures.

Genetic Correlations:

Another useful thing to mention is that instead of just calculating the heritability of a specific trait, everything discussed thus far can also be applied to a statistic called the genetic correlation. Say for the sake of argument that one twin's IQ can be used to predict the second twin's income. If this prediction is more successful in MZ twins than it is in DZ twins, and the EEA is true, then it is known that the genotype involved in IQ is correlated with the genotype involved in income.

Alternatively to the twin method, molecular genetic studies can test the degree to which genotypes which are associated with IQ are also associated with income. The genetic contribution to the raw phenotypic correlation can be derived as the product of the genetic correlation and the square roots of the heritabilities of the two phenotypes.

The Convergence Of Methods:

In addition to twins reared together, twins reared apart, GWAS, and GCTA methods, heritability is further confirmed via censuses, identity by descent, and by virtual twin studies where unrelated children of similar age are adopted into the same family in a way that resembles normal siblings [more here]. With all methods converging upon the same finding, and the tenability of the assumptions behind these methods, the evidence behind heritability can be taken as very reliable.

Conclusions:

All in all, if assumption violations are taken into account, heritability figures would have no such conflict with common sense definitions of environmental effects as those who peddle the redhead oppression analogy would have us believe they do. Indeed, heritability figures should actually rise somewhat when all assumption violations are accounted for.

Assumption Violations:

The Equal Environments Assumption:

The Equal Environments Assumption (EEA) was first tested in source 296 which measured the degree to which parents treated twins the same way, the degree to which they were dressed alike, whether they had been put into the same classes, whether they slept in the same room, etc. They then measured the correlation between how similarly the twins were treated by their parents to how similarly they were in IQ. The paper found that increased similarity of treatment predicted almost no increased similarity in IQ.

Since then, source 117 comprehensively reviewed the evidence on the EEA, and did its own analysis with the most comprehensive set of controls to date. Correcting for EEA violations adjusted heritability figures downwards only very modestly; heritability figures, at most, go down by about 10%. However, this line of research is often merely correlational: Correcting twin class correlations for "environmental" similarity should be done with caution because corrections may commit the sociologist's fallacy [more here]. The entire goal is to root out **causality**. Phenotypic similarity may cause "environmental" similarity rather than the other way around. For example, the evidence on assortative mating [more here] shows that people want to live around other people who are similar to them, and that this also influences the rate at which twins choose to live together. The various supposed EEA violations should have their respective environmental variables tested for phenotypic causality to establish trait relevance.

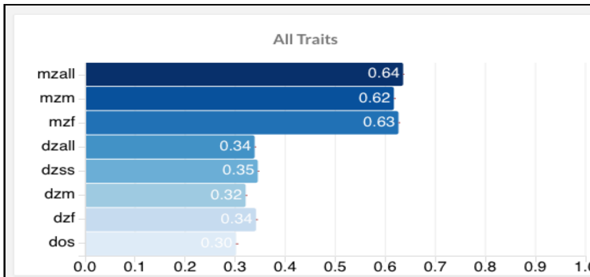
Should such differential similarity in environment be in terms of trait-relevant variables, it could still be the case that twins create their environments, and that genotype affects phenotype by causing environment. To rigorously test the classical twin method for genetic causality, we must ask why identical twins would have more similar environments than fraternal twins if not for reasons of genotype creating environment. This leaves us with essentially three options:

1. In terms of physical appearance, identical twins look more similar to each other than do fraternal twins, the phenotypic similarity is caused by people discriminating based on appearance.
2. The linguistic label of "identical twins" causes people to apply more similar treatment to such twin pairs than they do to fraternal twin pairs.
3. Identical twins have more similar prenatal environments than fraternal twins have, and this causes greater trait similarity.

Option 3 is not an issue; identical twins are not more similar than fraternal twins because of prenatal effects [more here].

Option 2 is also not an issue; identical twins who are accidentally classified as fraternal twins throughout their entire lifetime actually turn out more phenotypically similar than correctly classified twins [298, & 297] (perhaps the label "identical" makes people strive for individuation).

Option 3 is a bit more tricky to assess, but we have a few things we can look at. First, the review cited earlier [117] included tests for physical appearance. Second, it is well established that physically attractive people are thought of as more intelligent, yet attractiveness is slightly negatively correlated, if not uncorrelated, with IQ [407]. Given this, we don't even need to assess the causality of such a correlation. Third, we have the sanity test of sex differences: Same-sex twins look more similar, are more likely to be treated similarly by their parents, are more likely to wear similar clothes, are more likely to spend time together, etc. It should be noted that any effects on twin class correlations could just be a reflection of the effects of innate sex differences, but regardless, this can be readily investigated with data from a recent, gargantuan meta-analysis of every twin study ever done on thousands of traits and millions of twin pairs [490]. For all traits, the correlations are as follows:

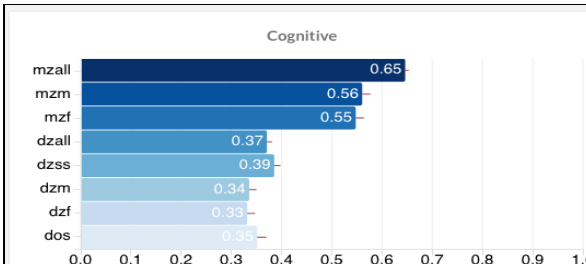


TWIN CORRELATIONS

All Traits

	Est.	SE	Ntraits	Npairs
mzall	0.636	0.002	9568	2563627
mzm	0.617	0.004	4518	1070962
mzf	0.626	0.004	4360	1171841
dzall	0.339	0.003	5220	2606252
dzss	0.345	0.003	6108	1752951
dzm	0.321	0.003	4412	1039238
dzf	0.342	0.004	4255	1068562
dos	0.302	0.005	2342	898610

For cognitive traits, we see the following:



TWIN CORRELATIONS

Cognitive

	Est.	SE	Ntraits	Npairs
mzall	0.646	0.007	931	288866
mzm	0.560	0.017	248	54974
mzf	0.548	0.016	199	38940
dzall	0.371	0.010	454	304719
dzss	0.385	0.012	545	116058
dzm	0.335	0.015	228	50647
dzf	0.331	0.016	186	28929
dos	0.351	0.018	155	37645

As we can see, sex effects are dwarfed by zygosity effects. For assessing the impact of these differences in correlation of heritability coefficients, it should also be noted that only ½ of fraternal twin pairs are mixed-sex.

Moreover, source 531 meta-analyzed sibling pairs and all combinations correlated equally at .49.

Further evidence for the tenability of the EEA includes sources 354, 486, 487, 488, and 485.

Gene-environment interaction effects, especially novel ones, also mostly fail replication [868 & 869], and are inflated by publication bias [868]. This has led to top journals requiring replication of novel GxE effects before papers are considered for publication [868].

-The Heritability Of "Environment":

Several "environmental" variables which correlate with IQ, and which are fallaciously assumed to causally influence IQ, are themselves highly heritable.

Source 624 puts the heritability of IQ at 66%, the heritability of income at 42%, and the heritability of educational attainment at 40%. A review of 19 twin studies [695] also puts the heritability of income in the USA at 41%. Source 324 meta-analyzed data on more than 13,000 twins and put the heritability of GCSE scores at 62%. Source 325 meta-analyzed 34 twin studies from 9 nations and found that 40% of variation in educational attainment was attributable to genetics. Source 326 found that lifetime income had a heritability of 24% for women and 54% for men. Source 326 also reviewed 19 previous samples from which the heritability of income has been estimated. The typical finding is that about 42% of income variation is caused by genetics while about 9% is explained by shared environmental effects.

Source 350 puts the heritability of independent reading at .62 for 10 year olds and .55 for 11 year olds. Source 351 puts the heritability of potato consumption by men at .68, the heritability of vegetable consumption at .24, and red meat at .34. Source 352 put the heritability of voluntary non-sports exercise at 0.63 for males and 0.32 for females, and the heritability of sports exercise at 0.684 for males and 0.398 for females. This replicated source 353 which found the heritability of sports exercise at 0.83 for males and 0.35 for females, and non-sports exercise at 0.62 for males and 0.29 for females. Source 354 gave an overall heritability of exercise of 0.49 and showed that the EEA is tenable for exercise. Most psychological traits in general have substantial genetic components [308].

Here is the degree of genetic mediation for the relationship between IQ and SES:

Age	Correlation	% Genetic Mediation	Source #
7	.31	94%	624
12	.32	56%	624
16	.50	50%	417

Source 330 did the same for education and found a genetic correlation of .95.

Assortative Mating:

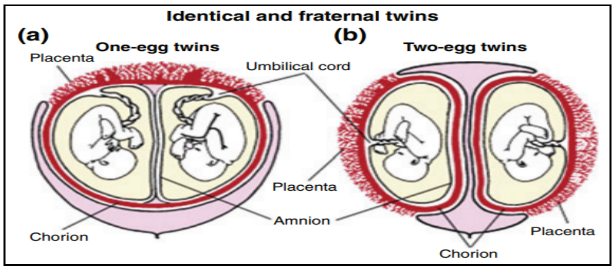
This is the strongest violation of the assumptions that go into heritability estimates. There is a phenomenon where people like each other more when they are more genetically similar:

- Marital Partners are psychologically [312] and genetically [316] similar to each other.
- Friends are genetically similar to each other, and the genetic similarity of the communities that friend groups are contained within does not account for all their similarity [307].
- Pretty much all psychological traits have at least some genetic component [308].
- Friends are most similar to each other in terms of the most heritable traits [309].
- Similarity of personality is predictive of successful marriage [313], and the more heritable traits are better predictors [310].
- If you ask somebody to imagine a fictional person who is similar to themselves in various ways, the more heritable the trait in question, the more the person will think that they would like the fictional person [311].
- The friends of one twin are similar to the friends of the counterpart twin. This trend is stronger in identical twins than in non-identical twins. This lets us directly calculate the heritability of choice in friends. Heritability is .31 for choice of spouse, and .21 for choice of friends [309].
- The fact of assortative mating is robust to various controls, and assortative mating selects upon intelligence [314, 315, & 316].
- There is a positive association between kinship and fertility. Historically, in Iceland, the ideal was 3rd degree cousins [317].
- One piece of evidence which tried to test the EEA is also relevant to assortative mating. Sources 483 and 484 show that MZ twins who have greater contact with each other have more similar personalities than MZ twins who are less in touch. This seems convincing on its face, but this is just a classic example of the Sociologist's Fallacy. It was thought that this is a violation of the equal environments assumption, but as it turns out, twin similarity causes cohabitation rather than the other way around [485]; more similar twins want to live together.

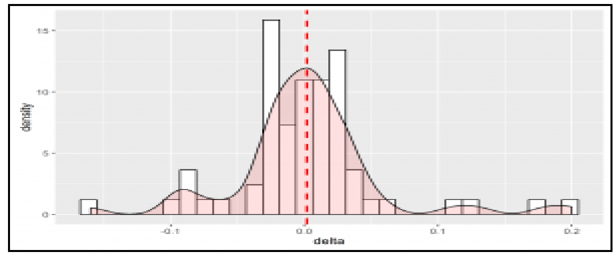
Obviously, correcting for assortative mating would mean that non-identical twins are more genetically similar than previously expected, which means that a smaller than previously supposed increase in genetic similarity is what has been producing the previously observed increases in phenotypic similarity the entire time, meaning a downward bias for heritability figures.

Prenatal Effects:

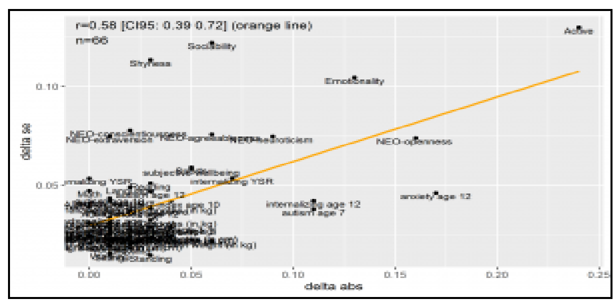
Many MZ twins share the same placenta and have a single chorion. What if more similar womb environments are part of the cause of increased phenotypic similarity?



About 1 in 4 MZ twins do not share a single chorion and so have separate placentas. A large body of evidence shows that "monochorionic" (MC) monozygotic twins are no more similar to each other than "dichorionic" (DC) monozygotic twins are to each other [299]. But aren't some traits in this study affected? Technically, but a study which examines 100 traits would likely find positive and negative effects for a couple of random traits due to random sampling error even if no effects actually existed. The proper investigation is to look at all effects at once. The following analysis simply calculates the correlation between MC twins minus the correlation between DC twins for all effect sizes in the supplementary materials of source 299 with the x axis being effect size and the y axis being statistical power:



The mean is 0.00, tightly clustered around 0.0, and evenly distributed around the meta-analytic effect size. In fact, we can go further. A null model + sampling error model also predicts that the larger effects in either direction should be the less precisely measured effects. So, here are the standard errors of the deltas plotted against the absolute effect size:



Here is the data [626 warning, auto-download] and code [627] for the above two tables.

One could invoke the trait specific context defense that perhaps prenatal effects matter for some traits but not others, but this is usually a post-hoc argument levied by those whose favorite pet effects fail to replicate, so we should be skeptical. For IQ in particular, such effects are small and inconsistent, if at all existent [299 & 625]. If existent, we know that these influences fade with age given the convergence in similarity between DZ twins and normal siblings [318 & 532], so MZ-specific influences should as well. There also seem to be signs of this in the chorionicity tests as well [625]. Of final note regarding the importance of chorionicity over the lifespan is that even if it were a given that chorionicity effects had persistence, this would not be able to explain the rise in rMZ-rDZ differences generally found with age [more here].

Furthermore, if the prenatal environment matters much for IQ in adulthood, then presumably, there would be lasting effects of prenatal interventions. However, the evidence here is scant [629]. Also worth noting is that maternal genotype may influence the prenatal environment [628].

Non-Total Separation:

Some would argue that prenatal effects aren't the only bias in the adoption method. It is argued that twins are often adopted considerably after birth, and so contrary to what adoption studies assume, they have abnormally similar shared environments to some degree.

It is true that some adoption studies have had less than perfect separation criteria, but multiple studies have shown that the amount of time that twins adopted into separate homes spend together prior to a study does not impact their IQ similarity and so does not inflate heritability figures [481 & 482].

"Identical" twins:

Monozygotic (MZ) twins aren't necessarily completely genetically identical; one twin may carry some mutations which the other lacks, and twin studies would model these as nonshared environment effects [844].

"Find The Genes!"

Hopefully, by now it should be clear to the "Find The Genes!" people that the twin studies work just fine, but many people have a vague impression that molecular genetic evidence is somehow comprehensively better to the point that the twin studies, that quantitative genetic evidence, is worthless. This attitude is deeply mistaken. Do not take this as reason to be against the use of molecular genetic evidence on sheer principle, it's just that molecular genetic evidence has some limitations which should be noted.

The main problem with looking at things through molecular genetic evidence is the sheer amount of statistical power which is needed for it. There are over 3 billion base pairs in the human genome with roughly 40% of the genome involved in cognition [672 & 673]. Each nucleotide is its own variable which has to be considered individually. IQ is an incredibly polygenic trait [329 & 331] meaning that millions of individual Single Nucleotide Polymorphisms (SNPs) have an effect, so the independent effect that any single SNP has will be incredibly small. The smaller a variable's effect size, the more statistical power you need to accurately measure it.

To illustrate this, consider height, another incredibly polygenic trait. Source 335 was a genome-wide association study (GWAS) about height which utilized a sample of 100,000 people, and in the regions of the genome studied, 98 loci were found which explained less than 10% of the variance in height. Should we say that this kind of result from GWA proves that the twin studies are wrong about height and that height is less than 10% heritable? No, doing so would be an obvious sanity test failure. By contrast, source 336 was able to find 700 variants associated with height using a sample of 250,000 people. It would seem that the search for molecular genetic heritability of complex, polygenic traits is just a search for larger sample sizes.

Similarly, with educational attainment as a sort of a proxy for intelligence, source 337 was able to find 3 new associated genetic variants using a sample of 125,000 people. By contrast, source 338 was able to find 74 associated variants using a sample of about 300,000 people, and ~160 variants using their combined sample of about 400,000 people. The variants source 338 found were disproportionately found in genomic regions regulating gene expression in fetal brains.

Polygenic scores computed from current GWAS are currently able to account for 12% of variance in g [1158].

It's also important to note many kinds of theoretical genetic effects that genome-wide association would not be able to measure:

- Non-additive effects (gene-gene interactions / recessive effects where gene A only affects intelligence in the presence of gene B): Identical twins share non-additive effects so twin studies can account for these effects while GWAS cannot do so.
- Rare gene-variants, copy-number variations, and other exotic kinds of genetic variants: Say that there are a bunch of rare gene-variants, so many that finding some which are unique to specific people is easy, but each individual gene variant is so rare that you are unlikely to find them in two people. GWAS can't measure these effects while twin studies can measure their net effect since identical twins would share many rare variants.

-Genome-wide Complex Trait Analysis:

Further evidence for the additive heritability of intelligence being so polygenic that GWAS is currently insufficient to capture all of even the additive genetic effects comes from another technique called Genome-wide Complex Trait Analysis (GCTA). GCTA attempts to directly measure genetic similarity among non-family members to see how random variation in genetic similarity predicts variation in trait similarity. Again, non-additive effects can't be accounted for and neither can unmeasured parts of the genome, rare variants, etc be. GCTA studies do not measure genetic similarity on the entire genome. Instead, they measure similarity on a portion of the genome and assume that unmeasured portions of the genome are "unrelated" ("unrelated" being defined as the average genetic similarity of the general population). The unmeasured parts in some will be more similar than "unrelated" and some will be less similar than "unrelated", but it's assumed that the deviations from "unrelated" will be evenly distributed around being both higher and lower than "unrelated", so the deviations will cancel each other out and make the assumption true with enough statistical power to average out a large enough from of people. Gwern has meta-analyzed GCTA studies for IQ [341], and the overall estimate about .32.

However some have suggested that assuming the unmeasured part of the genome averages out to 0 percent is an incorrect assumption and that it biases GCTA heritability downwards. Say genetic similarity is a result of parents passing down large portions of their genome to their kids all at once which means that genetic similarity on one portion of the genome will be predictive of genetic similarity on all portions of the genome. If true, assuming unrelatedness on unmeasured portions of the genome would yield a similar violation of assumptions as assortative mating, and taking the violations into account would push GCTA heritability estimates upwards. For example, source 339 finds that aggressive use of imputation for unobserved genetic information expands the GCTA heritability of height from 45% up to 56%. For intelligence specifically, source 340 expands GCTA to also look at some rarer variants which expanded heritability from 30% to 53%. More systematically, source 342 across 19 traits finds overall 42% higher heritabilities, which if we apply to Gwern's estimate, gives us a GCTA IQ heritability of 45.44%.

Source 322:

This is a good GCTA study of IQ to consider because it measures heritability using both GCTA and twin methods in the same sample, and it followed participants as they aged. It utilized participants in the Twins Early Development Study (TEDS) which included over 11,000 twin pairs born in England between 1993 and 1996. Funds were available to genotype 3665 people, 3152 of which survived quality control criteria, and of them, 2875 had g measured at least for one age, and 1344 had g measured for two ages. 700,000 SNPs were directly genotyped for these people, and with imputation, similarity for a further 1,000,000 unobserved SNPs was estimated. GCTA heritability rose from .26 at age 7 to .45 at age 12. Twin based heritability rose from .36 at age 7 to .49 at age 12. Thus, GCTA lends further support for the Wilson effect, and its estimates accounted for 74% of the twin estimate at age 7 and 94% of the twin estimate at age 12.

Source 329:

This paper genotyped a sample of 18,000 children which were broken into several samples and also did imputation for some unobserved SNPs. GCTA based heritability ranged from .22 to .46. Source 329 cites source 319 as the study giving a heritability estimate of a similar twin sample, and the twin based heritability of source 319 was .41. Therefore, we would say that the heritability of .34 is 83% of the heritability of .41.

Source 330:

This paper genotyped 6815 individuals with a median age of 57. The traditional heritability estimate was .54 while the GCTA based heritability estimate was .29. Thus, the GCTA estimate accounted for 54% of the traditional heritability estimate. The paper doesn't seem to mention imputation.

Source 331:

This paper genotyped 3511 unrelated adults and found that the GCTA based heritability of crystallized intelligence was .44 and the heritability of fluid intelligence was .51. (Crystallized intelligence refers to people's level of stored knowledge while fluid intelligence refers to their ability to perform more novel cognitive tasks). The paper suggests a heritability of full scale IQ in the high 40s, so I'll say .47. They gave no twin heritability to compare to, so I'll give them one. Based on the Wilson effect, we know that the heritability of IQ rises in adulthood up to about 80% [318]. 0.47 is about 59% of 0.80, so I'll say they detected 59% of the twin heritability.

Overall, SNP heritability accounts for 70-90% of twin based heritability, or sometimes 50% without imputation for the unmeasured genome.

Hopefully a few things have been made clear:

- Some questions about the heritability of height are not silly to analogize to the same questions about the heritability of intelligence.
- Molecular genetic methods, like all methods, are not without their flaws. It's not as if the inability of GWAS to explain much of the variance in intelligence is evidence, by itself, that the twin studies overinflate heritability. You don't need to find the specific genes to figure out the heritability of a trait within a population.
- The quantitative genetic evidence works just fine, the twin studies are mostly consistent with imputed GCTA.

The Convergence Of Methods:

In addition to twins reared together, twins reared apart, GWAS, and GCTA methods, heritability is further confirmed via censuses [491], identity by descent [534], and by virtual twin studies where unrelated children of similar age are adopted into the same family in a way that resembles normal siblings [655 & 535]. With all methods converging upon the same finding, and the tenability of the assumptions behind these methods, the evidence behind heritability can be taken as very reliable.

The Heritability Of What?

Measurement Error:

Much of the variance in IQ which is counted as "nonshared environment" is just failure in measurement reliability. When somebody takes an IQ test, and then takes the same IQ test again (controlling for learning effects, etc), the two test scores do not perfectly correlate. If, for example, you've ever taken a poorly designed test where you can tell what the correct answer is "supposed to be" but you're 100% sure that the supposed "correct" answer is incorrect, this may be low reliability on the part of the test. The reliability of an IQ test battery is not 100% [274]. When merely counting the heritability of the reliable variance in a test battery, the direct heritability of the latent g factor is .91, but only .86 before correcting for reliability [493 & 843].

g:

The heritability of various different IQ subtests vary with the heritability of g in particular being .86 [493]. Unsurprisingly, IQ subtest heritabilities are highly correlated with subtest g-loadings [355, 356, 357, 358, & 359]. After correction for measurement reliability, the heritability of g is .91 [843]. Correct for the twin misclassification EEA violation, random mating assumption violations, and violations of the assumption of genetically identical MZ twins, and the heritability of g would likely be found to be even higher [more here].

Heritability Between Who?

Heritability figures tell us the proportion of variance between individuals in a trait which is caused by genetic influences. Given this, and given that we are measuring the heritability of the correct traits, which individual differences are we measuring the heritability of? When nationally representative samples are used to assess heritability, the same heritability figures are derived [more here]. Our heritability figures also apply to both the rich and in the poor, [more here]; to Blacks, Whites, and Hispanics [more here]; to the high and low end of the ability distributions [more here], and to Western countries, to Soviet countries, to poor rural India, and even to sub-saharan African countries [more here]. DZ twins can be same sex or opposite sex, but MZ twins can only be the opposite sex; this does not affect heritability figures [more here]; Findings on twins are also generalizable to the non-twins of the population [more here].

The heritability of IQ is however non-constant across age. It rises from about .5 in childhood to about .8 in adulthood [more here].

Sign Up Bias:

A common objection is that particularly abusive or poor families don't sign up for psychological studies or don't want to, whatever the reason be. Obviously, heritability estimates are population specific, they measure how much of the phenotypic variance within a particular population is explained by genetics and if the sample is limited, the results are not necessarily generalizable. This argument is reasonable, however it has been refuted by studies which use the military or national school system to measure representative samples of either the entire population, or every male in the population. Such studies produce heritability figures which are totally consistent with the rest of the IQ literature [302 & 303]. Source 533 also examined unrelated children adopted together with the nationally representative Danish adoption register and found no correlation just like the other studies of the same experiment.

Restriction Of Range:

Similarly, some argue that adoption agencies favor middle-upper class married couples with no criminal background and with a basic understanding of parenting knowledge,

They argue that this selectivity biases heritability upwards. Even if true, this criticism only applies to studies of twins reared apart, and we have twins reared together studies which are better and cheaper to conduct. The only known study to have ever compared adoptive and non-adoptive families from the same sample found that yes, adoptive families were better, but statistically correcting for this didn't change heritability figures one iota because said variables were not seen to affect IQ in the adoptive sample [304]. Moreover, IQ gains from adoption are not g-loaded [306], and the subtests which are more heritable are the ones which are more g-loaded [355, 356, 357, 358, & 359].

Twins Versus Non-Twins:

Twins are more similar than non-twins during childhood, but this is an age effect of genetic development. As age goes up, DZ twins resemble normal siblings [318 & 532].

Wealth (Scarr Rowe):

Say that differences in wealth explain some of the variation in intelligence. Would the difference in income difference between $0 per year and $10,000 per year be as heritable as the difference between $50,000 per year and $60,000 per year? Maybe not. The difference between $0 and $10,000 is the difference between food and no food while the difference between $50,000 and $60,000 is not. To put it short, more nurturing environments would mean more people reaching their genetic potential meaning that phenotypic variance would be more of a function of genetic components, or so the story goes (this is called the Scarr-Rowe Hypothesis [165 & 166]).

On the other hand, if this either isn't true, or if an entire country has a wealth floor which is too high for this to matter, there may be no relationship.

An early study on this with a small sample size and a massive effect size was Turkheimer et al. 2003 [343]. The study is greatly over-cited, with 1546 citations on google scholar as of the time of writing this [168]. Here is figure 3 from source 343 (a=additive genetic, c=shared environment, e=unshared environment):

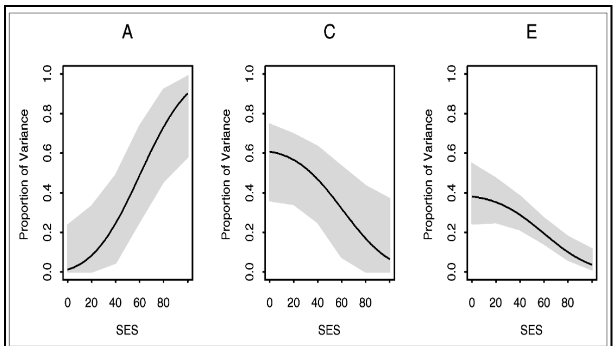


Fig. 3. Proportion of total Full-Scale IQ variance accounted for by A, C, and E plotted as a function of observed socioeconomic status (SES). Shading indicates 95% confidence intervals.

It's important to note that this paper [343] is a humongous outlier. Source 250 did a meta-analysis with regards to the Scarr-Rowe hypothesis for socioeconomic status, and Turkheimer's study [343] is the black dot furthest to the right on the funnel plots.

Source [250](#) - Figure 2:

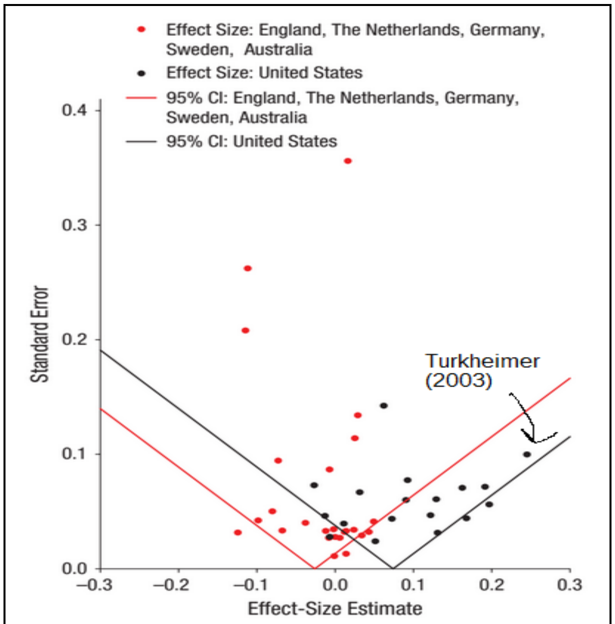
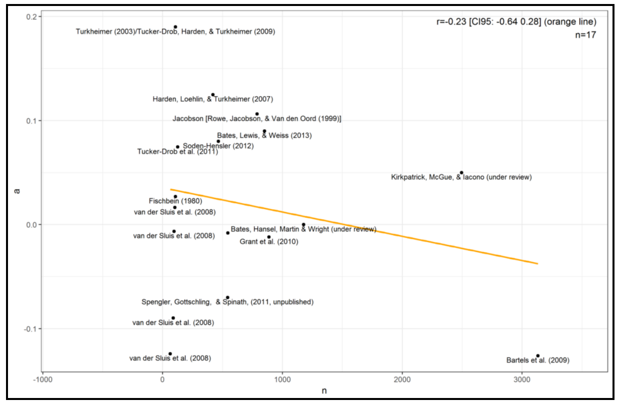
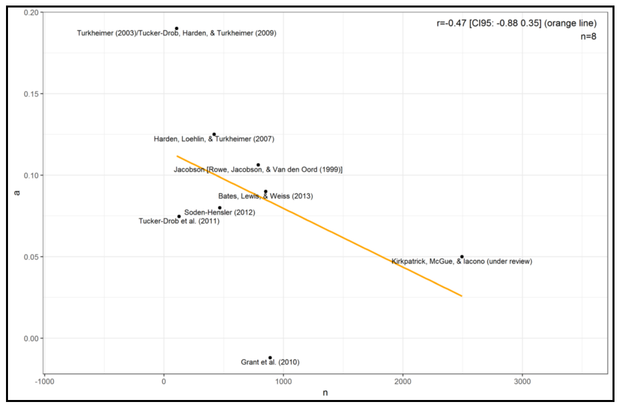


Fig. 2. Funnel plot of effect-size estimates for the Gene × Socioeconomic Status interaction in the U.S. and non-U.S. samples. Each plotted point represents the standard error and effect-size estimate for a study included in the meta-analysis. The triangle-shaped regions indicate where 95% of the data points should lie if there is no heterogeneity in population effect sizes. CI = confidence interval.

What funnel plots do is they look at the relationship between effect size and standard error. The red and black triangles mark the meta-analytic 95% confidence intervals. What the triangles do is they basically say that if the meta-analytic effect size is true, then given a study with a specific amount of statistical power, we would predict with 95% confidence that the effect size would go inside of the triangle. The studies outside of the meta-analytic 95% confidence intervals overwhelmingly push the effect size towards heritability being smaller within the poorer samples. Since the meta-analysis, further evidence has come out against Scarr-Rowe effects in Australia [915].

Isn't it clearly shown that there is a Scarr-Rowe effect, albeit a small one, which is limited to the United States? Publication bias is strongest for the USA samples:





The top scatterplot is for all samples while the bottom one is for the USA only.

Of course, the scatterplot for the USA is not conclusive because of the low amount of data points. Source 497, with 3,203 twin pairs found no Scarr-Rowe effects. Source 498 with 2,494 twin pairs found a very weak Scarr-Rowe effect with the largest difference in heritability being ~.05. These two studies (497 & 498) made up slightly more than half of source 250's full USA sample. Moreover, there are many studies either released after the meta-analysis, or missed by the meta-analysis the first time around.

After the meta-analysis was released, a large study with better methods from Florida, a good state for representativeness of the broader country, was released [167]. It found no consistent relationship between socioeconomic status and heritability. In fact, most relationships were negative. With a sample size of 34,432, it is more than 3 times the size of source 250's full USA sample.

Source 499's sample size is only slightly larger than 343, no Scarr-Rowe effect is found.

The evidence on range restriction for the adoption method is relevant, source 304 demonstrated that range-restriction of environments did not matter to heritability from adopted children, which goes against Scarr-Rowe.

Source 495 had an okay sample size (N=1,349) and it uses a decent measure of both g and SES. Scarr-Rowe effect failed to replicate, but there is one major issue with this study; it did not analyze the twin based heritability of g, but the parent-child correlation.

Using biometric models in the NLSY, source 501 found no evidence that the heritability of a variety of cognitive abilities was any lower in the bottom 20% than the normal group. The sample-size here is fairly large and the sample itself is racially diverse and oversampling of lower SES individuals. Some may not like this paper because of how it tests the Scarr-Rowe hypothesis, but this is a somewhat superior method in that it sidesteps any complaints about how SES is poorly operationalized in other studies, etc. Though this shouldn't be necessary because crude SES measures are good proxies for most shared environment effects [328 & 425]. All of the measures of intelligence that source 501 used seem to correlate with g above .7 [502].

As source 502 shows, reading comprehension is a robust correlate of g. Building on that point, a giant meta-analysis [500] found that the heritability of reading comprehension was not modified by SES, Racial composition, or nationality. This is powerful evidence against the Scarr-Rowe hypothesis.

Source 503 is perhaps the first study to use PGS to test for Scarr-Rowe in the U.S. While there was a Scarr-Rowe effect, the effect-size was meager (B=.02 on a log-scale). It also used a cohort born in the 40's, when the range of environments was likely much more variables than it currently is.

Overall, Scarr-Rowe effects in the USA seem weak at best, probably nonexistent, and inflated by publication bias.

High heritabilities of IQ have also been recorded in poorer, more primitive countries and time periods. Source 503 for example found no Scarr-Rowe effect for a U.S. cohort from the 40's. The total variance in intelligence, by itself, doesn't necessarily tell us the heritability of intelligence, but given a bunch of people prevented by the environment from reaching their genetic potential, we would theoretically expect the variance in intelligence to go down as the heritability of intelligence goes up. However, the amount of variance in intelligence has not meaningfully changed over long periods of time [846], in which enormous improvements in material quality of life and concomitant reductions in inequality of health and material well being have occurred [845]. Additionally, social class, a proxy for intelligence, has consistently been found to be 50%-80% heritable across countries [847], and, in the case of England, over time [848]. Additionally, the same heritabilities of IQ are found in Soviet Russia, East Germany, rural India [849, p. 196], and Africa [960] despite the regions' problems.

Race (Scarr-Rowe):

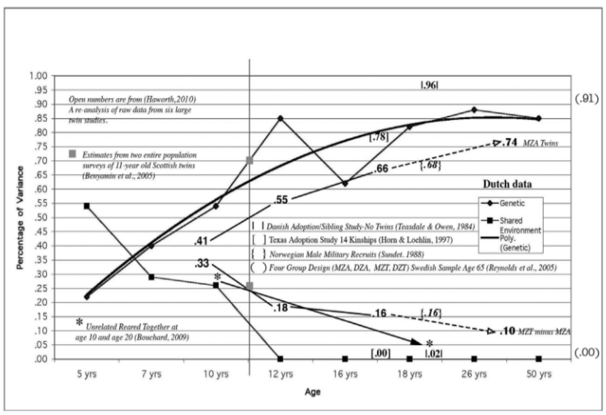
What is meant to be implied by economic Scarr-Rowe effects is that low SES is to be a proxy for the environments experienced by racial minorities. There is a meta-analysis specific to this question as well [300]. It shows that the heritability of differences between Whites and other Whites is the same as the heritability of differences between Blacks and other Blacks and is the same as the heritability of differences between Hispanics and other Hispanics. Source 167 did not report the results of tests for a Racial Scarr-Rowe effect, but source 300 reanalyzed source 167's data and found it consistent with source 300's broader meta-analysis. Source 300 meta-analyzed the Scarr Rowe hypothesis specifically with regards to whether White heritability is different from Black heritability or Hispanic heritability. All within group heritabilities were equal. Source 300 also tested for publication bias, and publication bias "wants" the heritability of differences between Whites and other Whites to be higher than the other within-group heritabilities.

Again, the fact of within group heritabilities being equal does not tell us the heritability of between group differences. However, the two kinds of heritabilities do have formal relationships [see source 344 & page 445 of source 7]. If the within group heritability is lower for the worse performing group, that would mean that the magnitude of environmental difference required for the heritability of the group differences to be zero would be a smaller magnitude than previously assumed.

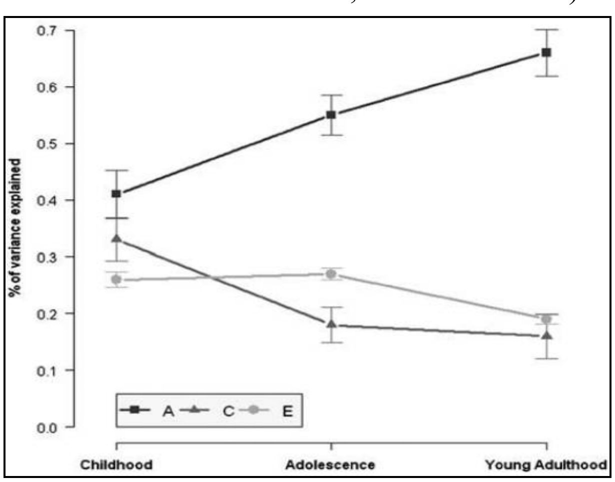
Age (The Wilson Effect):

There is a well replicated phenomenon called the Wilson effect where the heritability of IQ rises with age, usually from about .5 in childhood to .8 in adulthood. The Wilson effect has been shown in studies using a variety of methods (Twins reared together, twins reared apart, unrelated siblings adopted into the same home) over several decades utilizing data on thousands of twins and siblings [318]:

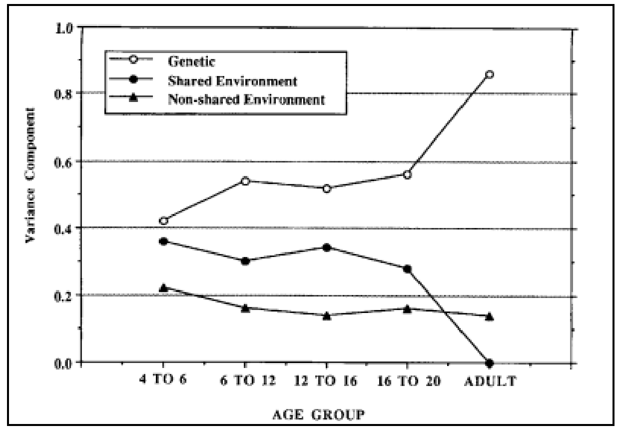
Source 318 - Figure 2 (source 308 related):



Source 319 - Figure 1: (A = additive genetic; C = Common Environment; E = Non-shared):



Source 320:



Source 321 - Figure 2:

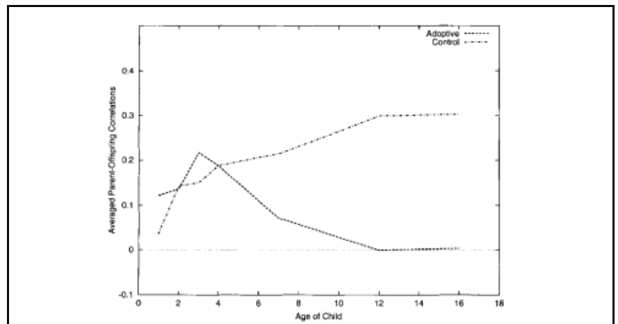


Figure 6 Correlations between parents' IQ and children's IQ in adoptive and control (i.e., biologically related) families at 1, 2, 3, 4, 7, 12, and 16 years (from Plomin et al., 1997).

A recent meta-analysis with ~150k MZ and ~150k DZ twins puts IQ heritability at .8 in the 18-64 cohort [490]. There is also molecular genetic evidence for the Wilson Effect [322]. Overall, IQ is ~50% heritable within children, and ~80% heritable within adults. Many people find this evidence to be highly counter intuitive. Surely as life goes on, and as you gain more life experience, the effect of that life experience should accrue, thereby driving twin correlations up for non-genetic reasons, thereby driving heritability downwards? Right? Why does the opposite happen?

Two things are happening. The first is that all of the longitudinal data, taken together, shows that part of the story is simply new genes activating during development. It would make sense that people are selected for how they end up as adults rather than what they were like as children. Cross-time genetic correlations are low during early childhood, they increase sharply over childhood development, and remain high from adolescence through late adulthood [332]. While the **influence** of shared environment lowers to near zero with age, shared environment factors become more **stable** with age (high cross-time shared environment correlations), just like the genetic influences. Nonshared environment correlations rise too, but they only end up at modest levels which means that they are constantly changing throughout life.

The Fadeout Effect is likely another part of the story; the effects of various environmental variables on IQ fade with time [more here]. If IQ is a function of whatever currently affects it, and genotype is the only omnipresent factor, then the fadeout of shared environment effects should be absorbed by genotype effects and by nonshared environment effects.

A third part of the story could be that —to the degree that genotype affects phenotype by affecting the environment—, heritability is driven upwards by people slowly being more and more acquainted to the environments that their genotype "wants" them to be in.

High-g Versus Low-g:

From Charles Spearman's Law Of Diminishing Returns, the Worst Performance Rule [261], and from the high correlation between g-loading and heritability [355, 356, 357, 358, & 359], we may expect that differences in g would be more heritable for between-low-g differences than for between-high-g differences. This doesn't happen. Between-high-g differences have about the same heritability as between-low-g differences [496]. Moreover, the finding that the most g-loaded tests are the most heritable is true for high-g people [496]. Also worth noting is that IQ is better at predicting job performance in the high end of the distribution than it is at predicting job performance in the low end of the distribution [64].

Predictive Validity:

List Of Outcomes:

As summarized in this useful chart from source 365, meta-analyses of hundreds of studies have demonstrated that IQ is predictive of life success across many domains.

Source 365 - Table 25.1:

Measure of Success:	r:	k:	n:	Source #
academic performance in primary education	.58	4	1,791	391
educational attainment	.56	59	84,828	253
job performance (supervisory rating)	.53	425	32,124	393
occupational attainment	.43	45	72,290	253
job performance (work sample)	.38	36	16,480	394
skill acquisition in work training	.38	17	6,713	395
degree attainment speed in graduate school	.35	5	1,700	396
group leadership success (group productivity)	.33	14	-	381
promotions at work	.28	9	21,290	397
interview success (interviewer rating of applicant)	.27	40	11,317	398
reading performance among problem children	.26	8	944	399
becoming a leader in group	.25	65	-	381
academic performance in secondary education	.24	17	12,606	391
academic performance in tertiary education	.23	26	17,588	391
income	.20	31	58,758	253
having anorexia nervosa	.20	16	484	401
research productivity in graduate school	.19	4	314	396
participation in group activities	.18	36	-	402
group leadership success (group member ratings)	.17	64	-	381
creativity	.17	447	45,880	403

Source [365] - Table 25.1 - Continued:

Measure of Success:	r	k	n	Source #
popularity among group members	.18	38	-	[402]
happiness	.05	19	2,546	[404]
procrastination	.03	14	2,151	[405]
changing jobs	.01	7	6,062	[406]
physical attractiveness	-.04	31	3,497	[407]
recidivism (repeated criminal behavior)	-.07	32	21,369	[408]
number of children	-.11	3	-	[400]
traffic accident involvement	-.12	10	1,020	[409]
persuaded by conformism	-.12	7	-	[378]
communication anxiety	-.13	8	2,548	[411]
having schizophrenia	-.26	18	-	[410]

r = correlation coefficient; k = # of studies; n = # of participants; study name replaced with source number

Measurement Quality:

One thing to keep in mind is that all of these meta-analytic correlations are probably limited by the quality of the measurements they use. For example, measuring income can be tricky since temporary events like unemployment or selling a house can cause a person's income to significantly differ from what it usually is. If income is averaged over several years, the correlation with IQ raises to .36 meaning that IQ explains 13 percent of variation in income and that a one point increase in IQ predicts a 2.5% increase in income [412].

-g:

Next, the g factor is responsible for the power of IQ tests to predict job performance [413] and academic achievement [502]. The best predictors are the most g-loaded. Therefore, studies looking at life outcomes which use more g-loaded tests and larger, more diverse test batteries should find larger effect sizes.

-Job Performance:

Source 64 reanalyzed the evidence on job performance and highlighted some interesting detail:

Table 1
Mean GCT Standard Scores, Standard Deviations, and Range of Scores of 18,782 AAF White Enlisted Men by Civilian Occupation (From Harrell & Harrell, 1945, pp. 231–232)

Occupation	*N*	*M*	*Mdn*	*SD*	Range
Accountant	172	128.1	128.1	11.7	94–157
Lawyer	94	127.6	126.8	10.9	96–157
Engineer	39	126.6	125.8	11.7	100–151
Public-relations man	42	126.0	125.5	11.4	100–149
Auditor	62	125.9	125.5	11.2	98–151
Chemist	21	124.8	124.5	13.8	102–153
Reporter	45	124.5	125.7	11.7	100–157
Chief clerk	165	124.2	124.5	11.7	88–153
Teacher	256	122.8	123.7	12.8	76–155
Draftsman	153	122.0	121.7	12.8	74–155
Stenographer	147	121.0	121.4	12.5	66–151
Pharmacist	58	120.5	124.0	15.2	76–149
Tabulating-machine operator	140	120.1	119.8	13.3	80–151
Bookkeeper	272	120.0	119.7	13.1	70–157
Manager, sales	42	119.0	120.7	11.5	90–137
Purchasing agent	98	118.7	119.2	12.9	82–153
Manager, production	34	118.1	117.0	16.0	82–153
Photographer	95	117.6	119.8	13.9	66–147
Clerk, general	496	117.5	117.9	13.0	68–155
Clerk–typist	468	116.8	117.3	12.0	80–147
Manager, miscellaneous	235	116.0	117.5	14.8	60–151
Installer–repairman, tel. & tel.	96	115.8	116.8	13.1	76–149
Cashier	111	115.8	116.8	11.9	80–145
Instrument repairman	47	115.5	115.8	11.9	82–141
Radio repairman	267	115.3	116.5	14.5	56–151
Printer, job pressman, lithographic pressman	132	115.1	116.7	14.3	60–149
Salesman	494	115.1	116.2	15.7	60–153
Artist	48	114.9	115.4	11.2	82–139
Manager, retail store	420	114.0	116.2	15.7	52–151
Laboratory assistant	128	113.4	114.0	14.6	76–147
Tool-maker	60	112.5	111.6	12.5	76–143
Inspector	358	112.3	113.1	15.7	54–147
Stock clerk	490	111.8	113.0	16.3	54–151
Receiving and shipping clerk	486	111.3	113.4	16.4	58–155
Musician	157	110.9	112.8	15.9	56–147
Machinist	456	110.1	110.8	16.1	38–153
Foreman	298	109.8	111.4	16.7	60–151
Watchmaker	56	109.8	113.0	14.7	68–147
Airplane mechanic	235	109.3	110.5	14.9	66–147
Sales clerk	492	109.2	110.4	16.3	42–149
Electrician	289	109.0	110.6	15.2	64–149
Lathe operator	172	108.5	109.4	15.5	64–147
Receiving & shipping checker	281	107.6	108.9	15.8	52–151
Sheet metal worker	498	107.5	108.1	15.3	62–153
Lineman, power and tel. & tel.	77	107.1	108.8	15.5	70–133
Assembler	498	106.3	106.6	14.6	48–145
Mechanic	421	106.3	108.3	16.0	60–155
Machine-operator	486	104.8	105.7	17.1	42–151
Auto serviceman	539	104.2	105.9	16.7	30–141
Riveter	239	104.1	105.3	15.1	50–141
Cabinetmaker	48	103.5	104.7	15.9	66–127
Upholsterer	59	103.3	105.8	14.5	68–131
Butcher	259	102.9	104.8	17.1	42–147
Plumber	128	102.7	104.8	16.0	56–139
Bartender	98	102.2	105.0	16.6	56–137
Carpenter, construction	451	102.1	104.1	19.5	42–147
Pipe-fitter	72	101.9	105.2	18.0	56–139
Welder	493	101.8	103.7	16.1	48–147
Auto mechanic	466	101.3	101.8	17.0	48–151
Molder	79	101.1	105.5	20.2	48–137
Chauffer	194	100.8	103.0	18.4	46–143
Tractor driver	354	99.5	101.6	19.1	42–147
Painter, general	440	98.3	100.1	18.7	38–147
Crane-hoist operator	99	97.9	99.1	16.6	58–147
Cook and baker	436	97.2	99.5	20.8	20–147

Table 1 (*continued*)

Occupation	*N*	*M*	*Mdn*	*SD*	Range
Weaver	56	97.0	97.3	17.7	50–135
Truck driver	817	96.2	97.8	19.7	16–149
Laborer	856	95.8	97.7	20.1	26–145
Barber	103	95.3	98.1	20.5	42–141
Lumberjack	59	94.7	96.5	19.8	46–137
Farmer	700	92.7	93.4	21.8	24–147
Farmhand	817	91.4	94.0	20.7	24–141
Miner	156	90.6	92.0	20.1	42–139
Teamster	77	87.7	89.0	19.6	46–145

Note. GCT = General Classification Test; AAF = Army Air Force; tel. & tel. = telephone and telegraph.

So, as seen above, Jobs become more complex (higher average IQ), the minimum required IQ increases, but there is no maximum IQ for any job, and the maximums in the recorded ranges are probably mostly just noise in the data.

In addition, we can see that when jobs are categorized according to their cognitive complexity, the validity of IQ is only .23 in the simplest of jobs and as high as .58 in the most complex jobs. In addition, the correlation for computer programmers specifically is .73. Third, intelligence is more related to success in job training than job performance:

Validity of the General Mental Ability (GMA) Measure in the General Aptitude Test Battery

Complexity level of job[a]	% of workforce	Performance measures: On the job	Performance measures: In training
1	14.7	.58	.59
2	2.5	.56	.65
3	62.7	.51	.57
4	17.7	.40	.54
5	2.4	.23	NR

After initial training however, the correlation between job performance and IQ raises with time as workers gain more experience up to .59 for people who have 12 or more years of experience:

Years of experience	Total sample size	GMA with performance correlation
0–3	4,424	.35
3–6	3,297	.37
6–9	570	.44
9–12	84	.44
12+	22	.59

Source 414 meta-analyzed 382 independent samples from the UK. It replicated previous findings, showing that IQ correlates at .42 with job performance, and .49 with training success. Interestingly, it also shows that IQ correlates at .32 with job performance among clerical workers and .69 with job performance among managers.

-School Year & Difficulty:

Meta-analyses which are larger than source 391 find the exact opposite pattern of what source 391 finds. Source 391 finds that the correlation between IQ and GPA decreases from primary school to secondary school to tertiary school while the larger analyses find the opposite, as shown by the table below:

Age	r	k	n	Source #
Elementary/Primary School	.45	71	18,584	245
Middle School	.54	75	49,771	245
High School	.58	71	15,427	245
High School	.65	32	13,290	415
College	.72	78	16,449	415

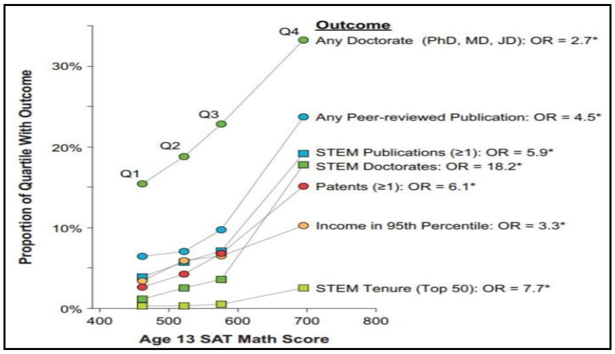
IQ also correlates much more strongly with standardized tests like the SAT, the ACT, and the GCSE than it does with grades:

Test:	Correlation with IQ:	Sample Size:	Source #
SAT-Verbal	0.80	339	246
SAT-Math	0.70	339	246
ACT	0.87	339	246
SAT	0.86	917	247
SAT	0.72	104	247
SAT	0.58	97	248
GCSE	0.81	70,000	249

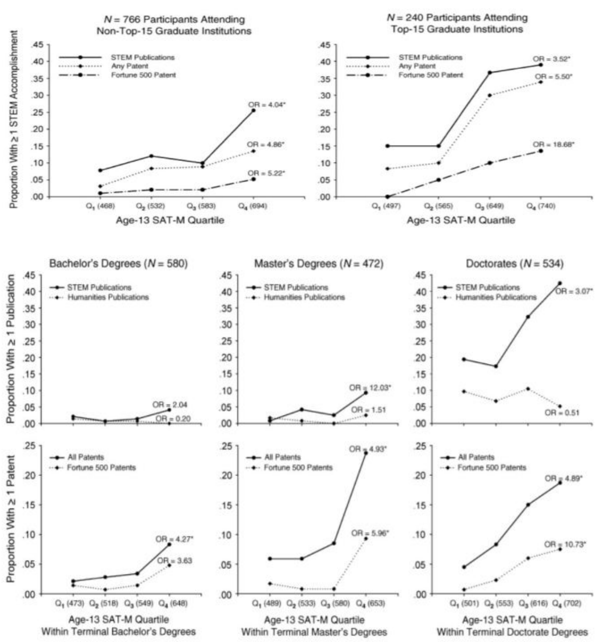
Given that the SAT is functionally an IQ subtest, we can take the following evidence as further support for the general finding that IQ can be a better predictor of life success at the high end of the spectrum than at the low end. For predicting outcomes ranging from income to educational attainment to scientific achievement, variation at the high end of the

SAT distribution corresponds to success more than variation at the low end does:

Source [251](#):



Source [252](#):



Socioeconomic Status & Causality:

Often, researchers want to know what effects IQ has on various outcomes after controlling for SES. However, this is to commit the Sociologist's Fallacy [more here] because of genetic confounding between the three variables. What's actually happening when IQ is related to life performance and the relationship is moderated by "environment", is that IQ causes life performance, and life performance causes "environment". The "environment" oftentimes is actually just caused by phenotype, and like phenotype, is substantially heritable [more here]. When the relationship between IQ and life performance is controlled for wealth, or whatever else, what the result is really saying is "When the relationship between genotype and phenotype is controlled for, genotype has no effect on phenotype!". IQ is the independent variable since it is substantially heritable [more here].

With that out of the way, the relationship between IQ and life performance is robust to controlling for SES, as shown by the table below:

Life Outcome:	Standardization Beta:	N	Source #
math scores	.60	7,147	421
reading scores	.51	7,147	421
scholastic achievement	.59	372	419
scholastic achievement	.60	100	419
scholastic achievement	.64	169	419
social class	.41	-	422
educational achievement	.47	-	423
income	.31	1,579	328
occupational status	.25	6,000	424

Along with providing the regression coefficient used in the table above, source 328 uses a siblings fixed-effects model to show that IQ predicts life outcomes within families. That is, within a given pair of siblings the sibling with the higher IQ typically ends up better educated, richer, and working a higher status occupation, than does their less intelligent sibling. This controls for all shared environment effects as well as some, but not all, genetic effects. The results of this sibling analysis are remarkably similar to regular regression results despite employing a much stronger control for the home environment:

Source 328 - Table 5-2:

TABLE 5–2
COMPARISON OF THE INDEPENDENT EFFECT OF IQ IN THE SIBLING SAMPLE USING *THE BELL CURVE*'S CONTROL FOR PARENTAL SES VERSUS A FIXED-EFFECT MODEL

	Bell Curve *Control for Parental SES*		*Siblings Fixed-Effects Model*	
Indicator	n	OLS or logit coefficient[a]	n	OLS or logit coefficient[a]
Annual earnings, year-round workers	1,579	5,548 (603)	1,579	5,317 (852)
Years of schooling	4,758	.59 (.02)	4,578	.45 (.02)
Attainment of BA	3,884	1.76 (.09)	309	1.87 (.23)
High-IQ occupation[b]	2,946	1.39 (.14)	94	1.72 (.43)
Out of labor force 1+ month[c]	1,096	−.34 (.10)	132	−.30 (.19)
Unemployed 1+ month[c]	720	−.52 (.14)	65	−.47 (.29)

Similar siblings fixed-effects results are found in 364,193 Danish men for income, grades, and welfare use [425]. This shows us that rather crude measures of SES actually do a good job of capturing most of whatever home environment variables actually matter, seeing as controlling for family by definition controls for whatever shared environmental variables actually affect.

Accordingly, straightforwardly taking a bunch of economic variables and factor analyzing them produces a so-called general socioeconomic (S) factor [797, 798, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811 & 953]; this s factor also correlates with the g factor.

Other:

Finally, turning to longitudinal research, source 253 meta-analyzed how IQ (and other predictors) correlated with income, occupational attainment, and educational attainment, with IQ measured first, and the life outcomes measured at least 3 years later making results predictive rather than retrodictive. Results are consistent with the rest of the literature, and IQ is consistently the best predictor. It is even slightly better at predicting educational attainment than grades are. IQ is also a better predictive variable in the studies with time gaps larger than 10 years:

Source 253 - Table 1:

Table 1
Predictors of socioeconomic success

	k	N	r	rw	p	S.D.r	S.D.p	CV 95%	CI 95%
Correlation with education									
Intelligence (all studies)	59	84,828	.46	.48	.56	.12	.10	.36/.75	.53/.58
Intelligence (best studies)[a]	20	26,504	.49	.48	.56	.10	.07	.42/.69	.52/.59
Father's education	72	156,360	.40	.42	.50	.14	.13	.25/.75	.47/.53
Mother's education	57	141,216	.37	.40	.48	.13	.13	.22/.73	.44/.51
Father's occupation	55	147,090	.34	.35	.42	.09	.07	.27/.56	.40/.44
Parental income	13	64,165	.29	.31	.39	.10	.11	.17/.61	.33/.46
SES index	17	69,082	.41	.44	.55	.12	.10	.35/.75	.50/.60
Academic performance	27	49,646	.48	.47	.53	.09	.07	.39/.68	.50/.56
Correlation with occupation									
Intelligence (all studies)	45	72,290	.37	.36	.43	.13	.08	.28/.57	.40/.45
Intelligence (best studies)[a]	21	43,304	.41	.38	.45	.09	.05	.35/.54	.42/.47
Father's education	52	132,591	.27	.26	.31	.08	.06	.19/.43	.29/.33
Mother's education	40	116,998	.24	.23	.27	.08	.07	.13/.41	.25/.30
Father's occupation	57	146,343	.28	.29	.35	.10	.08	.19/.51	.33/.37
Parental income	12	60,735	.19	.21	.27	.07	.10	.07/.46	.21/.32
SES index	16	74,925	.30	.31	.38	.08	.08	.22/.54	.34/.42
Academic performance	17	54,049	.33	.33	.37	.09	.07	.23/.51	.33/.41
Correlation with income									
Intelligence (all studies)	31	58,758	.21	.16	.20	.09	.11	−.01/.40	.16/.23
Intelligence (best studies)[a]	15	29,152	.22	.19	.23	.08	.06	.10/.35	.19/.26
Father's education	45	107,312	.16	.14	.17	.09	.08	.01/.32	.14/.19
Mother's education	37	93,616	.13	.11	.13	.10	.07	.00/.27	.11/.16
Father's occupation	31	98,812	.16	.15	.19	.08	.10	.00/.38	.15/.22
Parental income	17	395,562	.16	.16	.20	.06	.07	.06/.33	.16/.23
SES index	14	64,711	.15	.14	.18	.07	.08	.03/.33	.14/.22
Academic performance	14	41,937	.11	.08	.09	.07	.08	−.07/.24	.04/.13

Note. k — number of independent samples, N — number of individuals, r — average correlation, rw — sample size weighted average correlation, p — sample size weighted average correlation corrected for unreliability and dichotomization, S.D.r — standard deviation of r, S.D.p — corrected standard deviation of p, CV 95%–95% credibility intervals of p, CI 95%–95% confidence intervals of p, SES — socioeconomic status.
[a]Best studies are the ones where intelligence is tested before the age of 19, and socioeconomic success is measured after the age of 29.

In addition, a reanalysis of the evidence on job performance [426] gives us the following table:

Table 1
Predictive Validity for Overall Job Performance of General Mental Ability (GMA) Scores Combined With a Second Predictor Using (Standardized) Multiple Regression

Personnel measures	Validity (r)	Multiple R	Gain in validity from adding supplement	% increase in validity	Standardized regression weights: GMA	Standardized regression weights: Supplement
GMA tests[a]	.51					
Work sample tests[b]	.54	.63	.12	24%	.36	.41
Integrity tests[c]	.41	.65	.14	27%	.51	.41
Conscientiousness tests[d]	.31	.60	.09	18%	.51	.31
Employment interviews (structured)[e]	.51	.63	.12	24%	.39	.39
Employment interviews (unstructured)[f]	.38	.55	.04	8%	.43	.22
Job knowledge tests[g]	.48	.58	.07	14%	.36	.31
Job tryout procedure[h]	.44	.58	.07	14%	.40	.20
Peer ratings[i]	.49	.58	.07	14%	.35	.31
T & E behavioral consistency method[j]	.45	.58	.07	14%	.39	.31
Reference checks[k]	.26	.57	.06	12%	.51	.26
Job experience (years)[l]	.18	.54	.03	6%	.51	.18
Biographical data measures[m]	.35	.52	.01	2%	.45	.13
Assessment centers[n]	.37	.53	.02	4%	.43	.15
T & E point method[o]	.11	.52	.01	2%	.39	.29
Years of education[p]	.10	.52	.01	2%	.51	.10
Interests[q]	.10	.52	.01	2%	.51	.10
Graphology[r]	.02	.51	.00	0%	.51	.02
Age[s]	−.01	.51	.00	0%	.51	−.01

Much of the predictive power of other predictors of job performance is accounted for by IQ.

Miscellaneous Outcomes:

-Does IQ Measure Conformity?

With respect to leadership, source 381 meta-analysed 151 samples and found a weak positive relationship between a person's IQ and their effectiveness as, or probability of becoming, a leader. Source 380 also finds that IQ is positively correlated with the probability of someone being an entrepreneur.

With respect to risk taking behavior, which we may expect more conformist people to be less willing to engage in, greater intelligence is related to either no difference or more risk tolerance [379].

Intelligence is related to rationality and skepticism towards unfounded beliefs [286]. In 2016, Stanovich, West, and Toplac came up with a formal test of rationality in their book, source 376, which was supposed to be an attack on intelligence testing for not being the same thing as rationality. However, their own data (table 13.11) shows their Comprehensive Assessment of Rational Thinking (or CART test) to correlate with IQ at .695. So with respect to critical thinking, IQ is strongly correlated with formal tests of rationality that gauge people's propensity to incorrectly use mental heuristics or think in biased ways:

Source 376 - Table 13.11:

Table 13.11
Correlation comparisons between the full-form CART (20 subtests), the short-form CART (11 subtests), and the residual CART (9 subtests) in RT60

	Full-Form CART	Short-Form CART	Residual CART
Cognitive Ability Composite3—Turk	.695	.671	.620
Cognitive Ability Composite3—Lab	.567	.546	.474
SAT Total—Turk	.313	.319	.253
SAT Total—Lab	.495	.489	.384
Cognitive Ability Composite4—Turk	.713	.699	.638
Cognitive Ability Composite4—Lab	.614	.595	.506
Sample (Turk = 1; Lab = 2)	–.283	–.260	–.280
Sex (Male = 1; Female = 2)	–.322	–.320	–.265
Actively Open-Minded Thinking scale—Turk	.628	.631	.508
Actively Open-Minded Thinking scale—Lab	.554	.568	.387
Deliberative Thinking scale—Turk	.267	.281	.191
Deliberative Thinking scale—Lab	.472	.470	.360
Future Orientation scale—Turk	.311	.296	.286
Future Orientation scale—Lab	.297	.278	.267

For Cognitive Ability Composite3 (*N* = 747)
Correlations > .075 significant at the .05 level, two-tailed
Correlations > .126 significant at the .001 level, two-tailed
For Cognitive Ability Composite4 and SAT (*N* = 538)
Correlations > .086 significant at the .05 level, two-tailed
Correlations > .141 significant at the .001 level, two-tailed

One formal logical fallacy is the appeal to authority fallacy ("the government says it therefore it's true!"). Source 378 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.

Intelligence has been found to be related to humor ability [494].

With respect to real world problems as measured by situational judgement tests (SJTs), source 377 found a .46 correlation between people's scores on SJTs and IQ tests in a meta-analysis of the subject.

So, the short answer is no, it does not.

-Longevity:

Source 382 meta-analyzed 16 longitudinal studies totaling 1,107,022 participants and 22,453 deaths; smarter people are, in general, less likely to die of all causes. Adult SES and education somewhat mediates the relationship, but childhood SES doesn't which suggests that the reason for mediation is that adult SES is influenced by intelligence. Adding to this, there is also evidence that the relationship between intelligence and general lifespan is mostly genetically mediated [383].

For more specific associations, source 637 used data on 7,476 participants of the 1979 NLSY who had intelligence measured in the NLSY, and a variety of health outcomes measured ~20 years later at 40 years old. It also reviews some of the other literature for cognitive epidemiology at the start. Source 637's results are only slightly attenuated by parental SES. Of the 19 significant relationships, intelligence is associated with better outcomes on 15 of them including ulcers, severe tooth or gum trouble, epilepsy or fits, stomach or intestinal ulcers, lameness/paralysis/polio, sleeping trouble, headaches/dizziness/fainting, anemia, chest pain/palpitations, neuritis, leg pain / bursitis, depression/anxiety, asthma, foot and leg problems, and Kidney/Bladder problems.

Longitudinal data on a cohort of over 1,000,000 Swedish men shows fatal and non-fatal accidental injury to be related to lower intelligence [638 & 639]. Additionally, a small meta-analysis finds intelligence to be negatively related (-.12) to involvement in a car accident [409].

Given a pre-existing injury, people of higher intelligence are better at dealing with the situation. One experiment on the efficacy of a drug which also measured the IQ of participants found that the higher IQ participants persisted with taking the medication for longer periods of time indicating that they could better care for themselves [640]. Investigation of the link between health literacy and actual health also finds that the relationship is almost entirely mediated by intelligence [641]. Intelligent people also make use of more preventative medicine even when access to healthcare is equal [642].

Using longitudinal data from a nationally representative (for the U.K.) sample of 17,419, source 651 finds that high childhood IQ predicts lower BMI, less obesity, healthier food consumption, and more frequent exercise in adulthood after controlling for education, earnings, mother's BMI, father's BMI, childhood social class, and sex. However, before controls, IQ only explains 0.009% of variance. Food deserts (poor areas where healthy food is scarce or expensive) are also the result of insufficient demand for healthy foods [841].

-Self Control / Time Preference:

One concept from economics which has utility outside of economics is the concept of time preference. Imagine offering a child the option of having 1 chocolate bar now, or ten chocolate bars in one month's time. The child which prefers having 1 chocolate bar as soon as possible is the child with a higher time preference. Higher IQ people tend to have lower time preferences. In a meta-analysis looking at "delay discounting", which is defined the same as time preference, the correlation between IQ and low time preference was found to be -0.23 on the aggregate [871]. This relationship is genetically mediated [1115], however this genetic mediation cannot fully explain the heritability of self control because self control is about 50% heritable [1117, 1118, & 1119].

-Financial Decision Making:

Source 1160:

When inflation happens, the value of a dollar on any given day is less than the value of a dollar the previous day. Given this, a rational actor would respond to inflation by purchasing everything as soon as possible or buying a currency like gold which doesn't experience as much inflation. This paper found that above median IQ men to display 50% less errors in predicting when inflation would occur, and were also more likely to consume in the short term when inflation was happening.

Source 1161:

This paper found higher IQ investors to display superior market timing, stock-picking skill, and trade execution.

-Crime:

Chapter 16 of source 384 meta-analyzed research done on the relationship between IQ and crime, delinquency, and related variables. Out of 68 studies on IQ and delinquency, 60 found a negative relation (88%) and the remaining 8 found no significant relationship. Out of 19 studies on IQ and adult criminal offending, 15 (79%) found a negative correlation. Out of 17 studies on self-reported offending and IQ, 14 (82%) found a negative relationship. Out of 5 studies on IQ and antisocial personality disorder, and out of 14 studies on childhood conduct disorder, all 19 found a negative relationship. Thus, the vast majority of research establishes IQ as a correlate of crime and related constructs. On the other hand, only 7 of 19 (36%) of studies on recidivism and IQ found a negative relationship. The authors posit that this is explained by range restriction; to be able to be caught in 2 crimes you have to be dumb enough to commit the first one which means the population of interest has undergone significant range restriction. Source 408 however did a meta-analysis on recidivism going over 32 studies and 21,369 participants and found a -.07 correlation between intelligence and recidivism.

These findings are confirmed by large, representative birth cohort studies in the United States [387], Finland [385], and Sweden [386]. The massive (700,514 participants) study from Sweden [386] found that the negative -.19 correlation between IQ and crime only fell to -.18 when controlling for income and single motherhood.

With regards to the differential detection hypothesis, source 388 investigated the impact of neighborhood characteristics and found that the negative relationship with criminality held even after controlling for neighborhood poverty, unemployment, % Black, % female headed household, and % on public assistance, as well as individual age, sex, race, poverty, self-control, and age. Although, the relationship between IQ and criminality was much stronger in well-off areas than it was in disadvantaged areas. We also have evidence like source 389 which compares actual arrests to self report finding no difference in intelligence estimates between methods of assessing criminality. Perhaps self report isn't the best assessment, but the result is certainly not what you would predict if differential detection mattered. Either way, to whatever degree differential detection matters, the impact that IQ has on how your life is affected by run-ins with the law remains the same.

There is also longitudinal evidence linking IQ measured in early childhood to crime later in life. Source 390 conducted a 25-year longitudinal study on 1,625 participants. They found that IQ at age 8-9 predicted criminality in adulthood. This relationship was also found to be mediated by childhood conduct problems, which just tells us that IQ begins to have an effect on criminality at an early age.

A meta-analysis of over 27,000 people from four European twin cohorts [842] on academic performance (i.e. intelligence-proxy) and aggression (parental and self-ratings) finds both within-family associations and between-family associations, thus ending discussion of neighborhood characteristics &

shared environment. The twin data also shows genetic mediation between the two, but relationships are still found between MZ twins which implies a role of nonshared environment. The agreement of parental report and self report is also further evidence against the differential detection hypothesis.

On IQ & Human Value:

Intelligence is an incredibly unidimensional trait [more here, here, here & here], it is not very malleable [more here, here, & here], and individual differences in intelligence are mostly genetically caused [more here]. IQ is also the most important variable influencing life success across many domains [more here], however this does not mean that intelligence explains all or even a majority of the variance in success. Let's take the two life outcomes which intelligence is most predictive of: grades in high school (.58 [286]), and job performance (.58 in complex jobs [64]). In this case .58 squared is .3364, meaning that, at best, IQ explains 33.64% of variance, and in most life outcomes, it explains well below that. It also doesn't matter how smart a person is if they never put in the required effort to use their intelligence to solve tasks. Although IQ is a better predictor, conscientiousness, a personality trait from the big 5 test which measures work ethic among other things, also has validity independent of intelligence for predicting job performance [426]:

Table 1
Predictive Validity for Overall Job Performance of General Mental Ability (GMA) Scores Combined With a Second Predictor Using (Standardized) Multiple Regression

Personnel measures	Validity (r)	Multiple R	Gain in validity from adding supplement	% increase in validity	Standardized regression weights: GMA	Standardized regression weights: Supplement
GMA tests[a]	.51					
Work sample tests[b]	.54	.63	.12	24%	.36	.41
Integrity tests[c]	.41	.65	.14	27%	.51	.41
Conscientiousness tests[d]	.31	.60	.09	18%	.51	.31
Employment interviews (structured)[e]	.51	.63	.12	24%	.39	.39
Employment interviews (unstructured)[f]	.38	.55	.04	8%	.43	.22
Job knowledge tests[g]	.48	.58	.07	14%	.36	.31
Job tryout procedure[h]	.44	.58	.07	14%	.40	.20
Peer ratings[i]	.49	.58	.07	14%	.35	.31
T & E behavioral consistency method[j]	.45	.58	.07	14%	.39	.31
Reference checks[k]	.26	.57	.06	12%	.51	.26
Job experience (years)[l]	.18	.54	.03	6%	.51	.18
Biographical data measures[m]	.35	.52	.01	2%	.45	.13
Assessment centers[n]	.37	.53	.02	4%	.43	.15
T & E point method[o]	.11	.52	.01	2%	.39	.29
Years of education[p]	.10	.52	.01	2%	.51	.10
Interests[q]	.10	.52	.01	2%	.51	.10
Graphology[r]	.02	.51	.00	0%	.51	.02
Age[s]	−.01	.51	.00	0%	.51	−.01

Worth noting is that while Intelligence is substantially unidimensional [more here] and most of its predictive power is a result of its general dimension [413 & 502], g isn't the entire story and non-g residuals have some independent predictive power [1162].

One of the things which IQ is predictive of is the ability to think rationally, avoid using biased mental heuristics, and to believe correct thing [286, 376, & 378]:

Source 376 - Table 13.11:

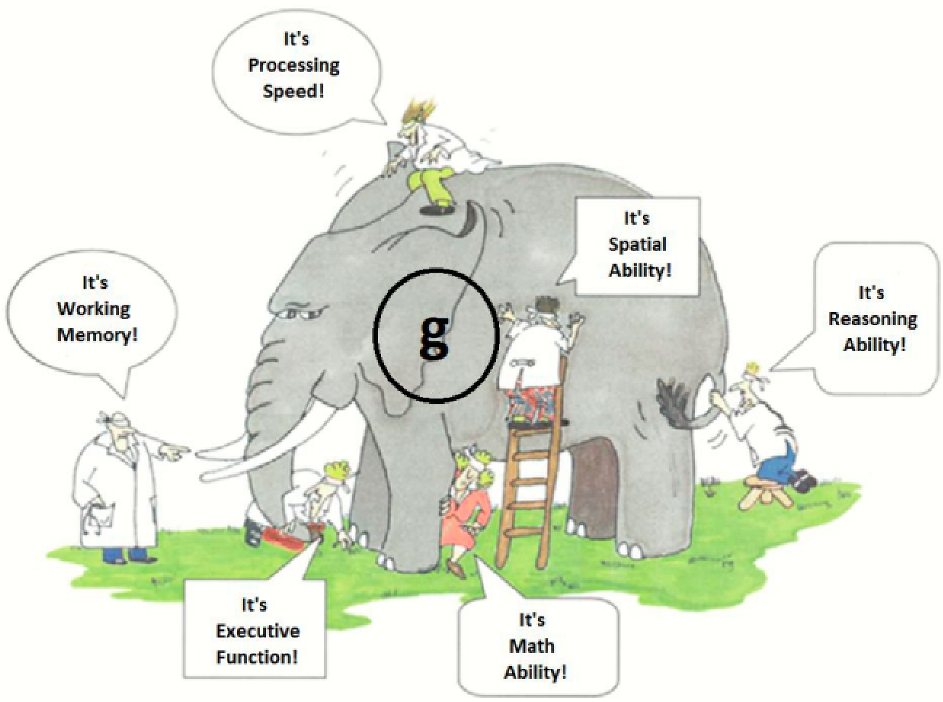
Table 13.11
Correlation comparisons between the full-form CART (20 subtests), the short-form CART (11 subtests), and the residual CART (9 subtests) in RT60

	Full-Form CART	Short-Form CART	Residual CART
Cognitive Ability Composite3—Turk	.695	.671	.620
Cognitive Ability Composite3—Lab	.567	.546	.474
SAT Total—Turk	.313	.319	.253
SAT Total—Lab	.495	.489	.384
Cognitive Ability Composite4—Turk	.713	.699	.638
Cognitive Ability Composite4—Lab	.614	.595	.506
Sample (Turk = 1; Lab = 2)	−.283	−.260	−.280
Sex (Male = 1; Female = 2)	−.322	−.320	−.265
Actively Open-Minded Thinking scale—Turk	.628	.631	.508
Actively Open-Minded Thinking scale—Lab	.554	.568	.387
Deliberative Thinking scale—Turk	.267	.281	.191
Deliberative Thinking scale—Lab	.472	.470	.360
Future Orientation scale—Turk	.311	.296	.286
Future Orientation scale—Lab	.297	.278	.267

For Cognitive Ability Composite3 (N = 747)
Correlations > .075 significant at the .05 level, two-tailed
Correlations > .126 significant at the .001 level, two-tailed
For Cognitive Ability Composite4 and SAT (N = 538)
Correlations > .086 significant at the .05 level, two-tailed
Correlations > .141 significant at the .001 level, two-tailed

This being stated, achieving rationality also requires the motivation to be rational [286]; it doesn't matter how smart you are if you don't stop to think.

Blind Men and g Elephant



4. Vanilla Privilege

Navigation:

Previous Chapter   Table Of Contents   Next Chapter

Summary:

In this chapter, we shall shamelessly play the blame game. To claim that an aspect of society is racially biased is to take upon oneself the burden of proof and to put oneself in a dangerous position; all that needs to happen for such a claim to be wrong is for enough confounding variables to be discovered that an inexplicable disparity does not exist once they are accounted for. Such a position is inherently dangerous because theoretically plausible confounders are infinite.

However of course, we are playing a game of hot potato. Once the blame has been removed from one aspect of society, that blame is simply moved onto either another aspect of society, or onto differences in behavior. From here, blame for the existence of differences in behavior, or for the existence of differences in treatment, can be passed on to yet more aspects of society or behavior.

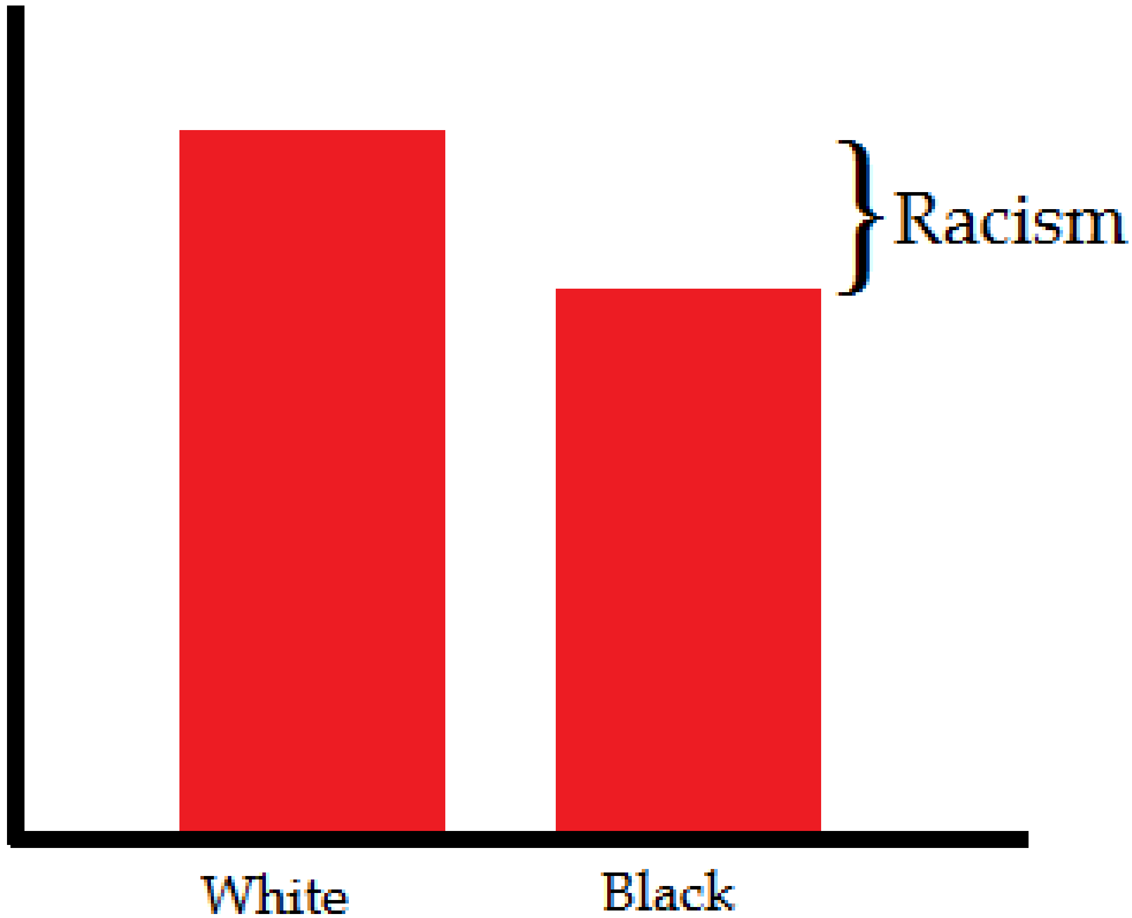
Lived Experience ≠ Evidence: In this subchapter, it is argued that it is epistemologically inappropriate to base claims of society level discrimination on anecdote, and that peoples recollections of their "lived experiences" are often epistemologically inadequate for discerning the existence of racial bias as the cause of even individual actions [more here]. It is also argued that levels of racial bias among and discrimination from Whites are low [more here], that the implicit associations test is a poor operationalization of racial bias [more here], and that whether or not people believe in stereotypes is a poor operationalization of levels of racial bias [more here]. An explanation grounded in evolutionary psychology is also offered as for why might racial biases exist [more here].

The Criminal Justice System: In this subchapter, it is argued that there is no appreciable anti-Black bias in criminal sentencing [more here], in arrests [more here], in use of force by police [more here], and in civilian stops and searches [more here]. This would mean that the Black-White crime gap really is a crime gap rather than just an arrest bias. It is also argued that the Black-White crime gap cannot be substantially explained by wealth, family structure, lead exposure, education, or child abuse [more here]; and that rather than these, it is likely mediated by differences in individual level factors such as self control, aggression, and IQ [more here].

Economic Gaps: In this subchapter, it is argued that where sufficiently studied, Blacks are afforded opportunity which is equal or superior to that afforded to Whites in various domains such as education [more here], lending [more here], and hiring [more here]. It is also argued that the modern day Black-White wealth gap cannot be explained by the historical Black-White wealth gap because the intergenerational effects of wealth usually fade to the point of negligibility within 2 generations, and that we have reason to think that this should have also applied to the Black-White wealth gap [more here]. Given the enduring presence of the various Black-White gaps and the infeasibility of modern day discrimination for explaining them, it is then argued that the modern day gaps are attributable to individual level factors such as self control and IQ [more here].

A substantial credit is due to @[Sean]__[Last].

Source Epic - Figure 13.50:



Racism deboonked.

Lived Experience ≠ Evidence:

When discussing the prevalence of discrimination, those claiming that discrimination is rampant sometimes appeal to their 'lived experience' as evidence for their view. Moreover, if statistical evidence is marshaled which suggests that discrimination is not prevalent, some people will take offense at the attempt to 'invalidate their lived experience'.

Traditionally, this kind of thinking is called "anecdotal reasoning" and people learn that it is problematic sometime in high-school or early college. Generally, it is said that anecdotal reasoning is to be avoided because human memory and judgement is highly fallible [1043], and because an individual's experience will often differ from peoples' typical experiences. For these reasons, while personal experience can be useful in the formulation of hypotheses, statistical evidence is preferred when it comes to judging the truth of such hypotheses. When better evidence isn't available, and personal experience is all we have, we should either avoid forming a view, or hold the view we form with a great deal of uncertainty.

This is all true and applicable to people's lived experience of discrimination. But there are even deeper problems here. Often, there is no evidence that discrimination took place in people's recollections of their 'lived experiences' even when those recollections are taken at face value. Frequently, these experiences merely consist of minorities being treated unfairly by particular Whites without reason to think the unfair treatment is based on race. Certain people are jerks, and in a society without racial discrimination, some Blacks would be jerks to some Blacks, some Blacks would be jerks to some Whites, some Whites would be jerks to some Whites, and yes, some Whites would even be jerks to some Blacks. Take the following two videos to more colorfully illustrate the flaws of this sort of reasoning: [1051 & 1052].

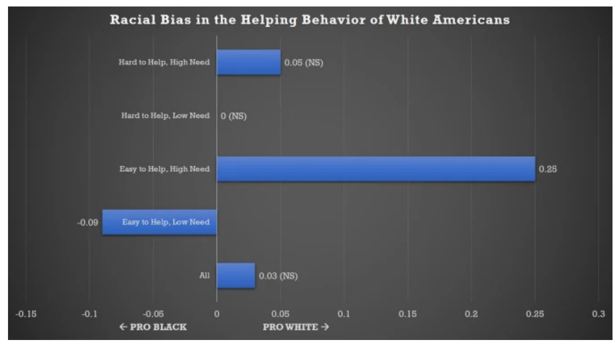
When this is pointed out, many will pivot to say that the evidence of discrimination is that Whites are disproportionately jerks to Blacks, with the general trend evidenced by the summation of lived experience, but the general trend can only properly be ascertained with empirical evidence. In fact, proper tests of discrimination generally find that Whites do not substantially discriminate [more here].

To illustrate the flaws of anecdotal reasoning as it applies to the question of discrimination in particular, take for example Kleck and Strata's experiments [1044]. In them, study participants were assigned a negative physical attribute. Some were given fake scars by make up artists while others had to fill out a biographical saying that they had epilepsy. These subjects then interacted with other people who were given said biography cards. Study participants reported that people liked them less, were patronizing, and tense, because of their assigned physical defects. What the participants didn't realize was that the people they were interacting with were not actually informed about their supposed epilepsy and a moisturizer that was applied to their scars after they viewed it in a hand mirror was actually a product that erased the whole thing. Thus, they perceived the discriminatiom they expected despite none actually taking place.

These are a few signs that this also applies to lived experiences of racial discrimination; that minorities' theories of society color their views of their various social interactions. Racial discrimination supposedly used to be overwhelming in the past before reforms to society, but younger minorities are more than or equally as likely as older minorities to say that they have experienced racial discrimination [1045, 1046, & 1047]. The same is also true for reports of discrimination by age among women [1048]. Younger women are also more likely to see men as being advantaged [1049]. Another sign is that reports of discrimination are highest among the most educated/privileged, and that reports of discrimination vary with partisan ideology, suggesting that many only believe in their discrimination when they are told that it happens [1045, 1047, 1048, 1049, & 1050]. Yet another sign of such inflated expectations is that foreign born Hispanics are less likely to report discrimination than Hispanics born in the USA [1047]. Finally, one more sign of such inflated expectations is that Blacks who live around less White people and should thus have less opportunities to experience discriminatory actions report experiencing more discriminatory behavior [1046].

How Biased Are Whites?

Experimental tests for discrimination generally find very little evidence that Whites racially discriminate against Blacks, and find much stronger evidence that Blacks discriminate against Whites. Source 478 meta-analyzed 17 such studies and found that Whites exhibited a statistically insignificant tendency to favor Blacks while Blacks exhibited a larger and statistically significant pro-Black bias. In an older meta-analysis of 31 studies totaling 48 hypothesis tests [1053], Whites showed no bias (d = .03, p = .103) for the main effect, but Blacks were not assessed. However, there were ways of cutting the data that caused differences to emerge. To produce this result, studies were separated based on how hard it was to help the stranger and how much they needed the help. When helping people was easy and no one was in dire need of help, Whites exhibited a slight bias in favor of Blacks. When helping people was easy and the people in question were in great need of help, there was a bias in favor of Whites. When helping people was hard, there was no difference in the propensity of Whites to help others:



Thinking about how such results may apply to the real world real world, we have to consider the frequency of each sort of incident. Intuitively, we may expect that the most common situations are small favors where people are easily helped in ways that slightly benefit them while situations in which help is easy and the need is high almost never happen. As for situations in which helping was difficult, statistically significant effects were not found. While source 1053 did not assess discrimination patterns by race, the previous review which source 1053 is based on did assess the behavior of Blacks, and noted that Blacks exhibited a larger in-group bias than did Whites [1054].

This is also consistent with studies which assess racial biases in experiments where people act as jurors and vote on whether or not a given defendant is guilty and on how long a convict's sentence should be.

Source 989 analyzed data from 34 such studies. It was found that Whites have nearly no bias in such decisions (0.028*d* & 0.096*d* for verdict and sentencing decisions respectively) while the Blacks exhibited a moderate in-group bias (0.428*d* & 0.731*d* for verdict & sentencing respectively).

A more recent meta-analysis [990] once again found White jurors to have no bias against Black defendants, but to have a moderate bias against Hispanics defendants. Black jurors, on the other hand, once again expressed a pro-Black/anti-White bias:

Source 990 - Table 1:

Table 1
Meta-Analytic Results for Bivariate Relationships Between Juror/Defendant Characteristics and Guilt Judgments

Juror/Defendant characteristics	N	k	$\bar{r}$	95% CI Lower	95% CI Upper	Q	I^2	BESD IV = 0 (Low)	BESD IV = 1 (High)
Defendant Attractiveness	1599	12	−.04	−.12	.04	22.37*	51	52	48
Defendant Gender	4172	25	.02	−.04	.08	78.06**	69	49	51
Defendant SES	4180	20	−.11**	−.14	−.07	20.61	8	56	45
Defendant Prior Criminal Record	2921	19	.12**	.05	.19	51.02**	65	44	56
Defendant Race	7076	51	.03	−.03	.09	267.58**	81	49	52
W Jurors with W/B or W/H Defendant	5793	39	.01	−.05	.07	224.05**	83	50	51
W Jurors with W/B Defendant	4476	32	−.02	−.09	.06	198.73**	84	51	49
W Jurors with W/H Defendant	1317	7	.11*	.03	.19	13.65*	56	45	56
B Jurors with W/B Defendant	1029	10	.13*	.01	.25	28.33**	68	44	57
Juror Need for Cognition	1747	10	−.07*	−.13	−.00	15.91	43	54	47
Juror Experience	7025	10	.03	−.01	.07	20.03*	55	49	52
Juror Education	8298	20	.00	−.03	.03	31.04*	39	50	50
Juror Authoritarianism	8205	36	.17**	.14	.20	67.72**	48	42	59
Juror Trust in Legal System									
JBS-Total Score	2763	17	.22**	.15	.28	46.24**	65	39	61
JBS-Probability of Commission	3080	12	.16**	.11	.22	24.67*	55	42	58
JBS-Reasonable Doubt	2938	11	.17**	.10	.25	43.55**	77	42	59
Juror Gender	60480	215	.08**	.06	.10	927.18**	77	46	54

In the experimental literature we can also look at studies which assess racially differential reactions when participants are assigned partners with which to complete tasks or engage in social interaction. Source 1055 meta-analyzed 108 samples from this literature and found that there was a weak, but statistically significant, tendency for each outcome to be more favorable among same race pairs of people:

Source 1055 - Table 2:

Table 2
Effect Size, Significance, and Heterogeneity Statistics for Interracial Interaction Outcome Variables

Outcome variable	k	Effect size and confidence intervals: Mean r	95% CI	Heterogeneity: $Q_w(k-1)$	I^2
Attitudes toward partner	41	.07	[.00, .14]	172.42***	76.80
Participants' emotional state	32	.10	[.05, .16]	77.45***	59.97
Nonverbal or observed behavior	37	.09	[.02, .15]	108.19***	66.73
Performance	55	.07	[.03, .11]	165.06***	67.29

Whites and minorities did not significantly differ in their degree of in group bias when this was measured in terms of their objective performance on a task or how they said they felt about their partners. However, among minorities, their reported general emotional state and body language did not differ according to the race of their partner while this was not true of Whites:

Source 1055 - Table 5:

Table 5
Average Correlations and Tests of Significance for Contextual Moderator Variables

Variable	k	Mean r	95% CI	Q_b or t(df)	Test statistic
Attitudes toward partner					
Interaction structure				$Q_b(1, 41)$	1.15
Free-form	26	.09	[.00, .18]		
Structured	15	.02	[−.07, .11]		
Location				$t(39)$	1.94ᵃ
Lab	31	.02	[−.06, .10]		
Field	10	.19	[.09, .29]		
Frequency				$t(39)$	1.03
One time	33	.05	[−.04, .13]		
Multiple	8	.15	[.03, .27]		
Salience of race				$Q_b(1, 40)$	5.27*
Race not salient	36	.08	[.01, .15]		
Race salient	4	−.10	[−.22, .04]		
Participants' emotional state					
Interaction structure				$Q_b(1, 32)$	2.84ᵃ
Free-form	21	.14	[.08, .21]		
Structured	11	.04	[−.06, .14]		
Location				$t(30)$	5.07***
Lab	24	.07	[.01, .13]		
Field	8	.21	[.10, .30]		
Frequency				$t(30)$	4.47***
One time	26	.14	[.08, .19]		
Multiple	6	−.06	[−.22, .11]		
Salience of race				$Q_b(1, 32)$	0.46
Race not salient	25	.11	[.04, .18]		
Race salient	7	.07	[−.01,.16]		
Nonverbal or observed behavior					
Interaction structure				$Q_b(1, 37)$	0.20
Free-form	21	.07	[−.04, .18]		
Structured	16	.10	[.02, .19]		
Location				$t(35)$	0.25
Lab	31	.10	[.01, .18]		
Field	6	.07	[.02, .12]		
Frequency				$t(35)$	0.39
One time	35	.09	[.02, .16]		
Multiple	2	−.09	[−.45, .28]		
Salience of race				$Q_b(1, 37)$	0.91
Race not salient	32	.07	[.00, .14]		
Race salient	5	.18	[−.03, .38]		
Performance					
Interaction structure				$Q_b(1, 55)$	5.88*
Free-form	14	.17	[.08, .26]		
Structured	41	.04	[.00, .09]		
Location				$t(53)$	0.53
Lab	30	.09	[.02, .17]		
Field	25	.06	[.01, .11]		
Frequency				$t(53)$	0.55
One time	53	.07	[.03, .11]		
Multiple	2	.09	[−.13, .30]		
Salience of race				$Q_b(1, 55)$	1.92
Race not salient	46	.05	[.01, .09]		
Race salient	8	.18	[.00, .35]		

Note. k indicates number of samples; r is the point estimate for mean effect size calculated using random-effects models; 95% CI provides the lower and upper bounds of the confidence interval for the point estimate; Q_b or $t(df)$ indicates whether a regression (Q_b) or a subgroup analysis (t) was performed and the degrees of freedom; Test statistic is the Q_b or t value, as indicated.
ᵃ $p \leq .09$.
* $p \leq .05$. *** $p \leq .001$.

Importantly, these effects have been declining with time. Studies done many decades ago found practically significant effects but research done within the last 15 years finds trivial effects on all outcomes with all measures reporting effect sizes of less than .15 by 2010:

Source [1055](#) - Table 3:

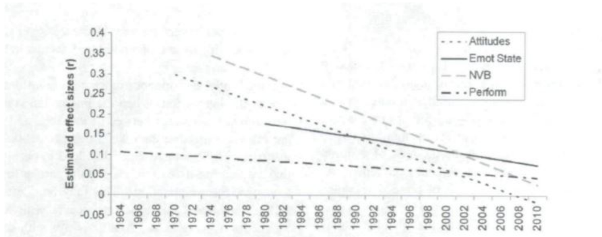


Figure 3. Historical trends in effect sizes. Higher values indicate more same-race bias; lower effect sizes indicate more egalitarian outcomes. Attitudes = attitudes toward partner; Emot State = participants' emotional state; NVB = nonverbal or observed behavior; Perform = performance.

It's also worth noting that people's explicit attitudes towards their partners, and their body language, used to exhibit the strongest effect sizes. Today, people's general emotional state and group performance are the strongest effects; really, these variables should be investigated as potential confounders of the racial effects. This is consistent with people learning to hide their discomfort with racial diversity, but it should be emphasized that even the strongest of these effects is quite weak. For all measures, around 1% or less of the variance in outcomes is explained by the racial homogeneity of the pair of people involved.

-Ethnic Identification:

While not the same as discrimination, the degree to which people say they identify with their ethnic group and consider their ethnic identity to be important is clearly related. Pew Research Center polling data finds that 74% of Blacks, 59% of Hispanics, and 56% of Asians consider their ethnicity to be an extremely/very important part of their identity while only 15% of Whites do [[1056](#)].

This is also consistent with various studies that employ more complex measures of ethnic identity. For instance:

Source [473](#) - Table 6:

TABLE 6: Ethnic Identity Item Mean Score by Ethnic Group

	Item[a] Mean Score (SD)	*Item Mean Difference Without Adjustment for SES[b] (SE)*	*Item[a] Mean Difference With Adjustment for SES[b] (SE)*
European American	2.71 (.59)	—	—
African American	3.07 (.56)	–.37*** (.03)	–.36*** (.03)
Mexican American	3.01 (.53)	–.31*** (.03)	–.32*** (.03)
Central American	3.03 (.52)	–.32*** (.04)	–.33*** (.04)
Vietnamese American	3.02 (.54)	–.32*** (.04)	–.33*** (.04)
Chinese American	3.04 (.50)	–.34*** (.05)	–.35*** (.05)
Indian American (India)	3.27 (.58)	–.56*** (.05)	–.57*** (.05)
Pakistani American	3.34 (.48)	–.64*** (.05)	–.62*** (.06)
Pacific Islander	3.11 (.55)	–.40*** (.06)	–.40*** (.06)
Mixed Ancestry	2.94 (.60)	–.23*** (.04)	–.24*** (.04)

a. European American as the comparison group.
b. SES = socioeconomic status.
****p* < .001.

Source [474](#) - Table 1:

TABLE 1: Main-Effect Differences Between Ethnic Groups on Self-Esteem, Authoritative Parenting Style, Family Stress, Teacher Support, and Ethnic Identity

	Means (SD) by Ethnic Group			
Outcome	*Hispanic[a]*	*African American[b]*	*White[c]*	F *value*
Self-esteem	3.50 (.66)†	3.93 (.68)‡	3.87 (.70)‡	19.91***
Authoritative style	1.93 (.62)†	2.12 (.60)‡	2.07 (.60)‡	3.33*
Family stress	.57 (.41)	.66 (.54)	.48 (.37)	2.48
Teacher support	2.16 (.48)†	2.24 (.49)	2.28 (.47)‡	3.09*
Ethnic identity	3.00 (.48)†	3.12 (.47)‡	2.92 (.58)†	5.49**

Source [476](#) - Table 1:

Table I. Means and Standard Deviations for Self-Esteem, Ethnic Identity, American Identity, and Other-Group Attitudes

	African Americans (*n* = 232)	Latinos (*n* = 372)	Whites (*n* = 65)
Self-esteem	3.37 (.47)	3.07 (.52)	3.12 (.57)
Males	3.41 (.47)	3.17 (.50)	3.28 (.47)
Females	3.33 (.51)	3.00 (.53)	2.93 (.63)
Ethnic identity	3.26 (.42)	3.16 (.45)	2.74 (.60)
American identity	3.23 (.88)	3.05 (.76)	3.39 (.74)
Other-group attitudes	3.07 (.66)	3.22 (.54)	3.53 (.58)

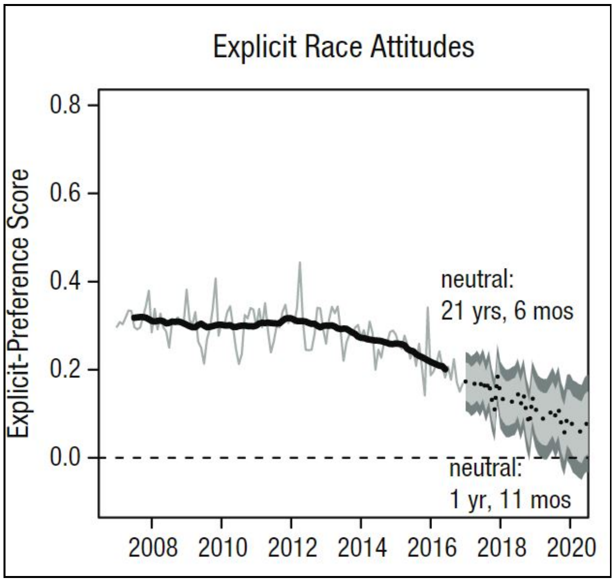
Source [477](#) - Table 4:

TABLE 4: Ethnic Identity Scores, by Ethnic Group

	High School			*College*		
	n	$\bar{X}$	SD	n	$\bar{X}$	SD
Asian	134	2.92	.49	35	3.02	.45
Black	131	3.04	.49	11	3.46	.43
Hispanic	89	2.91	.49	58	3.07	.62
White	12	2.42	.51	23	2.86	.60
Mixed	41	2.84	.51	8	2.62	.69

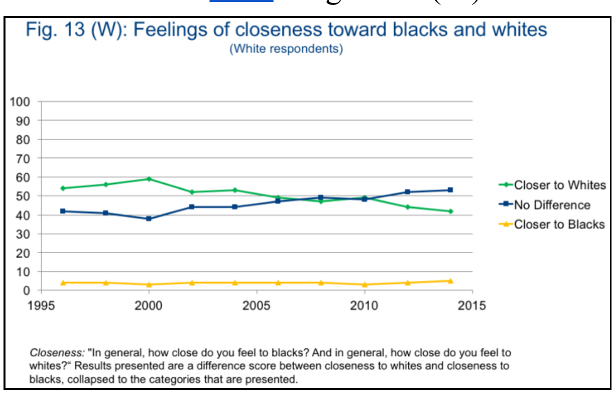
We can also look at explicit preferences where we ask people how much they like various ethnic groups and compare this to how much they say they like their own group. A meta-analysis of this sort of research [1057] finds that White Americans have a weak and declining preference for their own group equal to roughly .20 SD. The trend in this preference is such that it is expected to reach zero sometime between 2022 and 2040:

Source 1057 - Figure 1:

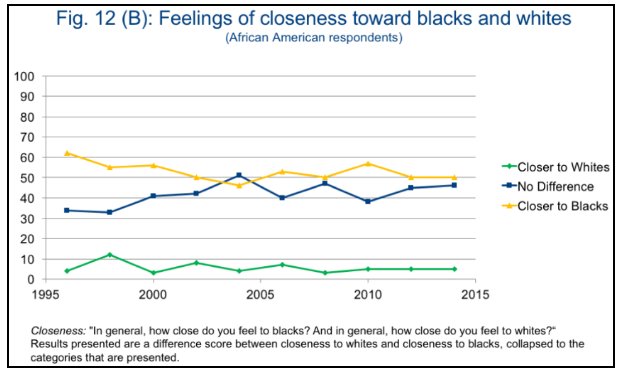


Similarly, we can look at which race Whites and Blacks say they feel the closest to. Whites generally feel about 8% less in-group closeness than do Blacks [1058]:

Source 1058 - Figure 13 (W):

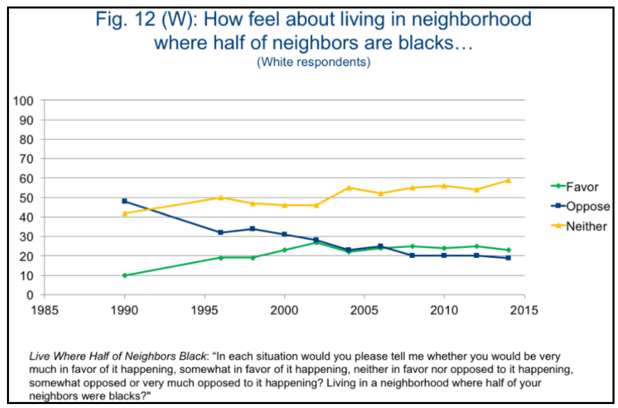


Source 1058 - Figure 12 (B):

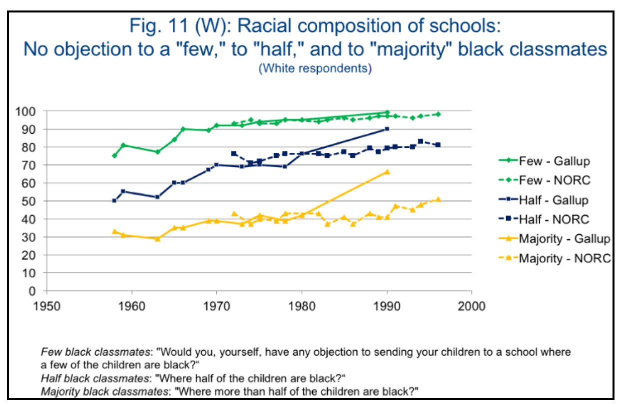


Again, the trend over time is a decrease in what would be considered the ethnocentric result. We see similar trends when we look at White opposition to things such as living in a Black neighborhood, going to a Black school, interracial marriage, etc with opposition to these things being low:

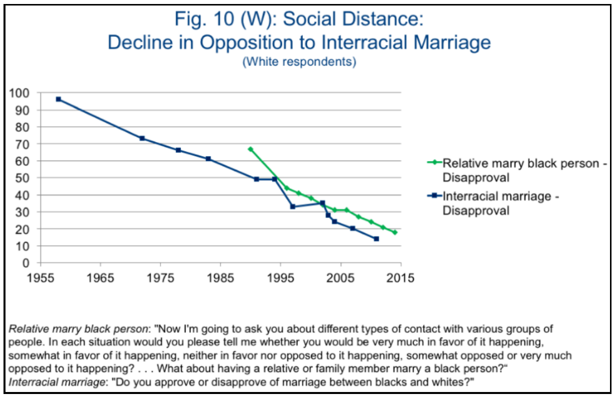
Source 1058 - Figure 12 (W):



Source 1058 - Figure 11 (W):

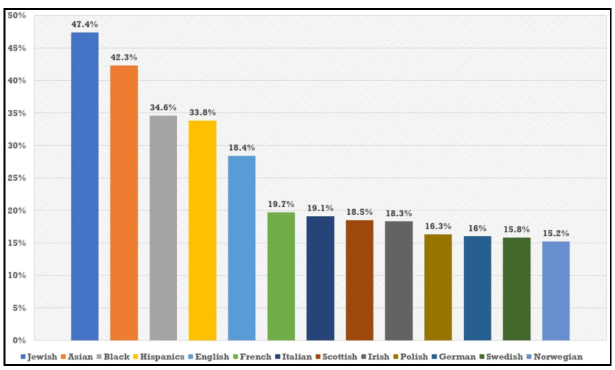


Source [1058](#) - Figure 10 (W):



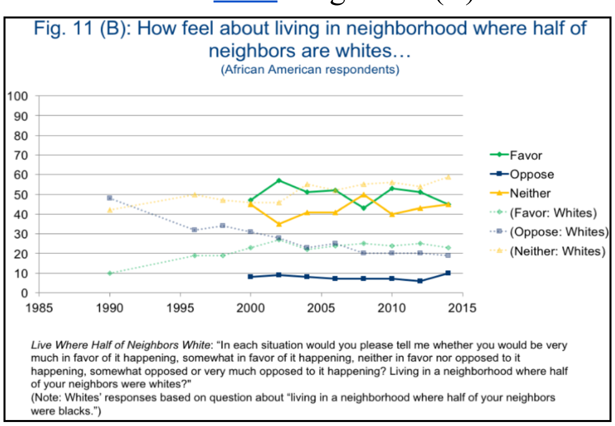
If we compare by ethnic group, Whites ethnic groups are less likely to say that marrying within the ethnic group has any importance, and this trend becomes stronger if Jews are not counted as White [[1059](#)]:

Percent who say marrying within the group is "very important" or "somewhat important" by ethnic group:

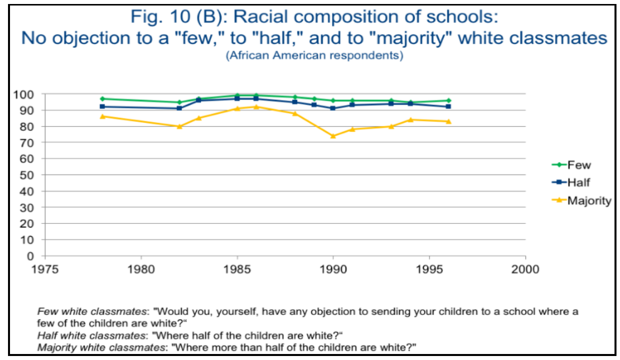


However, Blacks are more likely to want to live around Whites and go to school with Whites [[1058](#)]:

Source [1058](#) - Figure 11 (B):



Source [1058](#) - Figure 10 (B):



Perhaps perceived racial differences in school/neighborhood quality confound the results rather than people caring about the racial makeup of schools and neighborhoods in and of itself.

It is interesting that Whites act in an egalitarian way despite having a small but real in-group preference.

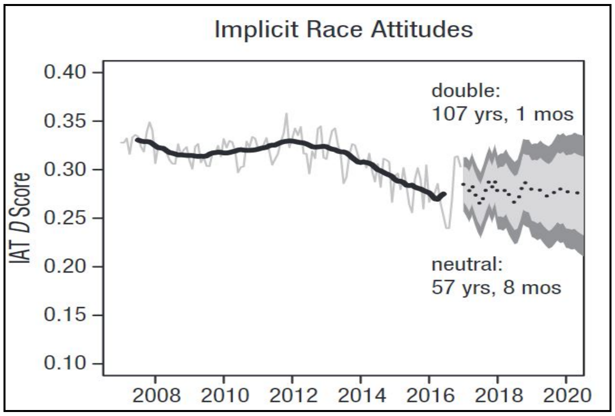
-Implicit Biases:

So far, in much of what we have looked at, participants have had control of their responses such that if they wanted to, they could manipulate the amount of racial bias which they exhibit in the experimental setting to be smaller than the amount of racial bias that they exhibit in real life. For this reason, many look to the Implicit Associations Test (IAT) as a robustness check.

In these tests, people see pairs of words or images and press a key to assign them as being “good” or “bad”. This good or bad decision is not entirely free; sometimes, when people are told to put words or images associated with Blacks into the “good” category they take something like half a second longer to press the “good” button than when Whites are paired with good items. Sometimes the opposite pattern occurs so that people take half a second longer to press the “negative” button for White faces than they do for Black faces. To the degree that this occurs, people are said to have an implicit, and possibly unconscious, bias against Blacks.

Consistent with the literature on explicit biases, implicit biases against Blacks have been declining with time [1057]. Roughly 17% of the total bias was eliminated just between the years 2007 and 2016:

Source 1057 - Figure 1:



It's noteworthy that the average degree of bias which is found (~0.3*d*), while statistically significant, is practically weak.

It is also important to mention that there is controversy concerning whether or not the IAT actually measures much of anything. Generally, researchers should use metrics that exhibit high reliability and validity. Reliability, meaning something close to consistency or precision, is often operationalized as the degree to which somebody taking the test multiple times will get roughly the same result each time. On the other hand, validity is high if our measures are measuring the things we are trying to measure, or if they correlate well with the things we think they should correlate with.

The IAT has a test-retest reliability in the range of 0.4 to 0.5 [1060 & 1061], which is lower than what is normally considered acceptable for a psychological test [1062]. Defenders of the IAT have pointed out that the test's internal reliability is higher than its test-retest reliability. So, for instance, if you arbitrarily divide the IAT test in half and score each half independently, the correlation between the two halves taken by the same person will be in the 0.6 – 0.7 range [1063]. This is better, but still questionable [1062]. The fact that the split-test reliability of the IAT is significantly greater than the test-retest reliability of the IAT implies that whatever the IAT measures changes a good deal within individuals over the course of weeks or months. These reliability estimates are low, but they are inconsistent with the view that the IAT doesn't measure anything. If that were true, then the test's reliability would be zero. But it is not.

With respect to the validity of the IAT, there is a good deal of variation depending on what we are trying to predict. The IAT does not correlate at all with experimental measures of racial bias in behavior [479 & 1064], so it has no validity in this area. So, whatever the IAT is measuring, it has nothing to do with whether people will treat Blacks differently than Whites, all else being equal. When IAT scores do predict a relevant criterion, the correlation is generally less than .20, meaning that IAT scores predict less than 4% of the variance in these outcomes [1065]:

Source 1065 - Table 1:

Table 1
Meta-Analysis of Implicit-Criterion Correlations (ICCs): Overall and by Subgroups

Criterion	k (s; N_{total})	$\hat{\rho}$ [95% CI]	$\hat{\tau}$	M	SD
All effects: Overall	298 (86; 17,470)	.14 [.10, .19]	.17	.12	.24
Interpersonal behavior	11 (6; 796)	.14 [.03, .26]	.12	.21	.15
Person perception	138 (46; 7,371)	.13 [.07, .18]	.13	.10	.21
Policy preference	21 (9; 4,677)	.13 [.07, .19]	.03	.14	.09
Microbehavior_a	96 (21; 3,879)	.07 [–.03, .18]	.19	.10	.24
Response time	6 (5; 300)	.19 [.02, .36]	.27	.31	.28
Brain activity_a	26 (8; 447)	.42 [.11, .73]	.68[b]	.26	.40
Black vs. White groups: Overall	206 (63; 9,899)	.15 [.09, .21]	.19	.13	.26
Interpersonal behavior	10 (5; 691)	.14 [.01, .28]	.14	.22	.16
Person perception	75 (30; 3,564)	.13 [.08, .19]	.12	.09	.22
$\text{Policy preference}_a$	8 (5; 1,855)	.10 [.02, .19]	.05	.09	.10
Microbehavior_b	87 (18; 3,162)	.07 [–.06, .19]	.22	.10	.25
Response time[a]	6 (5; 300)	.19 [.02, .37]	.27	.31	.28
$\text{Brain activity}_{a,b}$	20 (8; 327)	.43 [.12, .73]	.67[b]	.30	.42
Ethnic minority vs. majority groups: Overall	92 (24; 7,571)	.12 [.06, .19]	.12	.12	.18
$\text{Interpersonal behavior}_a$	1 (1; 105)	.19 [c]	—	.19	c
Person perception	63 (16; 3,807)	.11 [–.01, .23]	.15	.11	.19
Policy preference	13 (4; 2,822)	.16 [.08, .25]	.00	.17	.07
Microbehavior	9 (3; 717)	.11 [–.09, .31]	.14	.11	.19
Response time	—	—	—	—	—
Brain activity_a	6 (1; 120)	.11 [c]	—	.11	.27

Note. All effects were coded such that positive correlations are in the direction of promajority group or antiminority group responses or behaviors. With regard to Heider and Skowronski (2007) and Stanley et al. (2011), these analyses incorporate the difference score ICCs, not the Black-only and White-only ICCs. The correlation between dependent effects is assumed to be .50. The $\hat{\rho}$ for each category is based on a moderated meta-analysis across categories, where dependent effect sizes (both within and across categories) are accounted for (Hedges, Tipton, & Johnson, 2010), and the overall random-effects variance (tau-squared) weight is applied. $\hat{\tau}$ is also independently estimated within each category in separate analyses. Dashes indicate insufficient number of effects for computation purposes. Effects sharing subscripts within a category set are statistically significantly different from one another ($p < .05$). k = number of effects; s = number of independent samples within each category (this does not add up to the overall s because of sample overlap across categories); $\hat{\rho}$ = meta-analytically estimated population correlation; CI = confidence interval; $\hat{\tau}$ = random-effects standard deviation estimate; M = unweighted mean; SD = unweighted standard deviation.
[a] Even though this category in the overall analysis and in the Black-only analysis contains the same effects, results differ because estimates within categories are influenced by the effects, dependencies, and weighting across categories. [b] This extremely large value is in fact the estimated value. [c] An appropriate estimate cannot be computed due to the integrated analysis with limited effects in this category.

The major exception here is “brain activity”. The IAT is a reasonably good predictor of certain sorts of brain activity, normally amygdala response. Amygdala response is relevant because there is a separate literature linking discrimination to differences in how people’s amygdala’s respond to people based on race.

We might be tempted to interpret this as the IAT predicting the one variable that people really can’t hide, their neural responses. However, this neuro-imaging literature consists of many studies with tiny samples, as is typical of neuroscience [see more], normally less than 20 people, and most of the research has failed to find a link between amygdala response and racial bias [1066]:

Source 1066 - Table 1:

Table 1 | Reviewing amygdala activity across neuroimaging studies.

References	Contrast	Participants			Target stimuli			Task	Technical details	
		N	Age	Race	Expression	Gender	Race		Hemisphere	Resolution
(A) *UNILATERAL* AMYGDALA MODULATION										
(i) No significant effect										
Hart et al., 2000	Ingroup–Outgroup	8	20–35	B + W	Neutral	M + F	B + W	Male/Female	L	3.13 × 3.13 × 3
Phelps et al., 2000	Black–White†	14	–	W	Neutral	M	B + W	Same/Different	n/a	3.13 × 3.13 × 3
Phelps et al., 2000	Black–White (familiar)	13	–	W	Neutral	M	B + W	Same/Different	n/a	3.13 × 3.13 × 3
Cunningham et al., 2004	Black–White†	13	27	W	Neutral	–	B + W	Right/Left	n/a	3.13 × 3.13 × 6
Richeson et al., 2003	Black–White	15	21	W	Neutral	–	B + W	Right/Left	n/a	3 × 3 × 3
Richeson et al., 2003	Black–White	15	20	W	Neutral	–	B + W	Right/Left	n/a	3 × 3 × 3
Wheeler and Fiske, 2005	Black–White	7	–	W	Happy	–	B + W	Dot detection	n/a	3.75 × 3.75 × 5
Krill and Platek, 2009	Other-same race†	14	28	W	–	–	B + W	Cyberball game	n/a	–
(ii) Significant *increase*										
Lieberman et al., 2005	Black–White	20	24	B + W	Neutral	M	B + W	Same/Different	R	4 × 4 × 4
Cunningham et al., 2004	Black–White (subliminal)†	13	27	W	Neutral	–	B + W	Right/Left	R	3.13 × 3.13 × 6
Wheeler and Fiske, 2005	Black–White	7	–	W	Happy	–	B + W	Age >21?	L	3.75 × 3.75 × 5
Ronquillo et al., 2007	Black–White	11	18–36	W	Neutral	M	B + W	Age >24?	R	4.5 × 4.5 × 3.5
Ronquillo et al., 2007	Dark–Light skin	11	18–36	W	Neutral	M	B + W	Age >24?	R	4.5 × 4.5 × 3.5
(iii) Significant *decrease*										
Lieberman et al., 2005	Verbal Black–White	21	25	B + W	Neutral	M	B + W	*Verbal* Same/Diff.	R	4 × 4 × 4
Wheeler and Fiske, 2005	Black–White	7	–	W	Happy	–	B + W	Like/Dislike Veg?	R	3.75 × 3.75 × 5
(B) *BILATERAL* AMYGDALA MODULATION										
(i) Significant *increase*										
Richeson et al., 2008	Black–White (direct gaze)	9	19–23	W	Neutral	–	B + W	Right/Left	R + L	3 × 3 × 3
Demos et al., 2008	Dilated-normal pupil	27	22	–	Neutral	F	W	Passive viewing	R + L	3 × 3 × 3
Telzer et al., 2013	B-W correlation with age	32	4–17	M	Varied	–	B + W	Same/Different	R + L	3 × 3 × 3
(ii) No significant effect										
Richeson et al., 2008	Black–White (averted gaze)	9	19–23	W	Neutral	–	B + W	Right/Left	n/a	3 × 3 × 3

Section A highlights inconsistent, unilateral activity across race-related social neuroscience studies. Section B highlights results showing bilateral amygdala activity. Importantly, Target stimuli gender is rarely reported, precluding discussion of differences in activity toward Black men vs. Black women.
General: “–” for information not disclosed in articles. Contrast: “†” indicates reported positive correlation between B-W difference and Implicit Prejudice (IAT). Participants: Range given for age where mean not reported. Technical: Resolution as dimensions of one voxel in mm, compared to the ~1200 mm^3 volume of the amygdala. All studies conducted in USA.

Furthermore, we also know that the meta-analytic validity of the IAT is inflated by publication bias [479].

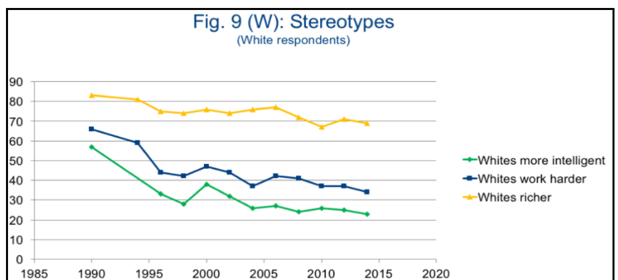
Given this, we have good reason to think that the IAT does not measure a person’s propensity to engage in racially biased behavior, and we don’t have any good reason to think that the IAT is even a good measure of racial bias that is not acted upon. There is some reason to think that it has some predictive power in this area, but that predictive power is very weak. Overall, it is not convincing evidence of significant racial bias among Whites.

-Stereotypes:

A final way that we might measure racial bias is with the degree to which Whites believe or endorse stereotypes about Blacks. One thing to consider is that some definitions of ‘stereotype’ condition whether or not a generalization is a stereotype on whether or not the generalization is accurate, and it is plausible that racial differences exist, and people form accurate stereotypes in response.

For example, Black Americans are poorer than Whites [1067]. Accordingly, Whites endorse the stereotype that Blacks are poorer than Whites [1058]:

Source 1058 - Figure 9 (W):



Stereotypes: Now, I have some questions about different groups in our society. I'm going to show you a seven-point scale on which the characteristics of people in a group can be rated. In the first statement, a score of 1 means that you think almost all of the people in that group are 'rich.' A score of 7 means that you think almost everyone in the group is 'poor.' A score of 4 means you think that the group is not towards one end or another, and of course you may choose any number in between that comes closest to where you think people in the group stand. . . . The second set of characteristics asks if people in the group tend to be hard-working or if they tend to be lazy. . . . Do people in these groups tend to be unintelligent or tend to be intelligent?“ Results presented are a difference score between evaluations of whites and blacks, collapsed to the categories that are presented.

Before interpreting the significance of the other two stereotypes, we must assess their empiricism. First, Whites complete more years of schooling than Blacks [728], and they score higher on IQ tests than Blacks [876], so whatever the causes of these differences, it is accurate to recognize the differences. Second, Blacks spend less time on homework [886], have a higher unemployment rate [1068], and spend less time working while at work [1069]. In general, literature reviews on stereotype accuracy find that stereotypes are accurate, and in the case of commonly shared stereotypes about race, are rated as highly empirically accurate more than 95% of the time [1070].

There is a research literature which attempts to assess whether or not stereotypes are harmful to the groups to which they apply which is known as the Stereotype Threat literature.

Stereotype threat occurs in a situation in which it is plausible that some members of a social group may exhibit behavior which is typical of a stereotype about their respective group. It is thought that belief in one's groups' stereotypes induces feelings of threat that cause the stereotypes to become a self-fulfilling prophecy, and that stereotype threat effects partially contribute to long standing racial and gender gaps in academic performance, intelligence, etc. It is thought that these effects can be tested with so-called "primes" in tests. For an example, let's say two groups are given a test, and for one group the start of their test says that racial groups consistently perform equally on the test, while the control group gets no such prime, or perhaps the prime says that some group performs worse. If the prime group and the control group have different performances, this is supposed to be evidence for stereotype threat.

Or at least that's the theory. The evidence? A bunch of small studies with various p-hacking issues and then some larger studies with null results. Stereotype threat effects do not exist meta-analytically [see more]. Logically, the stereotypes do not contribute to the group differences, and there is no harm in empirically evaluating the stereotypes.

-Genetic Self-Interests:

Why do people have in-group preferences? There is a well replicated phenomenon known as assortative mating. Marital Partners are psychologically [312] and genetically [316] more similar to each other than are two random members of the population. Friends are also genetically similar to each other (about as much as fourth degree cousins), and the genetic similarity of the communities that friend groups are contained within does not account for all their similarity [307]. Pretty much all psychological traits have at least some genetic component [308], and friends are most similar to each other in terms of the most heritable traits [309]. Similarity doesn't just induce contact either, it influences how much people like each other. Similarity of personality is predictive of marital satisfaction and duration [312 & 313], and the more heritable traits are better predictors [310]. There is also a positive association between kinship and fertility. Historically, in Iceland, the ideal for reproductive fitness was 3rd degree cousins [317] where the sweet spot of maximization partnership quality and minimization inbreeding was achieved. In addition, when somebody is asked to imagine a fictional person who is similar to themself in various ways, the more heritable the trait in question, the more the person will think that they would like the fictional person [311]. The friends of one twin are similar to the friends of the counterpart twin, and this trend is stronger in identical twins than in fraternal twins [309]. This lets us directly calculate the heritability of choice in friends; the heritability of choice in spouse choice is 31%, and the heritability of choice in friends choice is 21% [309]. The fact of assortative mating is robust to various controls, and assortative mating selects upon intelligence [314, 315, & 316]. Sources 483 and 484 show that MZ twins who have greater contact with each other have more similar personalities than MZ twins who are less in touch. This was thought to be a violation of the Equal Environments Assumption of the classical twin method [more here], but twin similarity causes cohabitation rather than the other way around [485].

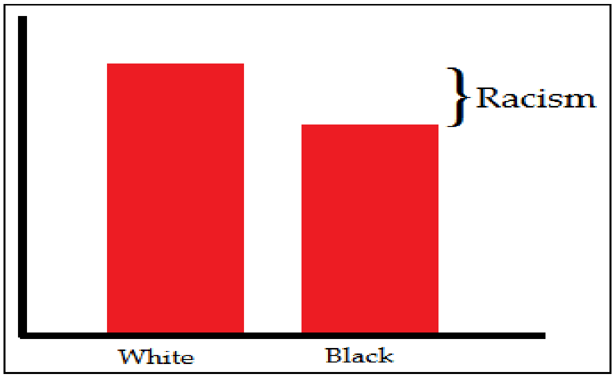
There is a sensible evolutionary logic of why people prefer similar marital partners. If people randomly mated, then the kinship coefficient between a parent and their child would be 0.5 on average. However, if the two parents are more similar than average, then the average kinship coefficient will be higher. In other words, a baby can be 60% similar to their parent instead of just 50%. For friendships and family, helping your kin will help similar genes be passed on. For greater degrees of relatedness, altruistically incurred hardships are more likely to pass a cost-benefit analysis. This is shown empirically; patterns of altruism between family members, both in humans and non-humans, showing that organisms are more willing to incur greater hardships when it benefits more genetically related family members, even controlling for the amount of contact between relatives [911].

Why all of this is relevant should be coming into picture. As we would expect from the genetics of race [see chapter 6], White + Hispanic couples are the most common interracial pairing [1071]. This makes sense because Hispanics are, on average, ~50% White [623]; this is the interracial pairing of greatest genetic similarity. The success of the relationships of similar partners extends to race as well, with monoracial marriages enduring longer than miscegenous ones [1072, 1073, 1144, 1145, 1146, & 1147]. Mixed race couples are also higher in psychological distress [1148], and are at over 2.3 times the risk of mutual assault of both monoracial White and monoracial Black couples [1074]. The evolutionary logic against mixed race relationships appears to be understood subconsciously, with women abstaining from interracial relationships more than normal during the parts of the menstrual cycle of greatest fertility [650]. Unsurprisingly, identification with one's ethnicity is associated with satisfaction and well being [473 & 1075], and diversity is associated with poorer mental health [1076].

Race is just an extended family; preference for one's own group is no more evil than love for one's own family.

The Criminal Justice System:

Source Epic - Figure 13.50:



Racism deboonked by independent fact checkers.

Despite making up only 13.6% of the population [1025], Blacks accounted for 37% of the male prison population in 2014 [1103]. This statistic grants us a useful perspective because should there be any anti-Black biases in stops and searches, arrests, or criminal sentencing, such biases would all factor into this statistic and help to explain the prison population disproportionality. We are thus left with the question: Do anti-Black biases explain the overrepresentation of Blacks among prisoners, and to the extent it doesn't, what does?

In this subchapter, it will be argued there there is no anti-Black bias in criminal sentencing decisions [more here], in arrests and police use of force [more here], and in civilian stops and searches [more here]. This would mean that the Black-White crime gap really is a crime gap rather than just an arrest bias. It will also be argued that the Black-White crime gap cannot be substantially explained by inequalities of wealth, educational attainment, family structure, lead exposure levels, or child abuse [more here]; and that rather than these, the Black-White crime gap is likely mediated by differences in individual level factors such as self control, aggression, and IQ [more here].

Stops & Searches:

One line of research is concerned with disparities in "hit rates", where a higher hit rate means that a larger proportion of people stopped and searched are found to actually have been engaging in criminal activity. If one group has a higher hit rate than another, this is said to mean that the group with a lower hit rate is held to a higher standard and is searched in response to far more minor offenses. For example, let's say that police hold Blacks to a higher standard, and that they search Blacks whenever there is evidence that there is a 40% chance of there being crime afoot, but they only search Whites if there is evidence that there is a 60% chance of a crime occuring. In this example, Whites would have a higher (60%) hit rate because of discrimination against Blacks.

Although there is also evidence against racial bias in pedestrian stops when confounds are accounted for [916 & 917], the vast majority of stops that actually happen are of cars. A review of 15 studies on the hit rate for car searches in various parts in the US finds that although there is a great deal of variation, the White hit rate is, on average, 15% higher than the Black hit rate and 47% higher than the Hispanic hit rate [918]:

Source 918 - Table 5:

Table 5
Summary of Hit Rate Findings for Racial Profiling Studies

	Hit Rates		
	Whites	Blacks	Hispanics
Wichita, KS (this study)	22.7	22.03	18.9
Maryland++	32	34	11
Florida§§	25.1	20.9	11.5
Tennessee§	20.1	19.2	10.3
New Jersey**	10.5	13.5	nr
Rhode Island+	23.5	17.8‡	17.8‡
New York (pedestrian)*	13	11	nr
Charlotte, NC¶	30.9	24.2	nr
Lansing, MI¶ ¶	6.8	8.7	nr
Missouri†	23.2	17.5	14.7
San Antonio, TX††	17.2	14.6	14.9
Denver, CO#	16.5	19.7	11.3
Denver, CO (pedestrian)#	18.7	20.6	14.6
Los Angeles, CA ##	23.8	18.2	17.2
Sacramento, CA***	26.5	22.4	28
San Diego, CA§§§	11	12	5
Washington State†††	32	21	nr

This does indeed lend plausibility to the idea that there is a small bias against Blacks and a moderate one against Hispanics. However, there are two things worth noting:

1. It does not necessarily follow that this finding is based directly on race rather than other variables that correlate with race, officers may discriminate based on a variable that happens to correlate with race, such as SES, or they may simply be assigned to Blacker areas due to higher volumes of traffic violations and general criminality. We may think this is the case since Black officers are just as likely as White ones to stop Blacks [920 & 921].
2. The 'bias' against Blacks is small, and is not most of the reason why Blacks are pulled over more often than Whites. Indeed, in addition to the hit rate disparity, the races also differ in the rate at which they commit traffic violations such as speeding and distracted driving [1006, 1007, 1008, 1009, & 1010].

Among pedestrian stops, the Black hit rate is only 6% higher, with the stop rate of Black pedestrians being 20-30% lower than the representation among crime suspect descriptions [916].

The Veil Of Darkness:

Another line of evidence concerns the so-called 'veil of darkness'. The idea is basically that day/night differences in stops are attributable to racial discrimination since officers cannot discern the races of drivers at night. Or so the story goes.

The overrepresentation of Blacks among those stopped by police does indeed remain at night, in some studies to a magnitude indicating no discrimination [919 & 995]. But additionally, proper operationalization of when officers cannot see drivers shows that Blacks are a larger percentage of those stopped during the day time in some studies [996 & 997].

However, given the hypothesis of no discrimination, one may still expect Blacks to be a larger percentage of day time stops than night time stops for two reasons. The first is that it could just be that Blacks are more likely than Whites to drive during the night than during the day; the veil of darkness method should be applied to hit rates. The second reason is that while daylight enables officers to discern race, it also enables officers to discern certain crimes. Indeed, Whites are more likely than Blacks and Hispanics to employ the use of seatbelts [998, 1001, 1002, & 1003]. In the study of 100 million stops [1011] for example, the miniscule 3.5% difference made by daylight may be explained by seat belt behavior alone given that the veil of darkness test was done in Texas, a state with a primary enforcement seat belt law. We may also expect that Blacks are more likely to keep drugs and contraband in areas which are more visible to officers because Blacks are more likely to use drugs in high-crime areas, to use and buy drugs outside, to buy drugs from strangers, and other behaviors that elevate the risk of a user being caught [1004 & 1005]. This may explain the effect [1011] of marijuana legalization on hit rate results. Additionally, Blacks are also overrepresented among crime suspects with a warrant for their arrest [916].

-More Cops = Less Crime:

A likely counterargument whenever location effects are found to be partially responsible for racial disparities may be that if the disparities are based on locational differences in the severity of policing, then that is worse than officer level discrimination because it is institutionalized.

However, it is similarly possible that what's actually selected for are municipality level variables that correlate with race. If there are racial differences in the distribution of criminal behavior (there are, [see more here]), police may target Black areas with high crime rates because of the high crime rates rather than the racial composition. Increasing police presence in an area is robustly found to decrease crime rates of targeted areas, so a larger percentage of crimes are stopped when police resources are more concentrated on areas with higher crime rates. The evidence for this is robust:

Source 922:

This analysis of data from 1990-2001 in 2074 cities finds that police added to the force by the COPS program led to statistically significant reductions in auto thefts, burglaries, robberies and assaults.

Source 923:

In this meta-analysis of "hot spot" policing, there was a small but robust and statistically significant effect size for moving police officers to high crime areas, though the meta-analytic effect was slightly inflated by publication bias.

Source 939:

Looking at federal funding for local police staffing that was associated with the 2009 stimulus bill, cities that got grants got 3.2% more police staff & saw a 3.5% lower crime rate again with a larger drop in violent crime. The finding of violent crime reducing more than property crime also replicates [949].

Source 950:

In the natural experiment of the University of Pennsylvania increasing its private police force, crime decreased in adjacent city blocks by 43–73%.

Source 955:

Conversely, utilizing data from the Dallas Police Department, it is found that following cuts to police presence, crime increased in response.

Source 1012:

Similarly, viral incidences of deadly police use of force are followed by rises in homicides because the increased scrutiny that departments undergo lead to decreased interaction with civilians. This has caused almost 900 excess homicides and almost 34,000 excess felonies.

Source 961:

In New Jersey, the two largest cities offer us a natural experiment. The Newark Police Department terminated 13% of the police force in late 2010 while Jersey City prevented any layoffs. The termination resulted in general increases in crime.

Source 418:

This paper, covering 242 large U.S. cities of above 50,000 inhabitants from 1981 to 2018, is the first to investigate racial differences in the effect of police presence on arrests and on crime. As usual, it is found that more police presence prevents crimes such as homicide. In addition, it is found that Black victimization is prevented twice as much as is White victimization. Ironically, greater presence also lowers the rate at which Blacks are arrested for serious charges, and the paper finds evidence that this is due to the deterrence of criminal activity. This makes sense because the likelihood of being caught is a much larger deterrent to criminal activity than severity of punishment [957]. The paper also finds that increased police presence leads to an increase in Blacks being arrested for low level crimes, though if this leads to less Black victimization by these lesser crimes, then it is just a value judgement as to which outcome is more important. There is also important regional variation in effects, but **on net**, increased police presence leads to better outcomes for Blacks.

Source 1001:

Turning to the effect of mandatory seatbelt laws, they increase seatbelt use by 45-80%, they reduce traffic fatalities by 8%, and they are particularly effective at protecting Blacks and minorities.

Source 924:

Predictive policing trials in Los Angeles and Kent are able to predict 1.8 times as much crime as conventional methods. Following implementation of predictive policing and the entailing changes to deployment, there was a 7.4% reduction in overall crime.

Source 925:

Similarly, one algorithm under attack for supposedly discriminating against Blacks is the Federal Post-Conviction Risk Assessment algorithm, which is used when considering what sentence lengths to assign to convicts based on recidivism, the likelihood of convicts to reoffend. Some of the variables used to assess risk include marital history, financial background, employment, educational level, criminal record, substance abuse, and criminal thinking patterns such as feelings of entitlement and rationalizing misbehavior. The algorithm is a very good predictor of recidivism, and though there are racial differences in recidivism, validity of its predictions does not differ by race which shows that the racial differences in recidivism are accounted for by variables which correlate with race:

Source 925 - Table 2:

Table 2. Predictive Utility of PCRA by Race

	Any Arrest			Violent Arrest		
Feature	**All**	**Black**	**White**	**All**	**Black**	**White**
% Arrested by PCRA Classification						
Low	11	12	10	2	2	2
Low/Moderate	29	30	27	7	8	7
Moderate	49	49	48	15	16	14
High	64	62	66	21	23	19
DIF-R, PCRA Categories	.83	.78	.85	.99	.91	1.01
AUC, PCRA Total	.73	.71	.74	.74	.72	.75

NOTE: N = 33,074.
ABBREVIATIONS: AUC = area under the ROC curve; DIF-R = dispersion index; PCRA = Post Conviction Risk Assessment.

This is important because even among Blacks, the majority of crime is committed by a small minority of the population [926]:

> "If violent careers could be stopped after 3 convictions, 53% of all violent convictions would be prevented. The recurrence rate increased from about 70% after 4 convictions to about 80% after 7 and to about 90% after 11 crimes per individual."

Arrests (13:50):

Introduction:
One approach to trying to ascertain the existence of racial bias in the criminal justice system involves comparing official data to various benchmarks. No racial bias against Blacks is shown when comparing The Uniform Crime Report to The National Crime Victimization Survey:

The Proportion of Rapes, Robberies, and Assaults, Committed by Blacks between 2000 and 2008, as estimated by the Uniform Crime Report and the National Victimization Survey

Rape		*Robbery*		*Assault*	
UCR	NCVS	UCR	NCVS	UCR	NCVS
34%	34%	56%	61%	33%	27%

NCVS [928]:
The National Crime Victimization Survey (NCVS) is a survey carried out yearly by the Department of Justice in which a random sample of ~150,000 individuals are asked about their experience with crime over the last 6 months, with a typical response rate above 80%. Participants are asked if they have been the victim of a violent crime in the last 6 months. If they have, then they are asked to answer various questions about the crime and the perpetrator of said crime. These two biennial interviews are combined on a yearly basis. The results are then weighted to eliminate bias in the sample based on demographic variables like sex and age and then used to estimate national crime rates.

UCR [928]:
The Uniform Crime Report (UCR) is an aggregation of data sent to the FBI every year by police stations all around the country (2). Not all police stations send in this data, but the UCR manages to get information for police stations which have jurisdiction over 277 million Americans (aprox. 94% of the total population). The data the FBI compiles includes information on the demographics of who is arrested every year.

Goal:
The aim of this analysis is to ascertain the proportion of violent crime committed by Blacks according to the NCVS, and to ascertain the proportion of violent crime committed by Blacks according to the UCR in order to compare the two for disparities.
Aside from homicide where there are no victims to be interviewed, the three largest categories of violent crime in both surveys from 2000 to 2008 are rape, assault, and robbery. These are thus the central focus of analysis. One unfortunate obstacle for this analysis to overcome is the fact that both the UCR and the NCVS fail to delineate Hispanics and Whites.

Analysis:
The first step is to calculate the number of rapes, assaults, and robberies, committed by Blacks and by Whites for each year. In the NCVS, tables 40 and 46 give us the total number of single offender and multiple offender crimes committed each year, and the proportion of those crimes that were committed by Blacks and by Whites. To find the total number of each criminal act committed by each race, we must **(**multiply (the total number of single offender crimes committed) by (the proportion that were committed by the race in question)**)**, and then add that to **(**(the total number of multiple offender instances of the same crime that were committed) multiplied by (the proportion of said acts that were committed by the race in question)**)**. The UCR provides us with the number of crimes committed by each race in table 43. However, we must make sure to add together "aggravated assault" and "other assault" in order to compare our numbers to the NCVS's assault categories which includes all (non sexual) forms of assault. Once we

have the number of rapes, assaults, and robberies, committed by each race we can determine how frequently each crime occurred among each race. We do this by dividing the total population size of each race during each year, taken from the census [929], by the number of crimes they committed. For instance, in 2008 there were 247221954 Whites in America, and Whites committed 2209699 assaults. This means that there was one assault committed for every 112 Whites. It should be noted that this isn't the same thing as saying that 1 in 112 Whites committed an assault because a single White person could have committed multiple assaults and therefore accounted for the 1 assault per 112 White people for several hundred people (Note: difference in total number of crimes recorded by each survey reflect the fact that the UCR doesn't cover the whole country.).

NCVS:

Year:	Source #:
2000	930
2001	931
2002	932
2003	933
2004	934
2005	935
2006	936
2007	937
2008	938

UCR:

Year:	Source #:
2000	940
2001	941
2002	942
2003	943
2004	944
2005	945
2006	946
2007	947
2008	948

Census Population Data: 929

NCVS:

Assault

Year	White population	Whte Crimes	Rate	Black population	Black crimes	Rate
2008	247112954	2209699.52	111.83	41126808	801660.64	51.3
2007	245202728	2430272.02	100.9	40598730	912089.7	44.51
2006	243168230	2867092.95	84.81	40047296	1132724.22	35.35
2005	241228151	1844451.28	130.79	39534132	824204.08	47.97
2004	239388844	2039678.1	117.37	39056228	762110.97	51.25
2003	237521836	2682631.26	88.54	38581169	908631.33	42.46
2002	235799309	2790568.93	84.5	38170579	907148.04	42.08
2001	233945047	2986313.29	78.34	37715327	1120603.41	33.66
2000	231965180	3263091.68	71.09	37224692	1151023.84	32.34

Rape

Year	White population	Whte Crimes	Rate	Black population	Black crimes	Rate
2008	247112954	104661.8	2361.06	41126808	65493.83	627.95
2007	245202728	145518.93	1685.02	40598730	40664.22	998.39
2006	243168230	115888.52	2098.29	40047296	62904.16	636.64
2005	241228151	55006.66	4385.44	39534132	86235.95	458.44
2004	239388844	100671.58	2377.92	39056228	53071.36	735.92
2003	237521836	98017.36	2423.26	38581169	47513.96	812
2002	235799309	128842.7	1830.13	38170579	84152.8	453.59
2001	233945047	139007.09	1682.97	37715327	55548.07	678.97
2000	231965180	172453.67	1345.09	37224692	46085.75	807.73

Robbery

Year	White population	Whte Crimes	Rate	Black population	Black crimes	Rate
2008	247112954	136445.46	1811.07	41126808	246326.2	166.96
2007	245202728	144514.1	1696.74	40598730	285866	142.02
2006	243168230	260109.82	934.87	40047296	263305.21	152.09
2005	241228151	132128.58	1825.71	39534132	287013.68	137.74
2004	239388844	138512.21	1728.29	39056228	226828.88	172.18
2003	237521836	156689.52	1515.88	38581169	240666.01	160.31
2002	235799309	168827.56	1396.69	38170579	222021.56	171.92
2001	233945047	195565.87	1196.25	37715327	340191.02	110.87
2000	231965180	210324.24	1102.89	37224692	337535.05	110.28

UCR:

Assault

Year	White Crimes	White Popultaion	White Rate	Black Crimes	Black Population	Black Rate
2008	853951	247112954	289.38	431823	41126808	95.24
2007	850753	245202728	288.22	426202	40598730	95.26
2006	826242	243168230	294.31	418723	40047296	95.64
2005	823521	241228151	292.92	418460	39534132	94.48
2004	809332	239388844	295.79	390641	39056228	99.98
2003	776554	237521836	305.87	381625	38581169	101.1
2002	825938	235799309	285.49	402576	38170579	94.82
2001	797316	233945047	293.42	399472	37715327	94.41
2000	765205	231965180	303.14	377230	37224692	98.68

Rape

Year	White Crimes	White Popultaion	White Rate	Black Crimes	Black Population	Black Rate
2008	10990	247112954	22485.26	5428	41126808	7576.79
2007	10984	245202728	22323.63	5708	40598730	7112.6
2006	11122	243168230	21863.71	5536	40047296	7233.98
2005	11980	241228151	20135.91	6015	39534132	6572.59
2004	12140	239388844	19719.02	5903	39056228	6616.34
2003	11766	237521836	20187.14	6114	38581169	6310.3
2002	12766	235799309	18470.88	6852	38170579	5570.72
2001	11617	233945047	20138.16	6446	37715327	5850.97
2000	11381	231965180	20381.79	6089	37224692	6113.43

Robbery

Year	White Crimes	White Popultaion	White Rate	Black Crimes	Black Population	Black Rate
2008	41962	247112954	5888.97	56948	41126808	722.18
2007	40573	245202728	6043.5	54774	40598730	741.2
2006	39419	243168230	6168.81	52541	40047296	762.21
2005	35796	241228151	6738.97	47700	39534132	828.81
2004	35439	239388844	6754.95	41774	39056228	934.94
2003	33070	237521836	7182.4	40993	38581169	941.16
2002	34109	235799309	6913.11	41837	38170579	912.36
2001	34099	233945047	6860.76	41228	37715327	914.8
2000	31921	231965180	7266.85	38897	37224692	957.01

To figure out the racial disparity between these rates, we divide the White rate by the Black rate. For instance, the NCVS shows that the White robbery rate in 2008 (1/1811 people) divided by the Black rate (1/167 people) is 11. This means that, per capita, Black people committed 11 times as many assaults as White people in 2008:

NCVS:

Assault

Year	White Rate	Black Rate	W/B Rate
2008	111.83	51.3	2.18
2007	100.9	44.51	2.27
2006	84.81	35.35	2.4
2005	130.79	47.97	2.73
2004	117.37	51.25	2.29
2003	88.54	42.46	2.09
2002	84.5	42.08	2.01
2001	78.34	33.66	2.33
2000	71.09	32.34	2.2

NCVS Continued:

Rape

Year	White Rate	Black Rate	W/B Rate
2008	2361.06	627.95	3.76
2007	1685.02	998.39	1.69
2006	2098.29	636.64	3.3
2005	4385.44	458.44	9.57
2004	2377.92	735.92	3.23
2003	2423.26	812	2.98
2002	1830.13	453.59	4.03
2001	1682.97	678.97	2.48
2000	1345.09	807.73	1.67

Robbery

Year	White Rate	Black Rate	W/B Rate
2008	1811.07	166.96	10.85
2007	1696.74	142.02	11.95
2006	934.87	152.09	6.15
2005	1825.71	137.74	13.25
2004	1728.29	172.18	10.04
2003	1515.88	160.31	9.46
2002	1396.69	171.92	8.12
2001	1196.25	110.87	10.79
2000	1102.89	110.28	10

UCR:

Assault

Year	White Rate	Black Rate	W/B Rate
2008	289.38	95.24	3.04
2007	288.22	95.26	3.03
2006	294.31	95.64	3.08
2005	292.92	94.48	3.1
2004	295.79	99.98	2.96
2003	305.87	101.1	3.03
2002	285.49	94.82	3.01
2001	293.42	94.41	3.11
2000	303.14	98.68	3.07

Rape

Year	White Rate	Black Rate	W/B Rate
2008	22485.26	7576.79	2.97
2007	22323.63	7112.6	3.14
2006	21863.71	7233.98	3.02
2005	20135.91	6572.59	3.06
2004	19719.02	6616.34	2.98
2003	20187.14	6310.3	3.2
2002	18470.88	5570.72	3.32
2001	20138.16	5850.97	3.44
2000	20381.79	6113.43	3.33

Robbery

Year	White Rate	Black Rate	W/B Rate
2008	5888.97	722.18	8.15
2007	6043.5	741.2	8.15
2006	6168.81	762.21	8.09
2005	6738.97	828.81	8.13
2004	6754.95	934.94	7.23
2003	7182.4	941.16	7.63
2002	6913.11	912.36	7.58
2001	6860.76	914.8	7.5
2000	7266.85	957.01	7.59

We can then measure how different the racial disparities reported by the NCVS and the UCR are by subtracting the NCVS disparity from the UCR disparity. A positive difference will indicate that the UCR overestimates Black crime relative to the NCVS. As can be seen in the right hand column, most of the differences are actually negative. This suggests that the UCR underestimates Black crime relative to the NCVS. In general, the two surveys match up very closely. The average differences are -0.47 for rape, .77 for assault, and -2.29 for robbery:

Assault

Year	NCVS	UCR	Difference
2008	2.18	3.04	0.86
2007	2.27	3.03	0.76
2006	2.4	3.08	0.68
2005	2.73	3.1	0.37
2004	2.29	2.96	0.67
2003	2.09	3.03	0.94
2002	2.01	3.01	1.0
2001	2.33	3.11	0.78
2000	2.2	3.07	0.87

Average = 0.77

Rape / Sexual Assault

Year	NCVS	UCR	Difference
2008	3.76	2.97	-0.79
2007	1.69	3.14	1.45
2006	3.3	3.02	-0.28
2005	9.57	3.06	-6.51
2004	3.23	2.98	-0.25
2003	2.98	3.2	0.22
2002	4.03	3.32	-0.71
2001	2.48	3.44	0.96
2000	1.67	3.33	1.66

Average = -0.47

Robbery

Year	NCVS	UCR	Difference
2008	10.85	8.15	-2.7
2007	11.95	8.15	-3.8
2006	6.15	8.09	1.94
2005	13.25	8.13	-5.12
2004	10.04	7.23	-2.81
2003	9.46	7.63	-1.83
2002	8.12	7.58	-0.54
2001	10.79	7.5	-3.29
2000	10	7.59	-2.41

Average = -2.29

Another metric to compare between the two surveys is to see if the NCVS and the UCR both report that Blacks commit roughly the same proportion of each crime. This sort of result is more quickly interpreted and understood by the Layman. In the case of the NCVS, we find the proportion of crime which is committed by Blacks by dividing the total number of crimes committed in a given year by the total number of crimes committed by Blacks, which as explained above, we get by combining proportions of single offender and multiple offender crimes on tables 40 and 46. Once again, the UCR just gives us the proportions on table 43. Such an analysis shows that the UCR tends to report that Blacks make up a somewhat higher proportion of violent criminals than the NCVS does:

Proportion of Assaults Commited by Blacks

Year	NCVS	UCR	Difference
2008	20.67%	34.20%	13.53%
2007	22.36%	33.70%	11.34%
2006	23.67%	34.50%	10.83%
2005	20.81%	56.30%	35.49%
2004	18.53%	32.70%	14.17%
2003	21.60%	33.00%	11.40%
2002	21.51%	34.20%	12.69%
2001	24.99%	33.30%	8.31%
2000	23.63%	34.00%	10.37%

Average difference: 14.2%

Proportion of Rape/Sexual Assaults Commited by Blacks

Year	NCVS	UCR	Difference
2008	32.66%	32.20%	0.46%
2007	16.37%	33.50%	17.13%
2006	24.60%	32.50%	7.90%
2005	46.63%	28.50%	-18.13%
2004	25.60%	31.90%	6.30%
2003	24.83%	33.30%	8.47%
2002	33.96%	34.00%	0.04%
2001	23.05%	34.30%	11.25%
2000	17.94%	34.10%	16.16%

Average difference: 5.5%

Proportion of Robbery Commited by Blacks

Year	NCVS	UCR	Difference
2008	48.86%	56.70%	7.84%
2007	50.48%	56.70%	6.22%
2006	40.76%	56.30%	15.54%
2005	50.40%	34.30%	-16.10%
2004	49.33%	27.00%	-22.33%
2003	43.53%	54.40%	10.87%
2002	48.42%	54.10%	5.68%
2001	57.62%	52.50%	-5.12%
2000	49.03%	53.90%	4.87%

Average: 0.83%

However, a closer look at the NCVS numbers reveals that oftentimes, the race of the offender is written down as "mixed" or "unknown". I think that many of these mixed and unknown offenders are Black, and that, as a result, the NCVS underestimates the proportion of violent crime committed by Blacks. We can get around this (and test this hypothesis) by simply subtracting all of the crimes committed by people who are neither White nor Black from both the UCR and the NCVS and then seeing if Blacks make up a similar proportion of the remaining criminals in each survey. As can be seen, they do:

Assault

Year	NCVS	UCR	Difference
2008	0.27	0.34	0.07
2007	0.27	0.33	0.06
2006	0.28	0.34	0.06
2005	0.31	0.34	0.03
2004	0.27	0.33	0.06
2003	0.25	0.33	0.08
2002	0.25	0.33	0.08
2001	0.27	0.33	0.06
2000	0.26	0.33	0.07
Average	0.27	0.33	0.06

Rape

Year	NCVS	UCR	Difference
2008	0.38	0.33	-0.05
2007	0.22	0.34	0.12
2006	0.35	0.33	-0.02
2005	0.61	0.33	-0.28
2004	0.35	0.33	-0.02
2003	0.33	0.34	0.01
2002	0.4	0.35	0.05
2001	0.29	0.36	0.07
2000	0.21	0.35	0.14
Average	0.35	0.34	0

Robbery

Year	NCVS	UCR	Difference
2008	0.64	0.58	-0.06
2007	0.66	0.57	-0.09
2006	0.5	0.57	0.07
2005	0.68	0.57	-0.11
2004	0.62	0.54	-0.08
2003	0.61	0.55	-0.06
2002	0.57	0.55	-0.02
2001	0.63	0.55	-0.08
2000	0.62	0.55	-0.07
Average	0.61	0.56	-0.05

This remains true if we aggregate the crime data for 2000-2008 and produce smaller charts that make the degree to which these surveys agree more obvious:

Assault 2000-2008

Total Crime (NVCS)	Total Black Crimes (NCVS)	Proportion of Crimes Commited by Blacks (NCVS)	Total Crimes (UCR)	Total Black Crimes (UCR)	Proportion of Crimes Commited by Blacks (UCR)	Difference (UCR-NCVS)
31633995.26	8520196.23	0.27	10975564	3646752	0.33	0.06

Rape 2000-2008

Total Crime (NVCS)	Total Black Crimes (NCVS)	Proportion of Crimes Commited by Blacks (NCVS)	Total Crimes (UCR)	Total Black Crimes (UCR)	Proportion of Crimes Commited by Blacks (UCR)	Difference (UCR-NCVS)
1601738.41	541670.1	0.34	158837	54091	0.34	0

Continued:

Robbery 2000-2008

Total Crime (NVCS)	Total Black Crimes (NCVS)	Proportion of Crimes Commited by Blacks (NCVS)	Total Crimes (UCR)	Total Black Crimes (UCR)	Proportion of Crimes Commited by Blacks (UCR)	Difference (UCR-NCVS)
3992870.97	2449753.61	0.61	743080	416692	0.56	-0.05

Conclusions:

Summarizing the main results further, we get the following table:

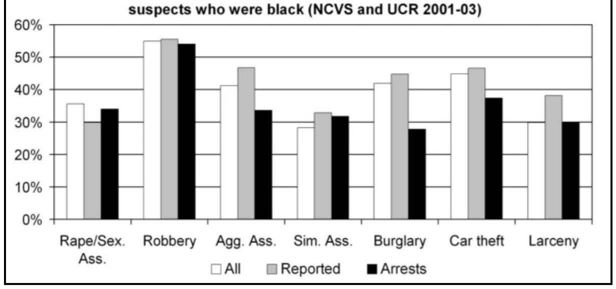
The Proportion of Rapes, Robberies, and Assaults, Committed by Blacks between 2000 and 2008, as estimated by the Uniform Crime Report and the National Victimization Survey					
Rape		*Robbery*		*Assault*	
UCR	NCVS	UCR	NCVS	UCR	NCVS
34%	34%	56%	61%	33%	27%

In conclusion, both the NCVS and the UCR report very similar racial differences in arrests for violent crime. Because of this, it is highly unlikely that UCR numbers can be explained by police bias in arrests. Instead, the most likely explanation for the UCR numbers is that Blacks really do commit far more crime than Whites. Why they do so is a separate conversation. Since police are demonstrably not biased when arresting people for most violent crimes, it is reasonable to infer that this generalizes to other crimes until evidence to the contrary is provided.

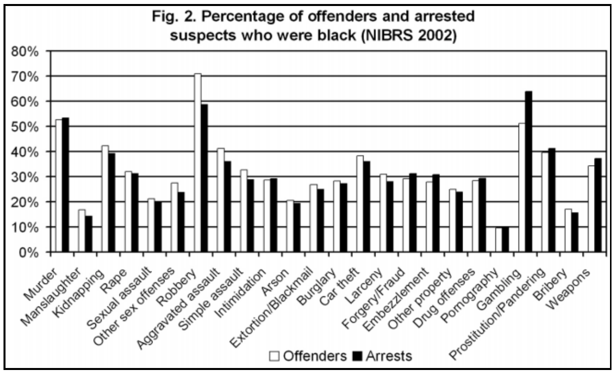
Independent analyses comparing arrest data to victimization data also produces the same general findings for more categories of offenses [1021 & 1022]:

Source 1021: Figure 1:

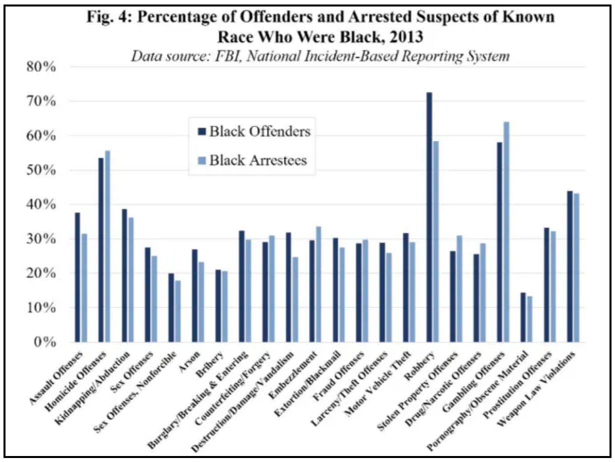
Fig. 1. Percentage of offenders who were black for all crimes and crimes reported to police, and percentage of arrested suspects who were black (NCVS and UCR 2001-03)



Source [1021](#): Figure 2:



Source [1022](#): Figure 4:



Thus, this appears to be a very robust finding. Further evidence against discrimination is the finding that Blacks are more likely to be arrested when the decision is made by a Black police officer [1023].

-Drug Arrests:

The only study I know of which attempts to assess the degree to which racial disparities in drug arrests are due to race-neutral variables is source 1005. It finds that although Blacks are 13% of the population, they make up 36% of those arrested for drug possession. According to Langan's data, Blacks are expected to be 23% of those arrested for drug possession when accounting for the types of drugs used, self report data for frequency of use, and whether or not residents live in metropolitan areas. However, these are not all of the relevant variables; Blacks are more likely to engage in risky drug purchasing behaviors such as buying from strangers, away from home, and in the outdoors [1004].

Also worth pointing out is that most evidence is based on misleading self-report data which is inappropriate because there is a myriad of evidence that Blacks under report drug usage in comparison to Whites. While self report data finds that the same percentage of Blacks 'use' drugs as do Whites, actual drug tests which run forensic analyses on people's hair, blood, urine, etc find that more Blacks use drugs [1013, 1014, 1015, 1016, 1017, & 1018]. Another sign that this happens is that sober Blacks are twice as likely as sober Whites to say that if they used drugs, they would not report it [1020].

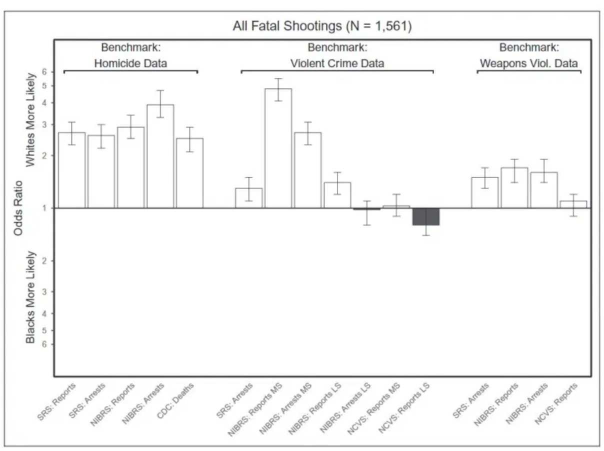
One tell that race does not affect drug arrests is the racial makeup of drug-related emergency room visits [1019]. Given these numbers in conjunction with the demographics of the United States [1024 & 1025], Blacks are 2.8 times more likely than Whites to end up in the ER because of marijuana. For cocaine, the odds ratio was 7, and for all drugs, the odds ratio was 3.5. Throwing drug arrests [1026] into the mix and directly comparing all three, in 2011 Blacks were 13.6% of the population, 30.7% of those in the ER due to drug use, and 31.7% of those arrested for drug abuse violations. Now, account for Blacks purchasing larger quantities away from home, outside, from a stranger, etc, and if anything, Whites are probably the ones who are 'discriminated' against. Another tell again relevant is that drug arrests are consistent with victimization reports in the same way as are other crimes [1021 & 1022].

-Shootings:

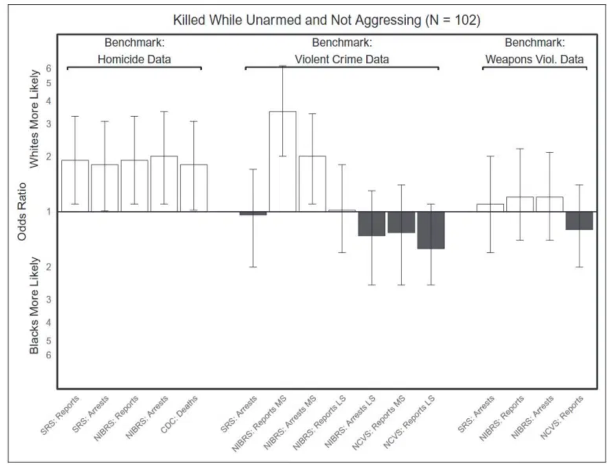
So there don't seem to be an anti-Black biases in searches [see more] or in arrests [see more], but given an arrest, are Blacks treated more harshly? Despite Blacks being 13.6% of the population [1025], they made up 31.8% of arrest related deaths from 2003-2009 [1028]. However, 13.6% is not the proper benchmark of comparison. As [previously evidenced], it is also true that despite being 13.6% of the population, Blacks account for roughly 30% of arrests for most crimes and for roughly 30% of most offenders for most crimes. So given the status of being 30% of arrestees, you would also expect them to be 30% of arrestees killed by police. Probably a better benchmark of which races offer officers more violent conflict when confronted, from 2001 to 2010, Blacks made up 44% of cop killers [1027].

Source 1029 also distinguishes between everyone killed by police and those who were killed by police while unarmed and not aggressing:

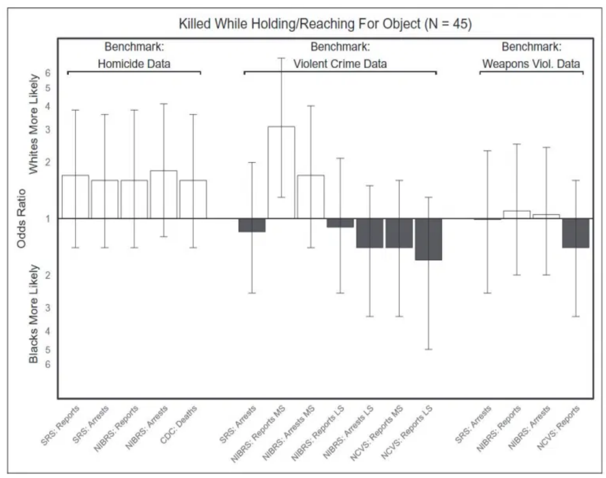
Source 1029 - Figure 1:



Source 1029 - Figure 2:



Source 1029 - Figure 3:



For the majority of estimates, Whites were overrepresented among such killings.

However, these sorts of analyses use national level FBI data, and the FBI is not reported to by 100% of police departments. So, some may have concerns that the data is incomplete, affecting results. To overcome this issue, we may simply look at more localized contexts where we know both local proportions of arrest related deaths and local benchmarks.

Multiple such local analyses have been done using local arrest rate benchmarks, consistently finding no anti-Black bias [1030, 1031, & 1032]. Of course, the best benchmark that we can use is the rate at which populations shoot at police officers. Such an analysis has been carried out and has found that using such a benchmark rendered the probability of a Black being shot 40% lower than the probability of a White being shot [1034]:

Source 1034 - Table 2:

Study #1 – Odds of black and hispanic citizens being fatally shot relative to white citizens using LEOKA benchmarks (2015–2017).

2015–2017	Black citizens	White citizens	Odds ratios	Confidence interval[a]
Fatally Shot by Police	715	1421	–	–
Benchmark				
Felonious Homicides	56	61	0.55	0.49–0.68
Non-fatal Assaults	92	109	0.60	0.53–0.74

2015–2017	Hisp. citizens	White citizens	Odds ratios	Confidence interval[a]
Fatally Shot by Police	511	1421	–	–
Benchmark				
Felonious Homicides	13	61	1.69	1.51–2.25
Non-fatal Assaults	51	109	0.77	0.69–1.03

In short, the best benchmark evidence available to us clearly does not evidence the idea of racial bias in police shootings.

One similar line of evidence to the benchmarking studies uses a detailed list of 120 relevant descriptors such as decedent characteristics, criminal activity, threat levels, police actions, and the setting of the lethal interaction to predict which race is more likely to be shot given equality among the descriptors. When this analysis is done, Blacks are found to be equally likely to be shot as are Whites [1036].

This is all also consistent with studies having to do with training simulations which measure whether or not police are quicker to shoot Blacks than Whites. Since this line of evidence is experimental, there cannot be any unspecified variables of relevance; the only potential concern is relevance to the real world. Police hesitate more before shooting Blacks, and shoot Whites more often [1037, 1038, & 1039].

Yet another line of evidence yields results which are contrary to the predictions made by the belief that racism causes the shooting inequality; the Black-White inequality in the rate at which people are killed by police is lowest in the South and highest in the Northeast and Midwest [1035]:

Source 1035 - Figure 2:

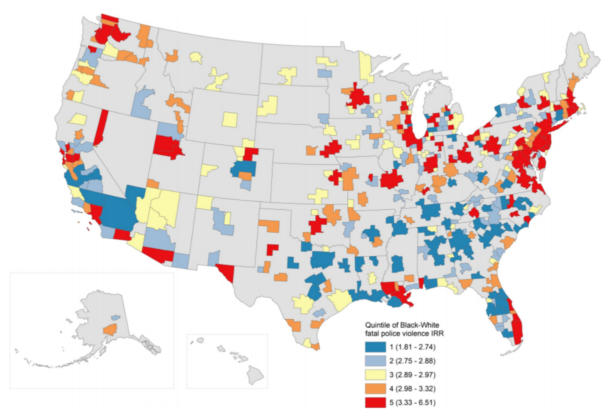


Fig 2. Estimated Black-White incidence rate ratios (annual) for fatal police violence, by MSA. Estimated MSA-specific incident rate ratios comparing rates of fatal police violence experienced by Black people relative to those experienced by White people are mapped. Quintiles are labeled in the legend along with, in parentheses, the range of IRR values included in that quintile.

https://doi.org/10.1371/journal.pone.0229686.g002

The final relevant line of evidence is the consistent finding that Black officers are as likely to use force against Blacks as are White officers. Source 1040 for instance finds that nationally, Blacks are 33% of those killed by non-White officers, and 28% of those killed by White officers. Source 1033 also finds that the race of officers involved in fatal shootings is unrelated to the probability of the target being Black or Hispanic, but use of the paper is controversial because the paper has been retracted [1041] due to concerns [1042] of its results being misinterpreted . This retraction however, is irrelevant to the current use of the paper because the paper is still equipped to address how the racial composition of officers relates to the shooting inequality [1041].

Sentencing:

Once stopped, searched, and arrested, there is of course the potential issue of bias in sentencing given equal cases and behavior. As will be seen, there does not seem to be reason to think that there is much of an anti-Black bias in criminal sentencing in general when the following is considered:

1. What isn't accounted for by the regression results of the general research literature [more here].
2. Evidence on racial biases from mock jury experiments [more here].
3. Evidence on the effects of Black judges and Black lawyers on sentencing decisions [more here].

There is also evidence against there being an appreciable anti-Black bias in the assignment of death penalty sentences [1137 & 1138], and there is evidence against the race of victims having an appreciable effect on sentencing outcomes [1139, 1140 & 1141].

-Pre-Trial Outcomes:

In this meta-analysis [927] (k=36), Wu argues that pre-trial decisions are very important because 80% of state cases and 90% of federal cases never actually go to trial, and he finds that Black defendants are 9% more likely than White defendants to be charged:

Source 927 - Table 3:

TABLE 3: The Effect Size Estimate and Q Statistic

Moderator Variable	Mean Effect Size[a]	SE	95% CI Lower[a]	95% CI Upper[a]	Q	k
Race and ethnicity	1.093 (0.089)**	0.031	1.028 (0.028)	1.162 (0.150)	95.730***	36

However, there are several interesting findings in the moderator analysis. The first is that this effect is only found in the South. This is consistent with the standard narratives about the distribution of racism throughout the United States. However, there are two other

Source 927 - Table 4:

TABLE 4: Effect Size Analyses of Random-Effects Mean Odds Ratios by Moderators

Moderator Variable	*M* ES	95% CI Lower	95% CI Upper	*z*	*p*	*Q*	τ^2	*k*
Panel A: Methodological (sample or analytic) moderators								
Type of publication					.603	0.271		
Nonrefereed publication	1.150	0.958	1.380	1.503	.133		.016	4
Refereed journal article	1.091	1.012	1.177	2.269	.023		.019	32
Region					.024	7.456*		
Non-South	1.061	0.997	1.130	1.852	.064		.009	25
South	1.411	1.108	1.799	2.786	.005		.068	8
Multiple/not reported	0.983	0.888	1.087	−0.339	.735		.000	3
Type of jurisdiction					.005	7.918**		
Single	1.159	1.054	1.275	3.048	.002		.027	26
Multiple	0.999	0.959	1.042	−0.042	.966		.000	10
Year of data					.908	0.013		
Prior to 1991	1.089	1.022	1.162	2.611	.009		.008	19
1991 or later	1.100	0.940	1.288	1.192	.233		.049	17
Type of standard error					.012	6.240*		
Provided by study	1.029	0.946	1.120	0.676	.499		.017	25
Estimated	1.224	1.100	1.361	3.717	.000		.014	11
Statistical method for effect size					.277	2.567		
Logistic	1.116	1.042	1.195	3.125	.002		.013	32
Probit	1.108	0.988	1.242	1.755	.079		.000	2
Hierarchical linear modeling	0.875	0.654	1.169	−0.905	.365		.032	2
Coding for race and ethnicity					.351	0.870		
Black or Hispanic vs. White (single)	1.063	0.965	1.172	1.239	.215		.020	16
Minority vs. White (combined)	1.132	1.036	1.237	2.745	.006		.012	20
Prosecutorial decision point					.015	5.947**		
Screening	1.205	1.071	1.356	3.105	.002		.029	21
Prosecution	1.021	0.958	1.087	0.634	.526		.006	15
Panel B: Theoretical moderators								
Controls for all three primary legal factors (i.e., crime severity, criminal history, and strength of evidence)					.048	3.912*		
Yes	1.275	1.053	1.542	2.494	.013		.046	14
No	1.043	0.985	1.104	1.445	.149		.007	22
Controls for evidentiary strength					.079	3.076		
Yes	1.249	1.031	1.515	2.268	.023		.049	16
No	1.044	0.986	1.105	1.490	.136		.007	20
Controls for victim characteristics					.991	0.000		
Yes	1.092	0.916	1.302	0.983	.325		.014	13
No	1.093	1.024	1.168	2.658	.008		.013	23
Controls for victim–offender relationship					.497	0.461		
Yes	1.158	0.931	1.441	1.322	.186		.067	12
No	1.071	1.008	1.139	2.207	.027		.009	24

findings from the moderator analysis (see Table 4 of source 927 in the right column) which cast doubt on the idea that the meta-analysis is detecting any real bias:

These findings are that:

1. Contrary to what we would expect if racial animus were the cause, the strength of this effect has not changed over time.
2. No bias was found in studies that reported their standard error.

Standard error is a statistic which is needed to put a result into a meta-analysis. Some studies used in this meta-analysis reported their standard errors while others did not. So, how did Wu use studies that don't report standard error statistics when standard error is a statistic required for meta-analysis? He did so by estimating what he thought that their standard error statistics probably were. When the standard error was not reported, Wu estimated what the standard error probably was. The 25 studies which reported their standard error statistics found no effect while the entire meta-analytic effect was driven by the 11 studies which did not report their standard error statistics. The racial bias detected when including the non-reporting studies was already unsubstantial, but this also suggests that the small bias that was found is just a result of upwardly biased estimation.

-Post-Trial Outcomes:

This line of research looks at real world sentencing outcomes and is concerned with whether or not there are racial disparities which cannot be attributed to non-race factors.

Source 608 looks at this meta-analytically, examining 116 sentencing contexts: 101 State level sentencing contexts and 15 Federal. This produced 282 effect sizes: 258 State, 24 Federal. Of these, 37% of admitted papers were unpublished. Of the unpublished studies, 50% were doctoral dissertations.

For State sentencing, the raw effect size when looking at all studies was that without controlling for anything, Blacks were 28% more likely than Whites to receive a harsh sentence. Of the unpublished studies, the raw effect size was that Blacks were 14% more likely to receive a harsh sentence. This indicates either that the main meta-analytic effect size is inflated by publication bias, that the doctoral dissertations have smaller effect sizes because they are more rigorous, or both.

For all studies, it is also found that controlling for criminal history and offense severity shrinks the disparity from 28% to 14%.

For Federal sentencing, the raw effect size was a 15% disparity with unpublished studies having larger effect sizes. The trend for unpublished Federal studies however is not noteworthy because they are small in number, and because they produce an enormous confidence interval ranging from 7% to 136%.

Other noteworthy findings are that:

- Smaller estimates of unwarranted sentencing disparity were found in analyses that controlled for more variables.
- Similarly, studies which use better measures of offense severity and criminal history find smaller percentages of the disparity to be inexplicable.
- When Judges have more personal discretion over sentencing outcomes, the racial disparities are larger. However this effect is weak, and is entirely moderated by confounders.
- In Southern jurisdictions, inexplicable disparities are larger, but this is accounted for by methodological characteristics of the Southern studies.
- Federal data prior to 1980 showed inexplicable disparities of 2% while more modern analyses show inexplicable disparities of 58%. This doesn't align with narratives of the criminal justice system being highly discriminatory in the past before reforms were made.

Looking at the State level sentencing disparity (28%), the part of the disparity which cannot be explained by criminal history or by offense severity (14%) was statistically significant, and the authors note in the conclusions that this doesn't look good for the thesis of there being no discrimination. However, this is odd for them to say because they extensively discuss the possibility of potential confounders other than criminal history and offense severity, and because they go through the work of showing that inexplicable disparities are smaller in the better analyses that control for more confounders and which control for better confounders.

Potential Confounders:

The belief that part of the sentencing disparity is inexplicable by relevant confounds and thus attributable to a direct effect of racial discrimination is a dangerous position to be in because one can always just control for more confounders. The authors themselves discuss many of these at length.

The first to consider are sample differences in various demographic variables such as age, sex, socioeconomic status, geographic location, etc. Older people tend to be sentenced for smaller periods [963], and to be convicted less often [963]. In addition, Blacks tend to be younger than Whites [964]:

	Totals		Race											
			White alone		Black or African American alone		American Indian and Alaska Native alone		Asian alone		Native Hawaiian and Other Pacific Islander alone		Two or more races	
	Persons		Persons		Persons		Persons		Persons		Persons		Persons	
	Sum	PCT	Sum	PCT	Sum	PCT	Sum	PCT	Sum	PCT	Sum	PCT	Sum	PCT
Totals	299,035	100.0%	238,692	79.8%	37,815	12.6%	2,642	0.9%	13,464	4.5%	767	0.3%	5,655	1.9%
Age														
00 to 17	74,304	100.0%	56,287	75.8%	11,333	15.3%	799	1.1%	3,076	4.1%	199	0.3%	2,611	3.5%
18 to 24	28,655	100.0%	22,174	77.4%	4,180	14.6%	311	1.1%	1,236	4.3%	107	0.4%	647	2.3%
25 to 34	40,167	100.0%	31,295	77.9%	5,297	13.2%	402	1.0%	2,321	5.8%	152	0.4%	699	1.7%
35 to 44	41,898	100.0%	33,249	79.4%	5,246	12.5%	398	1.0%	2,343	5.6%	117	0.3%	545	1.3%
45 to 54	43,586	100.0%	35,557	81.6%	5,180	11.9%	349	0.8%	1,904	4.4%	97	0.2%	498	1.1%
55 to 59	18,436	100.0%	15,377	83.4%	1,937	10.5%	123	0.7%	768	4.2%	30	0.2%	200	1.1%
60 to 64	14,887	100.0%	12,591	84.6%	1,473	9.9%	88	0.6%	561	3.8%	23	0.2%	152	1.0%
65 to 74	19,811	100.0%	16,920	85.4%	1,839	9.3%	105	0.5%	729	3.7%	27	0.1%	191	1.0%
75 to 80+	17,292	100.0%	15,242	88.1%	1,331	7.7%	66	0.4%	525	3.0%	14	0.1%	114	0.7%

Notes: 7 year averages, from 2005 to 2011. Sums in thousands.

This average age difference is due, at least in part, to Blacks producing a higher average amount of offspring than Whites [1086]:

	Totals		Race											
			White alone		Black or African American alone		American Indian and Alaska Native alone		Asian alone		Native Hawaiian and Other Pacific Islander alone		Two or more races	
	Persons		Persons		Persons		Persons		Persons		Persons		Persons	
	Sum	PCT	Sum	PCT	Sum	PCT	Sum	PCT	Sum	PCT	Sum	PCT	Sum	PCT
Totals	299,035	100.0%	238,692	100.0%	37,815	100.0%	2,642	100.0%	13,464	100.0%	767	100.0%	5,655	100.0%
Related Children Under 18: Number														
0	144,253	48.2%	119,128	49.9%	16,003	42.3%	998	37.8%	5,969	44.3%	284	37.1%	1,871	33.1%
1	52,819	17.7%	40,646	17.0%	7,635	20.2%	512	19.4%	2,720	20.2%	141	18.4%	1,165	20.6%
2	57,840	19.3%	45,459	19.0%	7,175	19.0%	515	19.5%	3,169	23.5%	156	20.3%	1,367	24.2%
3	28,567	9.6%	22,296	9.3%	3,999	10.6%	347	13.1%	1,056	7.8%	92	12.0%	778	13.8%
4	10,375	3.5%	7,639	3.2%	1,893	5.0%	150	5.7%	325	2.4%	58	7.5%	310	5.5%
5	3,296	1.1%	2,278	1.0%	687	1.8%	70	2.6%	132	1.0%	23	3.0%	106	1.9%
6	1,199	0.4%	791	0.3%	266	0.7%	30	1.1%	67	0.5%	7	0.9%	38	0.7%
7	431	0.1%	281	0.1%	103	0.3%	13	0.5%	12	0.1%	6	0.8%	15	0.3%
8 or More	255	0.1%	175	0.1%	54	0.1%	7	0.3%	13	0.1%	1	0.1%	6	0.1%

Notes: 7 year averages, from 2005 to 2011. Sums in thousands.

The next thing to consider is that having a private attorney is associated with less punitive sentences [970, 973, 974, & 975], and that Blacks are less likely to have private attorneys [973 & 976]. While arguably a flaw of the justice system, this influence of socioeconomic status is not a racial bias of the justice system [see more on the causes of the socioeconomic differences here]. Yet another thing to perhaps consider is that inequality of educational attainment, whatever the cause of the inequality [see more here], may also lead White defendants to more easily navigate the criminal justice system. To reiterate, these sorts of things are not flaws of the criminal justice system. Rather, their fault lies in whatever causes the non-justice-system inequalities and are to be investigated separately.

The next potential confounders to consider are various legal variables; there are other variables beyond just criminal history and offense severity to consider. These include the degree of premeditation, strength of evidence, differences in pre-trial release status, etc. While legally, strength of evidence isn't necessarily something to be considered in assigning sentence lengths, violent felony cases with forensic evidence and cases with more varied pieces of physical evidence result in longer custodial sentences for convicted defendants [1078]. Pre-trial release status has a strong positive relationship with sentence severity [970, 974, 977, & 978], and Whites are more likely to gain pre-trial release for whatever reason this may be [973 & 978]. Perhaps a result of some kind of bias in some other stage of the criminal justice system, as always, pre-trial release status is separate from sentencing, and it is important to isolate variables in order to properly investigate each one.

The final sort of confounders to look for are variables of court behavior such as good/bad defendant behavior, willingness to testify against partners, willingness to plead guilty, and ability to navigate the court system.

Defendants who plead guilty receive less severe sentences than defendants convicted by trial [979, 980, 981, 982, 983, & 984], and there is evidence that Blacks/minorities are less likely to plead guilty [979, 985, 986, 987]. Source 988 attempts to use verbal IQ as a proxy for court behavior, and finds that it mediates the disparity. However, the analysis was underpowered. The paper says based on NHST results that it finds no evidence of racial discrimination, but this is a type II error.

Lack of direct evidence aside, it is a reasonable, likely true hypothesis that verbal IQ moderates the disparity given that IQ is causally related to criminality [see more here], and given the IQ gap [see chapter 7].

If a variable legitimately confounds the sentencing disparities, and a paper with sufficient statistical power fails to account for it, then the paper will find a disparity which is supposedly inexplicable by factors other than race. This however, is a type I error.

-Mock Juries:

Mock jury experiments sidestep these problems of ambiguity because in them, no differences between defendants exist and there can thus be no omitted variables or concern of causality. However, this advantage is in exchange for concerns that experimental settings are not generalizable to the real world.

Source 989 analyzed data from 34 such studies where people acted as jurors and voted on whether or not a given defendant was guilty and on sentence length. It was found that Whites have nearly no bias in such decisions (0.028*d* & 0.096*d* for verdict and sentencing decisions respectively) while the Blacks exhibited a moderate in-group bias (0.428*d* & 0.731*d* for verdict & sentencing respectively).

A more recent meta-analysis [990] once again found White jurors to have no bias against Black defendants, but to have a moderate bias against Hispanics defendants. Black jurors, on the other hand, once again expressed a pro-Black/anti-White bias:

Source 990 - Table 1:

Table 1
Meta-Analytic Results for Bivariate Relationships Between Juror/Defendant Characteristics and Guilt Judgments

Juror/Defendant characteristics	N	k	$\bar{r}$	95% CI Lower	95% CI Upper	Q	I^2	BESD IV = 0 (Low)	BESD IV = 1 (High)
Defendant Attractiveness	1599	12	−.04	−.12	.04	22.37*	51	52	48
Defendant Gender	4172	25	.02	−.04	.08	78.06**	69	49	51
Defendant SES	4180	20	−.11**	−.14	−.07	20.61	8	56	45
Defendant Prior Criminal Record	2921	19	.12**	.05	.19	51.02**	65	44	56
Defendant Race	7076	51	.03	−.03	.09	267.58**	81	49	52
W Jurors with W/B or W/H Defendant	5793	39	.01	−.05	.07	224.05**	83	50	51
W Jurors with W/B Defendant	4476	32	−.02	−.09	.06	198.73**	84	51	49
W Jurors with W/H Defendant	1317	7	.11*	.03	.19	13.65*	56	45	56
B Jurors with W/B Defendant	1029	10	.13*	.01	.25	28.33**	68	44	57
Juror Need for Cognition	1747	10	−.07*	−.13	−.00	15.91	43	54	47
Juror Experience	7025	10	.03	−.01	.07	20.03*	55	49	52
Juror Education	8298	20	.00	−.03	.03	31.04*	39	50	50
Juror Authoritarianism	8205	36	.17**	.14	.20	67.72**	48	42	59
Juror Trust in Legal System									
JBS-Total Score	2763	17	.22**	.15	.28	46.24**	65	39	61
JBS-Probability of Commission	3080	12	.16**	.11	.22	24.67*	55	42	58
JBS-Reasonable Doubt	2938	11	.17**	.10	.25	43.55**	77	42	59
Juror Gender	60480	215	.08**	.06	.10	927.18**	77	46	54

This is also consistent with evidence on the degree to which Whites in general racially discriminate [see more here].

This may be taken as suggesting that unexplained parts of the disparity which are observed in the real world are a result of observational research being unable to control for all of the differences between Black criminals and White criminals, seeing as such disparities do not exist in experimental research where moderating variables do not exist.

On the other hand, it may be contended that experimental research is less representative of real people and/or real behavior than the observational research. I am not aware of evidence that this sort of problem affects the results, but there intuitively seems to be less plausibly for this to impede the experimental research than there seems to be for confounders to impede the results of the observational research.

-Black Judges & Black Lawyers:

Importantly, the observational research can be unambiguously taken to evidence that if racial bias exists and/or matters in the criminal justice system, then Blacks have the opposite bias of Whites. This is important because in the real world observational data, Black Judges and Black Lawyers have the same 'racial biases' as White ones do, or rather, both are acting on confounding variables in a race neutral manner while Whites and Blacks differ in these confounding variables.

Turning to lawyers, Black sounding names receive fewer callbacks from lawyers than do White sounding names, a problem which could impact a criminals' legal outcomes, but this tendency is the same among White and Black lawyers [991]. Perhaps also relevant here is the evidence pertaining to callback disparities in hiring [see more here].

Turning to judges, an analysis of 35,000 trials from 1968 to 1974 [993] found Black and White judges to exhibit equal degrees of racial bias both in terms of decisions about guilt and in terms of decisions about sentence length:

Source 993 - Table 2:

TABLE 2

Mean Sentence Severity
by Race of Judge and Race of Defendant*
(controlled for crime severity)

Black Judges		White Judges	
Black Defendants	White Defendants	Black Defendants	White Defendants
27.9 (4897) +9.4%	23.3 (1089) −8.6%	26.1 (19447) +2.4%	21.2 (4917) −16.8%
partial $r = .11$** interracial percentage difference = 18.0		partial $r = .14$** interracial percentage difference = 19.2	

*Based on the 93 point severity scale, the sentence mean = 25.5.
**Statistically significant at the .001 level.

Similarly, an analysis of 40,000 sentences that were given in Pennsylvania between 1991 and 1994 [992] finds the impact of being Black on a person's sentence to not significantly differ between Black and White judges:

Source 992 - Table 2:

TABLE 2

Race-of-Judge Partitioned Analysis of the Effects of Case Characteristics and Judge Characteristics on In/Out and Length-of-Term Outcomes

	In/Out Probability Effect			Length of Term Sentence Length (months)		
Variable	Black Judge	White Judge	Black/ White Difference[a]	Black Judge	White Judge	Black/ White Difference
Prior record score	.081	.070	.011$^{n.s}$	2.375	2.263	.111$^{n.s}$
Offense severity	.138	.114	.024	6.840	8.325	–1.485
Number of convictions	.006$^{n.s}$	.005$^{n.s.}$	.001$^{n.s}$	3.139	1.579	1.561
Trial	.069$^{n.s}$	.106	–.037$^{n.s}$	12.442	16.333	3.891$^{n.s}$
Female offender	–.069	–.108	.039$^{n.s}$	–2.943$^{n.s}$	–3.094	.151$^{n.s}$
Black offender	.019$^{n.s}$	.062	–.043$^{n.s}$	–.459$^{n.s}$	–.944$^{n.s.}$	.485$^{n.s}$
Age of offender	–.005	–.003	–.002$^{n.s}$	–.080$^{n.s}$	–.069	–.011$^{n.s}$
Violent offense	–.007$^{n.s}$	.015$^{n.s.}$	–.022$^{n.s}$	10.443	11.292	–.849$^{n.s}$
Property offense	.093	.056	.036$^{n.s}$	1.146$^{n.s}$	–.338$^{n.s.}$	1.484$^{n.s}$
Drug offense	.175	.119	.056$^{n.s}$	–2.706$^{n.s}$	–5.139	2.433$^{n.s}$
Age of judge	.005	.009	–.004$^{n.s}$	.129	.037$^{n.s.}$	.092$^{n.s}$
Time on bench	–.012	–.006	–.005$^{n.s}$	–.072$^{n.s}$	.457	–.529
Intercept	.446	.486	–.040	–34.582	–39.124	4.543$^{n.s}$
Model chi square	1615.96	10481.80				
df	12	12				
Percentage correctly placed	84.0	80.3				
R^2					.422	.425
Adjusted R^2					.420	.425

[a]Black minus white difference calculated from unrounded figures.
$^{n.s.}$*Not* statistically significant at $p < .01$.

What Of The Gaps?

As we have seen, according to crime victimization data, the Black-White crime gap really is a crime gap rather than just an arrest bias [more here], and there is evidence against racial biases in stops and searches [more here], in arrests [more here], and in criminal sentencing [more here]. Given this, we may wonder why the crime gap exists. There are a couple of plausible explanations which are to be investigated; here are a couple of them which are (✓) or are not (×) important:

1. Poverty (×):
    - While there is a correlation between poverty and crime, poverty does not cause crime [more here].
    - The Black-White crime gap is still existent when economic variables are held constant [more here].
    - The intergenerational effects of wealth generally fade within two generations of their onset [more here].

2. Family Structure [more here] (×):
    - Very little variance in criminality covaries with family structure.
    - The Black-White crime gap is still existent when family structure is held constant.
    - The causality of what little correlation there is, is questionable.

3. Lead [more here] (×):
    - The Black-White gap in lead exposure is very small and so should not account for much of the crime gap.

4. Child Abuse [more here] (×):
    - Child abuse has a substantial, causal effect on criminality, and Blacks are (relatively) substantially more victimized. However, child abuse is rare enough among both races that it only accounts for roughly 0.28624831% of the Black-White crime gap.

5. Education [more here] (×):
    - Blacks have more educational opportunity than Whites.

6. Aggression & Testosterone (✓ & ×):
    - The Black-White crime gap is partially mediated by differences in self reported aggression [more here].
    - This is not due to Black-White differences in testosterone levels because in general, testosterone does not cause aggression [more here].

7. IQ [more here] (✓):
    - With IQ held constant, the Black-White prison population gap is divided by 2.6.

8. Self Control [more here] (✓):
    - The Black-White crime gap is likely substantially moderated by Black-White differences in self control.

Finally, worth noting is that Black adoptees have more run-ins with the law than non-Black adoptees [1143].

-Poverty:

Blacks are poorer than Whites [1067], and the poor tend to commit more crime [1079]:

Source 1079 - Table 2.4.3:

TABLE 2.4.3 Individual Social Status (Income/Wealth) and Criminal/Delinquent Behavior.

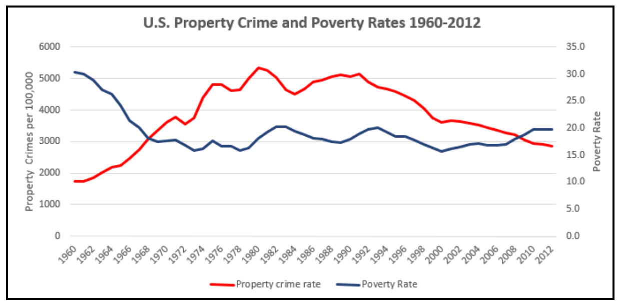
Nature of the relationship	Officially detected offenses				Self-reported offenses	
	Violent offenses	Property offenses	Delinquency	General & adult offenses	Overall offenses	Illegal drugs
Positive						
Not signif.						**EUROPE** *Britain*: Buchmueller & Zuvekos 1988 **NORTH AMERICA** *United States*: Gill & Michaels 1992; Register & Williams 1992; Kaestner 1994
Negative	**NORTH AMERICA** *United States*: Kaplan & Reich 1976	**NORTH AMERICA** *United States*: Cameron 1964 (shoplifters); E Yates 1986 (shoplifters)	**NORTH AMERICA** *United States*: Laub & Sampson 1994:245	**NORTH AMERICA** *United States*: Paez 1981:44	**NORTH AMERICA** *Canada*: SW Baron 2003 (self-control statistically controlled); *United States*: JB Ray et al. 1983 (shoplifters); RH Moore 1984 (shoplifters); Laub & Sampson 1994:245	

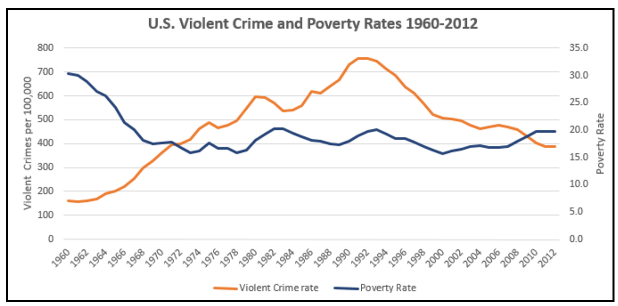
Meta-analyses on the subject, taken together, also show that while the literature is inconsistent, it falls more towards saying that areas with higher poverty have higher crime rates [1079, 1080, & 1081]. As for meta-analytic effect sizes, source 1082 meta-analyzed 153 studies on poverty and crime by geography, and found a correlation of .253. Similarly, source 1083 meta-analyzed 37 studies looking at predictors of national crime rates. For national wealth the mean effect size was -.055 and not statistically significant. For income inequality, the mean effect size ranged from .224 to .416, depending on how income inequality was measured. In both cases, the effect size was statistically significant. Unemployment's relationship with crime (across only 4 studies) was .043 and not significant.

However, correlation is not necessarily causation. There are alternative explanations to a raw correlation other than poverty causing crime. One may be that it is the opposite, that crime destroys wealth by destroying property and making business move away. Another may be that variables which are associated with crime (low self-control, aggression, stupidity, etc) cause both lower wealth and higher crime rates. If we look at trends over time, such as federal level poverty data [965] and crime data [966], we see that changes in poverty have historically been negatively correlated with changes in violent crime and property crime:

Correlation matrix:

	Violent Crime	Property Crime
Property Crime	0.89210775	
Poverty Rate	-0.59438495	-0.62145079





Source 1079 also analyzed 8 studies on the relationship between national wealth and crime over time and found the following:

The Relationship Between The States of the Economy and Crime over Time from Ellis, Beaver, and Wright 2009

Crime Type	Studies	Positive	Not Significant	Negative
Violent Crime	8	63%	25%	13%
Property Crime	8	38%	13%	50%
Overall Crime	8	25%	38%	63%

Looking at changes in unemployment, the following is also found:

The Relationship Between Unemployment and Crime over Time from Ellis, Beaver, and Wright 2009				
Crime Type	Studies	Positive	Not Significant	Negative
Violent Crime	21	38%	38%	24%
Property Crime	15	73%	0%	27%
Overall Crime	31	61%	13%	26%

Finally, source 1084 analyzed 35 reported national level time-series associations and found only 60% of them to be positive and statistically significant. In all, the time-series data inspires even less confidence than the raw effect sizes.

However, better evidence against causality for the poverty-crime correlation is evidence from Swedish family data [1085]. This study analyzed over half a million Swedes and how their childhood income levels related to their future criminality. In line with previous research, the study found that children from poor families were more likely than average to grow up and become criminals. However, some of these families became wealthier, and when this happened, the younger siblings who were only just then growing up were still more criminal. Since 'poor' families turn out more criminal whether or not they are actually impoverished, this indicates that the association between poverty in crime is caused entirely by family level factors other than poverty, whether they be genetic or environmental.

For the context of race, it is worth mentioning that even if we were to accept the association as causal, Blacks would still be substantially more criminal than Whites when economic variables are accounted for [967, 968, & 969].

-Family Structure:

It is popular among conservatives to point to the Black-White single motherhood gap as an explanation of the criminality gap. Indeed, there is a large Black-White gap in family structure [1087]:

Source 1087 - Figure 1:

THE ORIGINS OF AFRICAN-AMERICAN FAMILY STRUCTURE 141

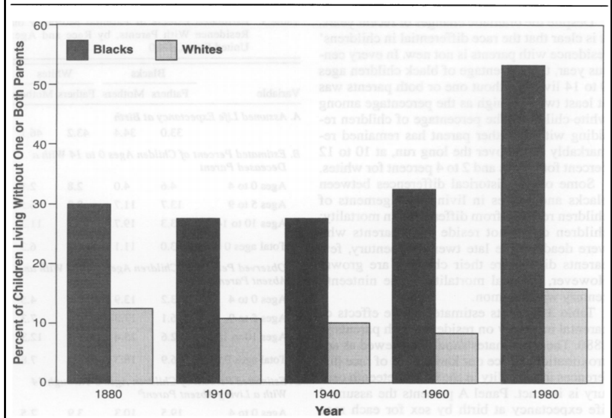


Figure 1. Percentages of Children Ages 0 to 14 With One or Both Parents Absent, by Race: United States, 1880–1980

This is driven by high out of wedlock births; of those married, divorce rates among Blacks and Whites are very similar [1091]:

Source 1091 - Table B:

Table B. Marriage Experience for Women, by Age, Race, and Hispanic Origin: 1975, 1980, 1985, and 1990
(Universe is women 20 to 54 years)

Category	All races				White				Black				Hispanic origin[1]		
	1975	1980	1985	1990	1975	1980	1985	1990	1975	1980	1985	1990	1980	1985	1990
Percent ever married															
20 to 24	62.5	49.5	43.3	38.5	64.9	52.2	46.6	41.3	47.5	33.3	23.9	23.5	55.4	56.7	45.8
25 to 29	87.2	78.6	74.0	69.0	88.8	81.0	77.4	73.2	76.5	62.3	53.4	45.0	80.2	78.4	69.6
30 to 34	93.1	89.9	85.8	82.2	93.9	91.6	88.1	85.6	87.1	77.9	70.9	61.1	88.3	88.0	83.0
35 to 39	95.5	94.3	91.6	89.4	96.2	95.3	93.1	91.4	90.1	87.4	80.7	74.9	91.2	91.6	88.9
40 to 44	95.8	95.1	94.6	92.0	95.9	95.8	95.6	93.4	95.1	89.7	86.1	82.1	94.2	90.3	92.8
45 to 49	95.9	95.9	94.4	94.4	95.9	96.4	95.1	95.1	95.4	92.5	88.4	89.7	94.4	91.1	91.7
50 to 54	95.8	95.3	95.2	95.5	96.0	95.8	95.4	96.1	94.6	92.1	93.4	91.9	95.0	92.5	91.8
Percent divorced after first marriage															
20 to 24	11.2	14.2	13.9	12.5	11.3	14.7	14.4	12.8	10.6	10.5	11.0	9.6	9.4	11.0	6.8
25 to 29	17.1	20.7	21.0	19.2	17.7	21.0	21.5	19.8	15.3	20.2	18.2	17.8	13.9	14.8	13.5
30 to 34	19.8	26.2	29.3	28.1	20.0	25.8	29.0	28.6	20.5	31.4	34.4	26.6	21.1	19.2	19.9
35 to 39	21.5	27.2	32.0	34.1	21.2	26.7	32.0	34.6	22.7	32.9	34.6	35.8	21.9	26.3	29.7
40 to 44	20.5	26.1	32.1	35.8	19.7	25.5	32.0	35.2	27.4	33.7	36.9	45.1	19.7	22.8	26.6
45 to 49	21.0	23.1	29.0	35.2	20.3	22.7	28.4	35.5	26.9	29.0	36.0	39.8	23.9	24.3	24.6
50 to 54	18.0	21.8	25.7	29.5	16.8	21.0	24.6	28.5	29.7	29.0	33.7	39.2	22.5	21.8	22.9
Percent remarried after divorce															
20 to 24	47.9	45.5	44.3	38.1	50.1	47.0	46.0	39.3	(B)	(B)	(B)	(B)	(B)	(B)	(B)
25 to 29	60.2	53.4	55.3	51.8	62.0	56.4	58.3	52.8	43.1	27.9	25.4	44.4	(B)	50.5	49.5
30 to 34	64.4	60.9	61.4	59.6	67.5	63.3	64.3	61.4	41.8	42.0	41.1	42.0	58.3	44.9	45.9
35 to 39	69.5	64.9	63.0	65.0	70.9	66.9	64.9	66.5	62.6	50.6	44.8	54.0	45.2	57.1	51.2
40 to 44	69.7	67.4	64.7	67.1	71.9	68.6	67.5	69.5	57.1	58.4	45.4	50.3	(B)	50.6	53.9
45 to 49	69.6	69.2	67.9	65.9	70.7	70.4	69.6	67.2	61.7	62.7	54.6	55.0	(B)	78.9	51.0
50 to 54	73.5	72.0	68.2	63.0	73.4	72.6	68.4	65.4	73.7	72.7	64.3	50.2	(B)	(B)	62.2
Percent redivorced after remarriage															
20 to 24	(NA)	8.5	8.7	13.1	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
25 to 29	(NA)	15.6	18.2	17.8	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
30 to 34	(NA)	19.1	20.0	22.7	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
35 to 39	(NA)	24.7	26.9	28.5	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
40 to 44	(NA)	28.4	33.0	30.6	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
45 to 49	(NA)	25.1	33.8	36.4	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
50 to 54	(NA)	29.0	27.3	34.5	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)

However, the correlation between single motherhood and delinquency, though existent, is rather small; source 1088 reviewed 5 previous meta-analyses, and the effect sizes were .07, .09, .09, .10, and .10, meaning that single motherhood explains, at most, 1% of individual level variance in criminality. One of the more recent meta-analyses [1089], covering 72 studies, also found the relationship to be weaker among older teenagers. As we'd expect from this, an association between race and crime remains when controlling for family structure [969].

A final thing to be considered with respect to crime and single motherhood is that the kinds of fathers who leave their kids behind tend not to be the most morally upright people. Empirically, fathers which don't live with their children are much more likely to be engaged with drug use, criminal activity, have high levels of psychopathy, etc [1090 & 1092]. Moreover, source 1092 finds that while kids who interacted with their fathers more were less likely to have conduct problems, this relationship only held for fathers who had low levels of antisocial behavior; fathers who had greater levels of antisocial behavior actually adversely affected their kids' level of conduct problems. More directly relevant, source 1093 finds that Black, inner-city children living without their fathers are actually less aggressive than their fathered counterparts.

-Lead:

A meta-analysis [1094] of 19 studies with an aggregated 8,561 participants found a statistically significant correlation of .19 between conduct problems and lead exposure among children and adolescents. The same is found when looking at lead exposure and criminality by region [1095, 1096, 1097, & 1098]. There also used to be a slight Black-White gap in lead exposure such that Blacks had a mean blood lead level that was ~1.4 ug/dl higher than that of Whites [726]. However, blood lead levels no longer significantly differ by race [727]. Given this, even though lead impacts crime, the fact that the races barely differ in terms of lead exposure suggests that lead probably plays little to no role in the Black-White crime gap. This is consistent with sources 1097 and 1098 which find that the proportion of an area which was Black continued to predict its crime rate even after its degree of lead exposure was controlled for.

-Child Abuse:

According to the U.S. Department of Health and Human Services' 2013 report on child maltreatment [1099], the rate at which children suffer from abuse is roughly 14.6 per 1,000 for Blacks, 8.5 per 1,000 for Hispanics, and 8.1 per 1,000 for Whites. These victimization rate differences are not explained by reporting biases, the report shows that Blacks are also overrepresented among those who die from child abuse.

Child abuse also causes criminality. The relationship remains in twins [1100 & 1163], meaning the more abused twin becomes more criminal. This rules out the possibility of genetic confounding. The relationship also remains when controlling for birth order, maternal education, paternal criminality, religion, and family structure [1100]. However, the degree to which being abused increases the likelihood of criminality is hard to estimate. Studies vary in their definitions of abuse, the set of statistical controls they employ, and their measurement of criminality. Because of this, estimates of how much a person's chance of criminality is increased by abuse range from 28% [1101] to 200% [1102]. No meta-analysis of this data has been done and so there is no simple way to judge the true effect. We can however say for sure that some of the Black-White crime gap is caused by the Black-White gap in child abuse.

With these effect sizes, we can devise a rough estimate of how much child abuse contributes to the crime gap, but we also need some perspective on how many people are imprisoned. Source 1103 gives us the numbers of people imprisoned per 100,000 U.S. residents by race and sex:

Source 1103 - Table 10:

TABLE 10
Imprisonment rate of sentenced state and federal prisoners per 100,000 U.S. residents, by demographic characteristics, December 31, 2014

		Male					Female				
Age group	Total[a]	All male[a]	White[b]	Black[b]	Hispanic	Other[b]	All female[a]	White[b]	Black[b]	Hispanic	Other[b]
Total[c]	471	890	465	2,724	1,091	968	65	53	109	64	93
18–19	169	317	102	1,072	349	542	14	8	32	17	12
20–24	746	1,365	584	3,868	1,521	1,755	96	72	152	94	109
25–29	1,055	1,912	958	5,434	2,245	2,022	170	150	244	165	208
30–34	1,161	2,129	1,111	6,412	2,457	2,193	185	163	264	174	225
35–39	1,067	1,982	1,029	6,122	2,272	1,878	155	138	229	137	189
40–44	904	1,689	942	5,105	1,933	1,619	132	119	213	107	174
45–49	758	1,417	815	4,352	1,602	1,444	111	90	203	94	161
50–54	567	1,081	633	3,331	1,320	1,112	72	57	128	67	124
55–59	358	698	400	2,178	978	832	37	27	72	42	63
60–64	212	422	252	1,265	680	483	20	15	37	25	37
65 or older	72	158	109	418	299	208	5	4	8	7	12
Number of sentenced prisoners[d]	1,508,636	1,402,404	453,500	516,900	308,700	123,300	106,232	53,100	22,600	17,800	12,800

Note: Counts based on prisoners with sentences of more than 1 year under the jurisdiction of state or federal correctional authorities. Imprisonment rate is the number of prisoners under state or federal jurisdiction with a sentence of more than 1 year per 100,000 U.S. residents of corresponding sex, age, and race or Hispanic origin. Resident population estimates are from the U.S. Census Bureau for January 1, 2015. Alaska did not submit 2014 data to the National Prisoner Statistics (NPS), so totals include imputed counts for this state. See *Methodology*.

[a]Includes American Indians and Alaska Natives; Asians, Native Hawaiians, and other Pacific Islanders; and persons of two or more races.

[b]Excludes persons of Hispanic or Latino origin.

[c]Includes persons age 17 or younger.

[d]Race totals are rounded to the nearest 100 to accommodate differences in data collection techniques between jurisdictions.

Sources: Bureau of Justice Statistics, National Prisoner Statistics, 2014; Federal Justice Statistics Program, 2014; National Corrections Reporting Program, 2013; Survey of Inmates in State and Federal Correctional Facilities, 2004; and U.S. Census Bureau, postcensal resident population estimates for January 1, 2015.

Assuming that males and females are both exactly 50% of the population for the sake of simplification, when we average imprisonment rates between the sexes, we get 259 Whites being imprisoned per 100,000 U.S. residents and 1416.5 Blacks being imprisoned per 100,000 U.S. residents.

To understand how to figure out how much of the gap is accounted for by child abuse, let us first understand the math of a simpler, fictional problem. Let's say for the sake of argument that we have group A and group B, and that they combine to create group T (T for total). Group A has 100 members and group B has 200 members. Group T thus has 300 members. 52% of group A dies, and 49% of group B dies. Therefore, 52 people in group A die, and 98 people in group B die. Therefore 150 total people die. Therefore, 150 out of 300 people died, or 50% of all people

Here is the information summarized in a table:

Group:	A	B	T(Total)
Initial #	100	200	300
% dead	52%	49%	50%
# dead	52	98	150

As we can see, the percentage of all people who died is just an average of the two death rates, but weighted by population size. We can just take ((100×52)+(200×49))÷300 to get 50.

Now let's make the same table but focused on the percentage of Whites who are imprisoned, by abuse status (Abused = # abused per 100k):

Group:	Abused	Non-Abused	Total
# of people	810	99,190	100,000
proportion imprisoned	X	Y	259 / 100,000

This is where the complexity comes from; we don't know X or Y. Rather, we only know the percentage of the total population which is imprisoned, and the size of X in terms of the size of Y (X is anywhere from 28% to 200% larger than y). Given the most generous estimate of effect size for child abuse (+200%), we can rewrite X in terms of Y:

Group:	Abused	Non-Abused	Total
# of people	810	99,190	100,000
proportion imprisoned	3Y	Y	259 / 100,000

We can now take the weighted average algebraically:

$259/100{,}000 = ((810\times3Y)+(99{,}190\times Y)) \div 100{,}000.$
$259/100{,}000 = 1.0162\times Y.$
$259/101{,}620 = Y.$

So, 259 in 101,620 non-abused Whites are imprisoned. Since we are assuming that people who suffer child abuse have three times the odds of being imprisoned, we'll say that 777 in 100,810 abused Whites are imprisoned. If we apply these numbers to the number of Whites who are abused and not abused, we would predict that 259 Whites would be imprisoned, which is empirically observed, so our math is correct.

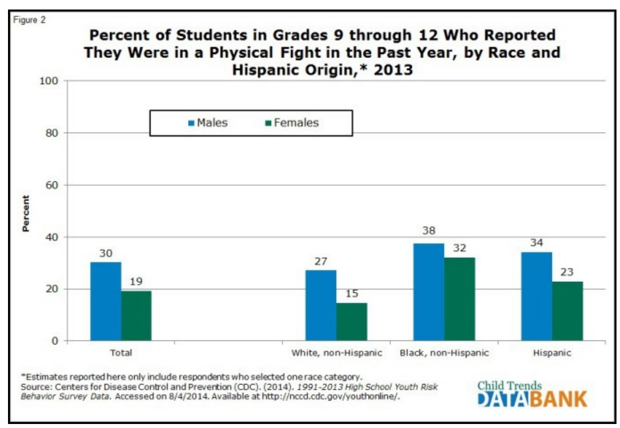
Now, how many Whites would be imprisoned if Whites were abused at the same rate that Blacks are abused? Well, 14.6 per 1,000 Blacks are abused [1099], or 1460 per 100,000. 777/101,620 of these 1460 people would be imprisoned, meaning that 11.16335367 of the 1460 people would be imprisoned. 98,540 of the 100,000 would not be abused. Of these 98,540 people, 259/101,620 would be imprisoned, meaning that 251.1499705 of the 98,540 people would be imprisoned. Adding the two together, we would expect 262.3133241 per 100,000 Whites to be imprisoned. Remember, before accounting for child abuse, 259/100,000 Whites were empirically shown to be imprisoned, and 1416.5 per 100,000 Blacks were empirically shown to be imprisoned. The gap, of 1157.5 people, is thus shrunk by only 3.313324148 people when child abuse is accounted for. In other terms, according to this rough calculation, only 0.28624831% of the Black-White crime gap is accounted for by child abuse rate differences.

In summary, child abuse has a substantial, causal impact on criminality, and Blacks suffer a relatively substantially higher rate of child abuse than Whites do. However, child abuse is rare enough among both races that it can only account for 0.28624831% of the Black-White crime gap.

-Aggression & Testosterone:

One fashionable explanation for criminality in general is that testosterone causes aggression and that aggression causes criminality. A meta-analysis of 45 independent studies totalling 9760 participants [1104] found a weak positive correlation of 0.14, which is already a bad sign for this explanation. The killing blow is that experimental studies which assess what effect there is on aggression when testosterone levels are manipulated find that testosterone is not causal [1105, 1106, & 1107]; aggression increases testosterone levels rather than the other way around.

This being said, Blacks do tend to be more aggressive for whatever reason, and this likely plays a role in the Black-White crime gap. There are multiple lines of evidence for this. The first is that Blacks are more likely than Whites to get into fights at school [1108]:



The second is that Blacks are more likely to bully others than are Whites [1109].

Blacks also score somewhat higher on measures of psychopathic personality. Source 1112 describes such a measure, the Psychopathic Deviate Scale, thusly:

> "This was constructed by writing a number of questions, giving them to criterion groups of those manifesting psychopathic behaviour and "normals", and selecting for the scale the questions best differentiating the two groups. The criterion group manifesting psychopathic behaviour consisted of 17–24 year olds appearing before the courts and referred for psychiatric examination because of their "long histories of delinquenttype behaviours such as stealing, lying, alcohol abuse, promiscuity, forgery and truancy" (Archer, 1997, p. 20). The common feature of this group has been described as their failure to "learn those anticipatory anxieties which operate to deter most people from committing anti-social behaviour" (Marks, Seeman, & Haller, 1974, p. 25). The manual describes those scoring high on the scale as follows: irresponsible, antisocial, aggressive, having recurrent marital and work problems, and underachieving (Hathaway & McKinley, 1989). A number of subsequent studies have shown that the Psychopathic Deviate scale differentiates delinquents and criminals from nondelinquents and non-criminals (e.g. Elion & Megargee, 1975)."

Source 1112 then reviewed 5 studies comparing racial groups on this measure; in Nigeria, Japan, and the United States, Blacks scored .29 to .5 standard deviations higher than Whites:

Source 1112 - Table 1:

Table 1
Psychopathic deviate scale of the MMP1 (*d*)

No.	Location	Test	Blacks	E. Asians	Hispanics	N. Americans	Whites	Reference
1	USA	MMP1	0.29	−0.31	0.00	0.44	0.00	Dahlstrom et al., 1986
2	USA	MMP1-2	0.48	−0.18	0.70	0.74	0.00	Hathaway & McKinley, 1989
3	Japan	MMP1-2		−0.36				Japanese MMP1, 1993
4	Nigeria	MMP1-2	0.50					Nzewi, 1998
5	USA	MMP1-A	0.33		0.36		0.00	Archer, 1997
6	Mean		0.40	−0.28	0.35	0.59	0.00	

Two meta-analyses [1113, & 1114] later reported statistically significant but practically negligible differences, but all samples were either clinical or correctional in nature, meaning they were unrepresentative due to threshold effects, which should downwardly bias differences. The Black-White crime gap does indeed seem to be partially mediated by differences in self reported aggression [988].

-IQ:

Chapter 16 of source 384 meta-analyzed research done on the relationship between IQ and crime, delinquency, and related variables. Of 68 studies on IQ and delinquency, 60 found a negative relation (88%) and the remaining 8 found no significant relationship. Out of 19 studies on IQ and adult criminal offending, 15 (79%) found a negative correlation. Out of 17 studies on self-reported offending and IQ, 14 (82%) found a negative relationship. Out of 5 studies on IQ and antisocial personality disorder, and out of 14 studies on childhood conduct disorder, all 19 found a negative relationship. Thus, the vast majority of research establishes IQ as a correlate of crime and related constructs. On the other hand, only 7 of 19 (36%) of studies on recidivism and IQ found a negative relationship. The authors posit that this is explained by range restriction; to be able to be caught in 2 crimes you have to be dumb enough to commit the first one which means the population of interest has undergone significant range restriction. Source 408 however did a meta-analysis on recidivism going over 32 studies and 21,369 participants and found a -.07 correlation between intelligence and recidivism.

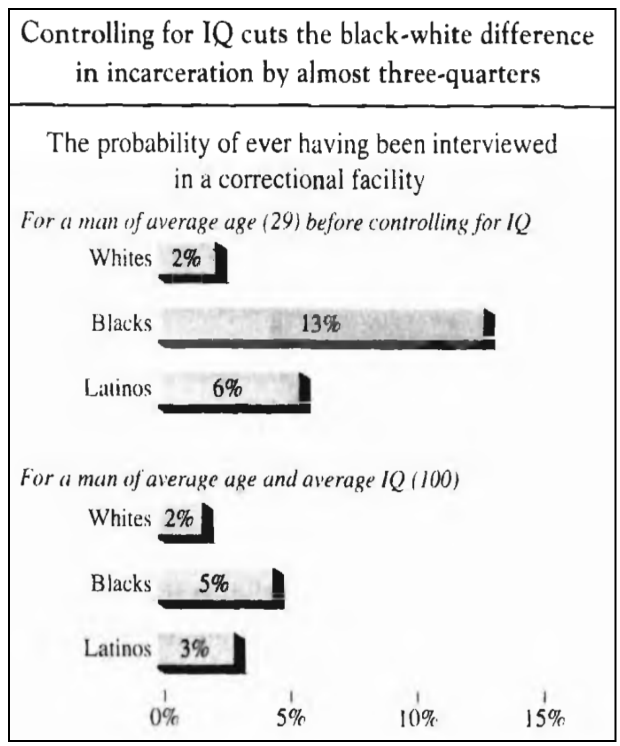
These findings are confirmed by large, representative birth cohort studies in Finland [385], Sweden [386], and the United States [387]. The massive (700,514 participants) study from Sweden [386] found that the negative -.19 correlation between IQ and crime only fell to -.18 when controlling for income and single motherhood.

With regards to the differential detection hypothesis, source [388](388) investigated the impact of neighborhood characteristics and found that the negative relationship with criminality held even after controlling for neighborhood poverty, unemployment, % Black, % female headed household, and % on public assistance, as well as individual age, sex, race, poverty, self-control, and age. Although, the relationship between IQ and criminality was much stronger in well-off areas than it was in disadvantaged areas. We also have evidence like source 389 which compares actual arrests to self report finding no difference in intelligence estimates between methods of assessing criminality. Perhaps self report isn't the best assessment, but the result is certainly not what you would predict if differential detection mattered. Either way, to whatever degree differential detection matters, the impact that IQ has on how your life is affected by run-ins with the law remains the same.

There is also longitudinal evidence linking IQ measured in early childhood to crime later in life. Source 390 conducted a 25-year longitudinal study on 1,625 participants. They found that IQ at age 8-9 predicted criminality in adulthood. This relationship was also found to be mediated by childhood conduct problems, which just tells us that IQ begins to have an effect on criminality at an early age.

A meta-analysis of over 27,000 people from four European twin cohorts [842] on academic performance (i.e. intelligence-proxy) and aggression (parental and self-ratings) finds both within-family associations and between-family associations, thus ending discussion of neighborhood characteristics & shared environment. The twin data also shows genetic mediation between the two, but relationships are still found between MZ twins which implies a role of nonshared environment. The agreement of parental report and self report is also further evidence against the differential detection hypothesis.

This is all of course relevant because there is a well established 1 standard deviation Black-White IQ gap [876, more here], and because when this is accounted for, the Black-White incarceration gap is divided by 2.6 666 - ch. 14]:



-Self Control:

IQ is negatively, though weakly associated with low self control [871], and this association is genetically mediated [1115]. However, this cannot fully explain the heritability of self control because self control is about 50% heritable [1117, 1118, & 1119]. Self control is important because it has power to predict life success which is independent of IQ and socioeconomic status. IQ is of course important to control for because of its predictive power and its collinearity with self control and success. Socioeconomic status is also an important control variable to include because people under emergency financial pressures may be influenced by said pressures to act in a way which is out of line with their true time preference.

Source 1110:

This paper looked at how well self-control measured in childhood (under the age of 10), based on self and peer reported behavior, predicted life outcomes at age 32 in comparison to childhood IQ and parental socio-economic status in a nationally representative sample. Higher childhood self-control was found to predict better health, more wealth, less criminality, and a lower chance of being a single parent in adulthood even controlling for IQ and parental SES. Particularly interesting is the fact that IQ was not predictive of criminality, drug abuse, or single parenthood when parental SES and self-control were controlled for. However, consistent with the past literature, the paper found IQ to be the best predictor of wealth and adult SES.

Source 1120:

Looking at how childhood self-control, IQ, and class predicted adult unemployment in a sample of 16,780 Brits, this paper finds holding the other two variables constant, high self control was related to lower unemployment while social class was not related to unemployment when the other two variables were held constant.

Source 1121:

This paper finds that self control is a better predictor of GPA than IQ and that self control was related to more time being spent on homework while IQ was related to less time being spent on homework.

Source 1123:

This meta-analysis confirms a correlation between self control and various life outcomes such as love, happiness, getting good grades, speeding, commitment in a relationship and lifetime delinquency, but did not assess the mediating roles of IQ or socioeconomic status.

Source 1159:

This meta-analysis found high self control to be related to lower deviancy, with cross-sectional and longitudinal effect sizes being r = .415 and 4 = .335 respectively.

Black-White Differences In Self Control:

Self control is of course relevant to Black-White inequalities in the things that self control is predictive of because there is evidence that Blacks have lower self control than Whites:

Source 1124:

This paper took advantage of a natural semi-experiment which came about due to the military. In the mid 1990s, the U.S. Government offered sufficiently experienced military personnel two options when they retired: they could take a large lump sum of money now or agree to get a yearly payment from the military for the rest of their lives which, over time, would add up to far more than the lump sum. Data was found on the choices of 66,000 individuals, and Blacks were 15% more likely than non-Blacks to take the lump-sum.

Source 893:

In this paper, the Black homes in a sample of 25,820 households were found to have lower savings rates than White homes even controlling for differences in income, age, family size, education, and marital status.

Source 888:

> "Blacks and Hispanics spend roughly 30 percent more on visible expenditures (cars, clothing, jewelry, and personal care items) than otherwise similar Whites."

Source 1122:

This paper utilized a sample of 5,291 university students from 45 countries and gave participants a chance to choose an immediate monetary reward or a larger long term reward; figure 3 shows the proportion of people from different regions that chose the larger and less immediate reward:

Source 1122 - Figure 3:

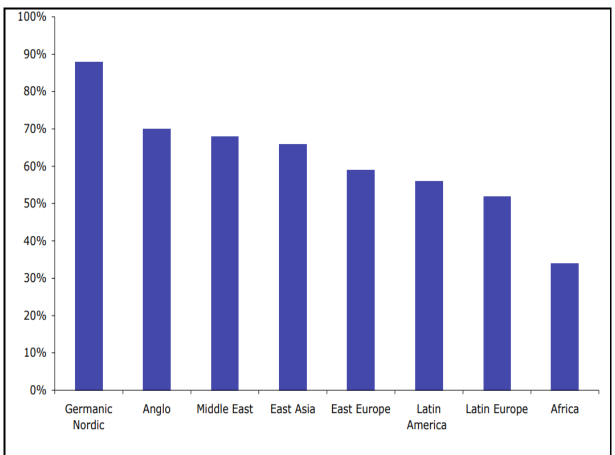


Figure 3: The percentage of choosing to wait grouped by cultural origin

Source 1125:

This paper looked at a sample of 317 individuals with gambling problems and found that White gambling addicts had more self-control than Black gambling addicts even after controlling for education, drug problems, and income.

Source 1126:

The authors of this paper describe their experiment as follows:

> "In our experiment, subjects are asked, orally and in writing, to make twenty decisions in total. For each decision, subjects are asked if they would prefer $49 one month from now or $49+$X seven months from now. The amount of money, $X, is strictly positive and increases over the twenty decisions."

Using this design in a sample consisting of 82% of the student population of 4 middle schools in a poor Georgia school district, the paper was able to measure at what point people began to prefer the later reward and, thus, the strength of their preference for immediate gratification. Blacks were found to have significantly less self-control than Whites.

Source 1127:

This paper looked at a sample of 100 4th grade school children and found that Blacks had lower self control than Whites even after controlling for socio-economic status.

While the within-group heritability of self-control does not necessarily guarantee an above zero between-group heritability of self control, a handful of gene variants which are related to impulsive behavior have also been found to be less common among Blacks than among Whites [1111].

Economic Gaps:

Slavery & Intergenerational Wealth:

Slavery and the intergenerational transfer of wealth acquired in the past cannot explain modern Black poverty because the intergenerational transfer of wealth cannot explain poverty in general; at least not for very long. The speed at which wealth effects fade is very quick. We know this because when families gain large sums of money or have property destroyed, the economic effects entirely fade in under 2 generations, and are mostly gone within a single generation.

The best evidence on this comes from comparison of the descendents of antebellum Blacks who were free before the civil war to the descendents of postbellum Blacks who were freed by the emancipation proclamation. The difference persisted for some time, but after two generations, the two groups of Blacks did not differ in terms of both education and economic success [1130]. This suggests that the direct economic effects of slavery had mostly faded for the grandchildren of slaves. This may seem surprising, but it is consistent with other data on the intergenerational effects of wealth in 19th century America and in the South. For instance, the descendants of those who won Georgia's land-lottery in the 1830s fared no better for it in terms of their income, wealth, and literacy rates [1131] than non-receiving applicants. Analyzing the opposite case, data on those whose wealth was destroyed during the civil war due to slave emancipation and war-related property destruction, a person's wealth being decimated by 10% predicted merely a 0.4% decrease in their child's income by the time the child reached age 50 [1132].

In modern day, data from the entire population of U.S. taxpayers shows that Black children born to parents in the top fifth of the income distribution are equally likely to occupy the top and bottom fifth of the income distribution when they grow up. By contrast, White children born into the top economic quintile are far more likely to stay there than to fall to the bottom [1133]. From 1984-2007, [872] a 10% increase in wealth among an American's grandparents predicted a 1.8% increase in their own wealth if they were White and a 0.2% increase in wealth if they were Black. This may be explained by self control [more here].

More broadly, it is also the case that the impact of various educational effects fade over time [305, 694, & 630].

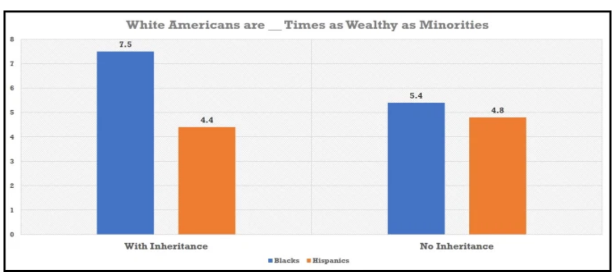
This may seem like a surprisingly short period of time in which to expect the economic effects of major events to vanish, but this is similar-to/greater-than the amount of time it seems to have taken for the Irish to rebound from extreme repression by the English, for the Jews to economically recover following emancipation, and for Japan to recover from the second world war and its damages.

This may seem hard to swallow, but people often overestimate the persistence of environmental effects because from their personal experience, children resemble their parents even well into adulthood, and group differences often persist across generations. However, this is not the appropriate kind of analysis because it is generally confounded. More appropriate would be twin studies that try to ascertain heritability, or adoption studies placing unrelated children into rich homes, or randomized experiments giving poor people large sums of money. A review of 19 twin studies puts the heritability of income in the

United States at 41%, the the contribution of shared environmental factors at just 9%, with the other 50% being explained by nonshared environmental factors such as random luck, measurement error, etc which is incompatible with intergenerational wealth transfer [695]. However, even variance attributed to shared environmental effects cannot automatically be attributed to the effects of intergenerational transfers of wealth since there are other theoretically plausible explanatory influences which are shared among siblings.

There is thus little, if any, non-genetically-mediated transmission of wealth and income within even a single generation. Perhaps slavery is a special case, but the data comparing antebellum free Blacks to postbellum free Blacks gives us reason to doubt this.

Also worth noting is just how much of a gap there is in wealth from raw inheritance. According to a paper from the federal reserve, among Americans who receive no inheritance, the Black-White gap is only 28% than the wealth gap among those who do receive inheritance [1067]:



Another thing to look at from the federal reserve paper is the rate of and median value of inheritance by race:

Race / Ethnicity	% w/ inheritance	Median Value of Inheritance	Average of Median Inheritance Per Person
White	22.9	$55,207	$12,642
Black	10.6	$49,441	$5,271
Hispanic	5.5	$28,708	$1,579

Sometimes it is noted that Black families were broken up in order to sell different family members to separate slave owners, and this is said to explain modern rates of single parenthood among Blacks. However, it is implausible that these old effects explain modern Black family structures because Black rates of single parenthood are far greater today than they were in the 19th century [1087]:

Source 1087 - Figure 1:

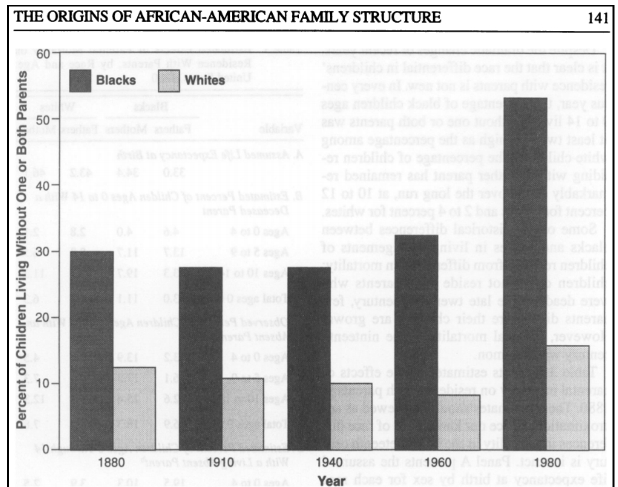
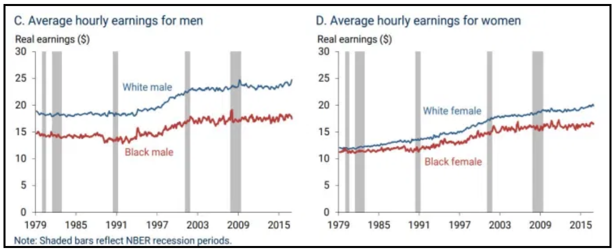


Figure 1. Percentages of Children Ages 0 to 14 With One or Both Parents Absent, by Race: United States, 1880–1980

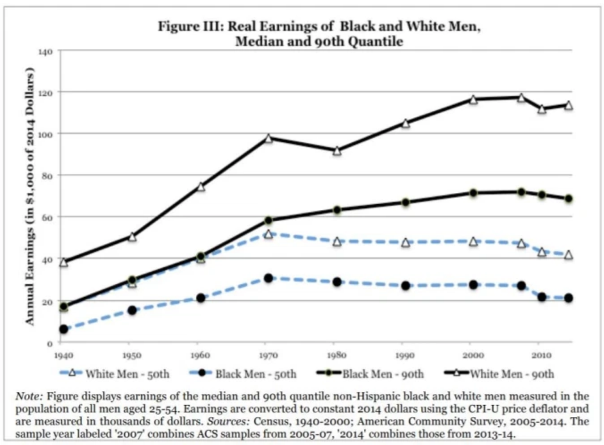
This brings us to yet another reason to doubt that the economic effects of slavery are still in the process of being eliminated: If this were true, then the economic effects of slavery should lessen with each generation, leading us to see slow and steady economic improvements among Blacks. However, nothing like this has taken place for the last half century. Instead, since intelligence is growing more and more valuable in the information age, the Black-White wealth gap has only grown. A 2017 Federal Reserve report [1129] shows that White and Black working women had roughly equal wages in the 1970s and 1980s, but since the 90s a gap has appeared which favors White women. The same report [1129] also shows that for males, there was already a wage gap present in the 1970s and it is even greater today:

Source [1129] - Figure 1 - C & D:



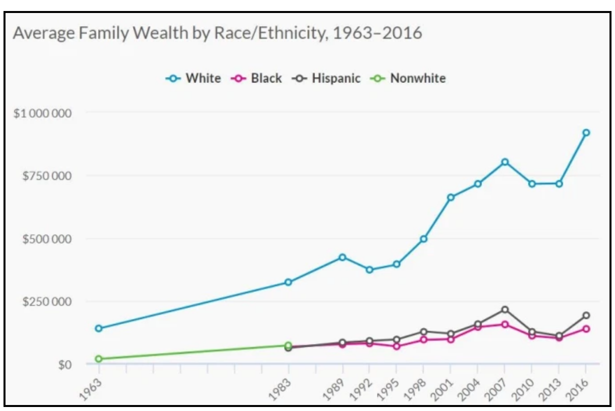
When looking at income for entire population rather than just those who are employed, the trend is more severe [1134]:

Source [1134] - Figure III:



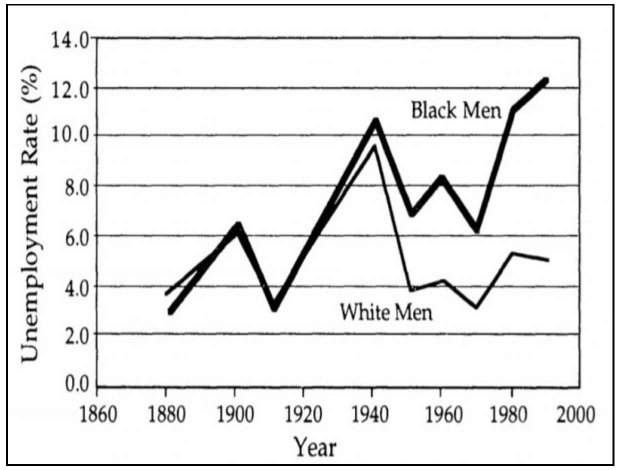
Additionally, the situation is yet more extreme when looking at net wealth instead of income. Since the 1960s, the Black-White Wealth gap has increased many times over [873]:

Source [873] - Figure 3:



Turning to employment, the Black-White unemployment gap appeared sometime in the 1940s and has widened since then [1135]:

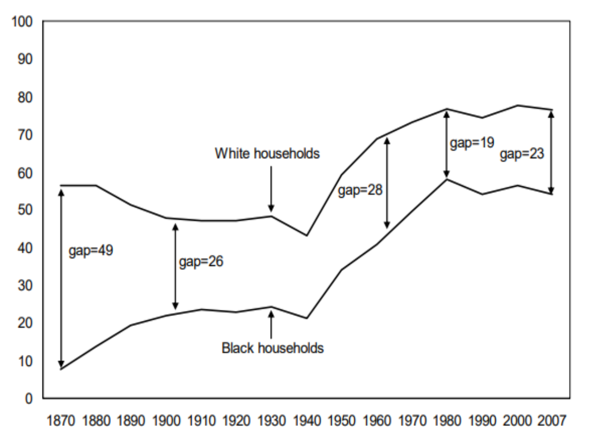
Source [1135] - Figure 1:



The one exception is that one could use home ownership to make a weak case for a slightly narrowing gap [903]:

Source [903] - Figure 1:

Figure 1: Rates of Owner-Occupancy, 1870-2007:
Households Headed by Males, Ages 25-64, in Labor Force, Not in School ("Core Sample")



Overall, there is not much support for the idea that that slavery, or the intergenerational transfer of wealth in general, is responsible for modern Black poverty. Modern Black poverty, therefore, must be explained some other set of factors that continues into modern day, whether it be discrimination, or [behavior].

Educational Opportunity:

There is a Black-White gap in the number years of completed schooling [728]:

Educational Attainment of the Population Aged 25 and Older by Age, Sex, Race and Hispanic Origin, and Other Selected Characteristics
(Numbers in thousands)

Characteristic		High school graduate or more		Some college or more		Associate's degree or more		Bachelor's degree or more		Advanced degree	
	Total	Percent	Margin of error[1] (±)	Percent	Margin of error[1] (±)	Percent	Margin of error[1] (±)	Percent	Margin of error[1] (±)	Percent	Margin of error[1] (±)
Population 25 and older	**212,132**	**88.4**	**0.3**	**58.9**	**0.5**	**42.3**	**0.5**	**32.5**	**0.5**	**12.0**	**0.3**
Age											
25 to 34	43,006	90.5	0.6	65.0	0.9	46.5	0.9	36.1	1.0	10.9	0.6
35 to 44	39,919	88.7	0.5	62.8	0.9	46.7	1.0	36.3	1.0	13.8	0.7
45 to 64	83,213	89.4	0.4	59.0	0.7	42.6	0.7	32.0	0.7	12.1	0.5
65 and older	45,994	84.3	0.7	49.7	0.9	34.1	0.9	26.7	0.8	11.3	0.7
Sex											
Male	101,888	88.0	0.4	57.6	0.7	41.2	0.7	32.3	0.6	12.0	0.4
Female	110,245	88.8	0.3	60.1	0.6	43.4	0.6	32.7	0.6	12.0	0.4
Race and Hispanic origin											
White alone	168,420	88.8	0.3	59.2	0.6	42.8	0.6	32.8	0.6	12.1	0.3
Non-Hispanic White alone	140,638	93.3	0.3	63.8	0.6	46.9	0.7	36.2	0.7	13.5	0.4
Black alone	25,420	87.0	0.9	52.9	1.4	32.4	1.4	22.5	1.2	8.2	0.7
Asian alone	12,331	89.1	1.2	70.0	1.9	60.4	2.0	53.9	2.0	21.4	1.5
Hispanic (of any race)	31,020	66.7	1.1	36.8	1.0	22.7	0.9	15.5	0.7	4.7	0.4
Nativity Status											
Native born	175,519	91.8	0.3	61.3	0.5	43.3	0.6	32.7	0.6	11.9	0.3
Foreign born	36,613	72.0	1.0	47.6	1.1	37.6	1.1	31.4	1.1	12.5	0.7
Disability Status											
With a disability	28,052	78.6	0.9	41.6	1.2	24.9	1.0	16.7	0.9	5.7	0.5
Without a disability	183,351	89.9	0.3	61.5	0.5	45.0	0.6	34.9	0.5	12.9	0.3

However, the question remains regarding whether this is a consequence of differences in educational opportunity, or other factors. Before discussion of gaps in school funding, it should be noted that the raw amount of available funding has little effect on student achievement [1000, 1116, & 1128; more here]. This stated, Black students in grade school now receive more funding. Black school districts receive less funding, but the Blacker schools within the Blacker districts get more funding than the Whiter schools in the Blacker districts [874]. Accounting for this, in 1972, Black students received $0.98 for every dollar spent on White students, and in 1982 this trend reversed such that Black students now receive more funding than White students [733]. This result has achieved replication [734].

One more replication [875] comes to the same finding, as shown in its second table:

Source 875 - Table 2:

Per pupil expenditures for each racial group expressed as a percentage of per pupil expenditures for white students, by state

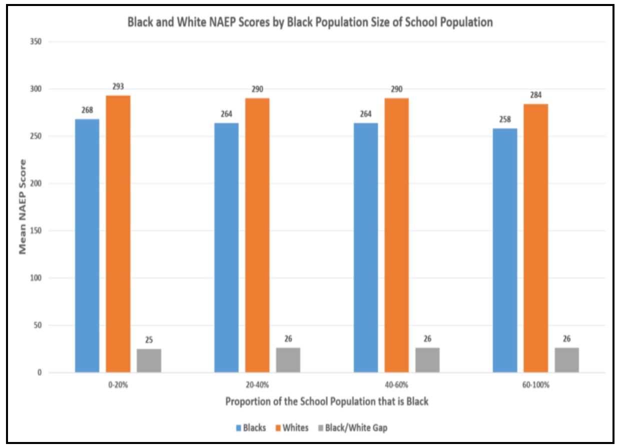
State	Asian	Black	Hispanic	Native American
Alabama	103	97	101	98
Alaska	101	96	100	120
Arizona	98	98	99	104
Arkansas	100	106	99	99
California	94	97	99	109
Colorado	98	103	102	103
Connecticut	100	103	101	102
Delaware	97	99	99	105
Florida	97	99	100	100
Georgia	98	102	100	100
Idaho	100	99	97	101
Illinois	104	93	91	98
Indiana	101	112	108	103
Iowa	99	99	100	100
Kansas	94	95	98	102
Kentucky	99	102	99	99
Louisiana	105	104	103	98
Maryland	106	97	101	97
Massachusetts	108	118	113	108
Michigan	107	107	104	105
Minnesota	104	107	106	116
Mississippi	101	103	99	100
Missouri	106	110	107	101
Montana	97	94	95	100
Nebraska	90	83	93	106
New Hampshire	96	89	83	97
New Jersey	100	117	110	109
New Mexico	94	96	102	86
New York	91	91	90	99
North Carolina	98	100	98	101
North Dakota	99	99	98	98
Ohio	109	111	104	102
Oklahoma	95	97	99	102
Oregon	97	101	99	105
Pennsylvania	96	89	85	98
Rhode Island	100	99	100	105
South Carolina	101	105	101	98
South Dakota	96	93	97	99
Tennessee	100	100	101	98
Texas	90	93	99	100
Utah	95	95	97	117
Vermont	103	101	102	92
Virginia	107	101	105	99
Washington	97	98	98	103
West Virginia	100	99	95	99
Wisconsin	102	98	98	105
Wyoming	98	97	98	97

However, the paper interprets [875] the finding in a bizarre fashion; the authors take issue with the fact that this figure is expressed as a nation-wide average, writing the following:

> “But racial disparities in education spending clearly exist in a host of other states. In Illinois, New York, and Pennsylvania, per pupil expenditures for black and Hispanic students hover around 90 percent of those for white students. This finding is a reflection of these states’ regressive funding tendencies, and the fact that people of color tend to be more concentrated in high-poverty districts. The flip side of this disturbing evidence comes from states such as Massachusetts and New Jersey in which high-poverty districts receive greater support from state and local sources than low-poverty districts.”

They express dismay at the fact that, in some states, Black students receive 10% less funding than White students, but seem relieved that in others Black students receive as much as 18% more funding than White students. Their language seems to imply a sort of anti-White bias on the part of the authors. In any case, if we are trying to explain why, on average, Black life outcomes differ from White life outcomes, and we are talking about national populations, then average spending per pupil across the nation is obviously the correct statistic to look at.

Also relevant is the fact that the Black-White test score gaps are consistent, regardless of schools’ racial makeup [909]. If the test score gap were due to Black schools getting less funding, this should not be the case [909]:



Turning to more specific measures of school quality, racial differences in class size were non-existent by the early 1970s [735]:

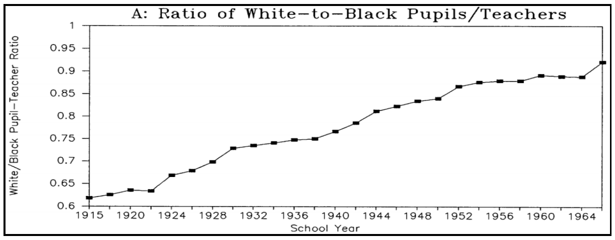
Source 735 - Table 6:

Table 6: Schooling Inputs by Demographic Characteristics 1972-1992

Category	Expenditures/Pupil (1992$) 1972	1982	1992	Pupils/Teacher 1972	1982	1992
By average white and non-white student in the district:						
(1) White	2,856	3,414	4,661	19.32	15.13	13.09
(2) Nonwhite	2,800	3,460	4,796	19.58	14.58	12.52
Ratio (1)/(2)	1.02	0.99	0.97	0.99	1.04	1.05
By median household income in the district:						
1st quartile	2,212	3,040	4,214	19.22	14.24	11.93
2nd quartile	2,388	3,381	4,324	19.24	14.56	12.56
3rd quartile	2,970	3,359	4,686	18.82	15.25	13.20
4th quartile	3,095	3,667	5,047	19.82	15.70	13.53
Ratio (4th)/(1st)	1.40	1.21	1.20	1.03	1.10	1.13
By poverty status:						
(1) Out of poverty	2,881	3,432	4,700	19.34	15.11	13.06
(2) In poverty	2,660	3,331	4,531	19.42	14.81	12.81
Ratio (1)/(2)	1.08	1.03	1.04	1.00	1.02	1.02

In fact, class size differences had been quickly equalizing, even during Southern segregation in the 1940s [736]:

Source 736 - Figure 1-A:



Class size is of course relevant because it has small to moderate effects on school achievement test scores [877, 878, 879, 880, 881, 882, & 883].

Moreover, Blacker schools have more experienced teachers with more formal education and more pay [735]:

Source 735 - Table 12:

Table 12: Characteristics of Newly Hired Teachers by Race and Income Composition of School Schools and Staffing Survey 1993-94

Percent of School Enrollment that is Black:	All	0-10%	10-50%	50-90%	90+%
N	3,643	2,656	696	181	110
Mean Years of Experience	1.48	1.48	1.49	1.49	1.51
Fraction Certified in Primary Teaching Field	91.4	93.8	88.8	87.3	86.8
Fraction with Bachelors Degree or Higher	99.5	99.4	99.7	99.8	99.7
Fraction with Masters Degree or Higher	16.7	15.6	15.1	26.2	28.4
Fraction Teaching Full-Time	86.0	83.6	88.1	94.7	94.2
Fraction Who Say They Would Teach Again	77.3	81.3	73.1	66.3	60.7
Fraction Who Plan to Exit Teaching as Soon as Possible	2.5	1.6	2.2	8.2	9.1
Fraction Who Plan to Exit Teaching at First Opportunity	14.3	13.1	12.9	27.2	21.7
Mean Academic Base Year Salary	23,083	22,741	23,509	23,943	24,209
Percent of School Enrollment Qualified for Free or Reduced-Price Lunch:	**All**	**0-10%**	**10-50%**	**50-90%**	**90+%**
N	3,643	834	1,878	729	202
Mean Years of Experience	-	1.47	1.47	1.49	1.58
Fraction Certified in Primary Teaching Field	-	95.6	93.1	86.7	80.9
Fraction with Bachelors Degree or Higher	-	99.3	99.6	99.5	99.6
Fraction with Masters Degree or Higher	-	22.9	14.3	16.3	14.7
Fraction Teaching Full-Time	-	82.6	84.4	91.1	90.5
Fraction Who Say They Would Teach Again	-	79.9	78.1	74.5	72.5
Fraction Who Plan to Exit Teaching as Soon as Possible	-	1.6	1.5	5.0	3.9
Fraction Who Plan to Exit Teaching at First Opportunity	-	13.1	13.5	17.8	11.2
Mean Academic Base Year Salary	-	24,282	22,331	23,232	24,268

This is not a recent development either; even during segregation in the South, the Black-White teacher pay gap equalized in the 1950's [736]:

Source 736 - Figure 1-C:

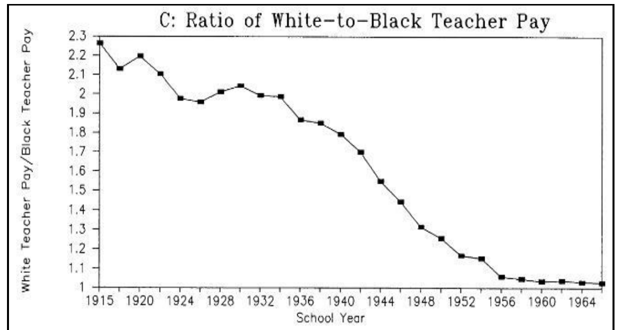


FIGURE I
Relative School Quality in Eighteen Segregated States, 1915–1966

Additionally, back in 1966 at the time of desegregation, a report written at the explicit request of the Supreme Court on thousands of schools and over 650,000 students [1000] found little difference between Black and White schools in terms of physical facilities, formal curricula, and other measurable criteria. It also found that these things did not appreciably align with school achievement differences, and that there was substantially more variation in achievement within schools than between schools.

Given the evidence, Black students are thus advantaged relative to White students in their pre-college education in modern day.

-Affirmative Action:

There is also a significant pro-Black bias in college admissions because of affirmative action. With equal qualifications, Black applicants are roughly 21 times more likely to be admitted into an American college, while Hispanics are 3 times as likely, and Asians are 6% less likely:

#:	School:	Black	Hispanic	Asian
[737](#)	Arizona State (Law)	1115.4	84.95	2.18
[737](#)	University of Nebraska (Law)	442.39	89.63	5.78
[737](#)	University of Arizona Law	250.03	18.15	2.54
[738](#)	University of Virginia (Law)	730.8	1.1	1.86
[738](#)	William and Mary (Law)	167.51	2.47	3.29
[738](#)	University of Maryland (Medical)	20.63	2.51	0.68
[738](#)	George Mason (Law)	1.13	1.09	1.74
[739](#)	William and Mary (Law)	267.0	0.66	0.66
[739](#)	University of Virginia (Undergrad)	106.0	2.81	0.94
[739](#)	North Carolina State (Undergrad)	13.0	1.93	0.64
[740](#)	Berkeley (Law)	121.6	18.2	1.6
[740](#)	UCLA (Undergrad)	5.15	1.92	0.85
[741](#)	University of Michigan	62.79	47.82	0.81
[742](#)	SUNY (Medical)	9.44	4.08	0.76

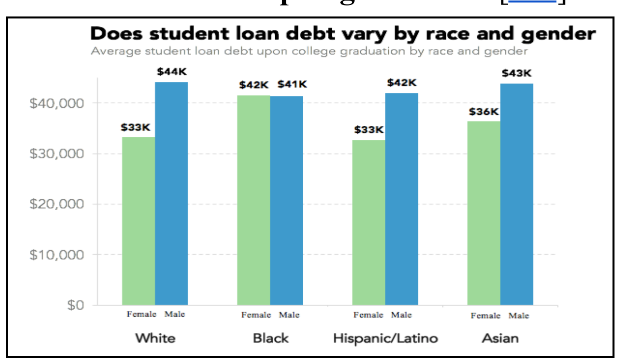
Continued:

#:	School:	Black	Hispanic	Asian
[742](#)	University of Washington (Medical)	4.01	4.86	0.9
[743](#)	Miami University (Undergrad)	7.99	2.16	2.14
[743](#)	Ohio State (Undergrad)	3.33	4.3	1.47
[744](#)	US Naval Academy	4.44	3.32	0.67
[744](#)	US Military Academy	1.94	1.2	0.68
All	All (Mean)	175.51	15.43	1.59
All	All (Median)	20.63	2.81	0.94

In selective colleges, it is estimated that the proportion of students who are White would increase from 66% to 75% if admissions were based solely on test scores [745]. Thinking about it another way, affirmative action gives Blacks a bonus worth the equivalent of 230 extra SAT points during admissions, Hispanics 185 points, legacies 160 points, and Asians -50 points [652].

-Debt / Inheritance:

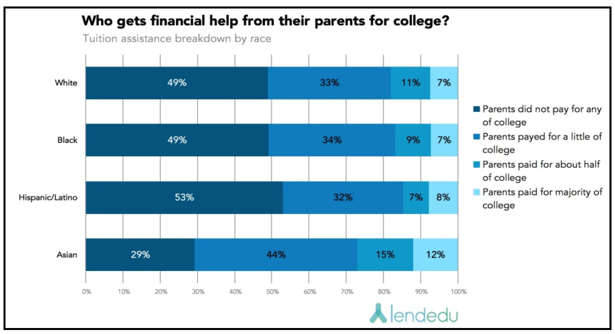
Does college debt disadvantage Blacks? The gap in debt is a function of Whites being more likely to pay it off; there is not really any gap in student loan debt **upon graduation** [746]:



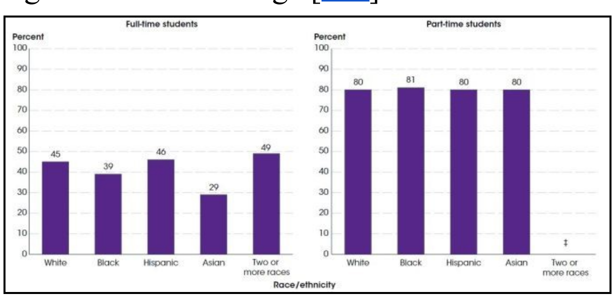
Once minorities get into college, they are given greater access to grants. Specifically, Minority students account for 38% of the student population and 40.4% of grant funding. White students account for 61.8% of all students and 59.3% of grant funding [749]:

Total Grants All Grants Race	Percentage Receiving Grants	Average Grant Amount Received	Total Grant Funding	Number of Grant Recipients	Percentage of Grant Recipients	Percentage of Total Grant Funding	Percentage of Student Population
Total	51.7%	$4,864	$52,646 million	10,822,900	100.0%	100.0%	100.0%
White	48.2%	$5,008	$31,230 million	6,235,700	57.6%	59.3%	61.8%
All Minority Students	57.4%	$4,672	$21,287 million	4,556,300	42.1%	40.4%	38.0%
Black or African-American	63.5%	$4,372	$8,113 million	1,855,800	17.1%	15.4%	14.0%
Hispanic or Latino	58.1%	$4,314	$7,425 million	1,720,900	15.9%	14.1%	14.1%
Asian	43.1%	$6,444	$3,430 million	532,200	4.9%	6.5%	5.9%
American Indian or Alaska Native	60.4%	$4,327	$461 million	106,600	1.0%	0.9%	0.8%
Native Hawaiian or Pacific Islander	49.3%	$4,097	$305 million	74,500	0.7%	0.6%	0.7%
More Than One Race	53.8%	$5,831	$1,553 million	266,300	2.5%	2.9%	2.4%

Black, Hispanic, and White students also have similar chances of their parents paying for a significant proportion of their college education while Asians are more likely than others to have parental aid [746]:

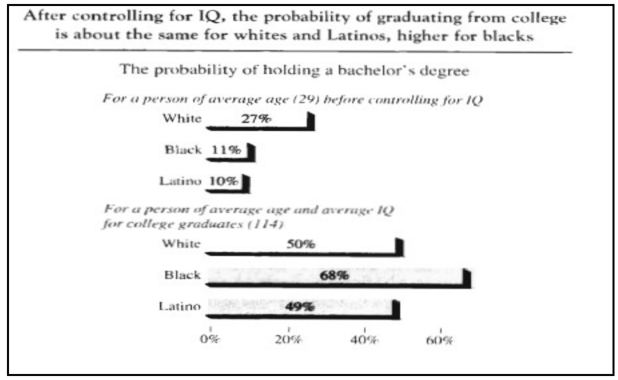


A related narrative is that Blacks can't focus as much on education because their poor financial situation means that they have to work to support themselves during college, but Whites are more likely to hold a job during high school and college [750]:



-Behavior:

So, given all of the financial privileges of Blacks, why are Whites more likely to graduate? Controlling for IQ, Whites and Hispanics are equally likely to graduate from college, and Blacks are more likely to graduate from college [666 - ch. 14 - p.320]:



This makes sense given the well documented pro-Black bias of universities. Whatever the causes of the IQ gap, this completely removes the blame from anything to do with the school system, and puts it onto whatever is the cause of the IQ gap. The case for the majority of the IQ gap being due to genetic differences is strong [see chapter 7], but even ignoring this, we can say even more strongly that the IQ gap cannot be explained by the schooling gaps at all, which means that causality goes from the IQ gap to the schooling gap [see chapter 7]. IQ is an absurdly good predictor of a variety of life outcomes [more here], including grades, test scores, and crime. This is manifested in the Black-White schooling gaps, which are moderated by Black-White differences in these behaviors. This is obviously relevant because one student may complete less years of school than another if they fail courses, drop out, or if they are expelled for poor behavior. The evidence for this moderation is fairly overwhelming:

First, data from the College Board shows there to be a widening Black-White gap in SAT scores [885]:

Source 885 - Figure 3:

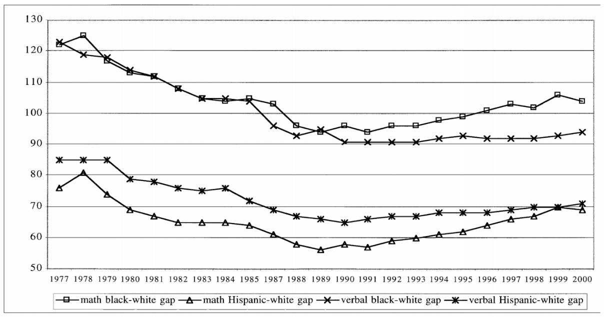
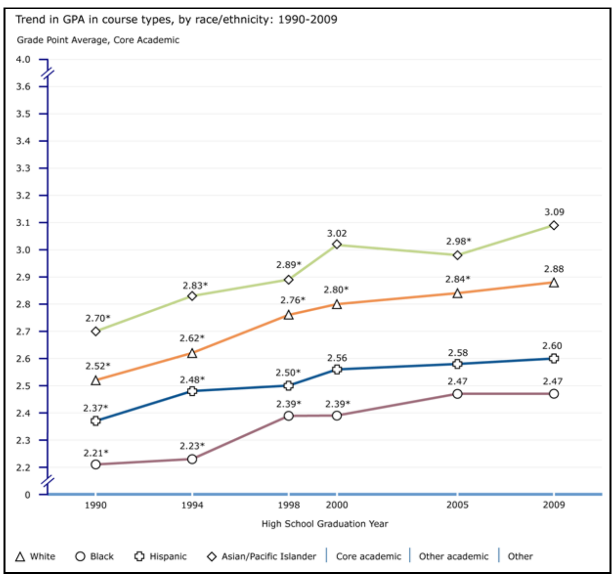


FIGURE 3. *Black-White and Hispanic-White average score gap trends in the 1977–2000 SAT verbal and mathematics.*

Second, government data shows there to be a widening Black-White gap in GPA [884]:



This is partially explained by White students spending more time on homework [886]:

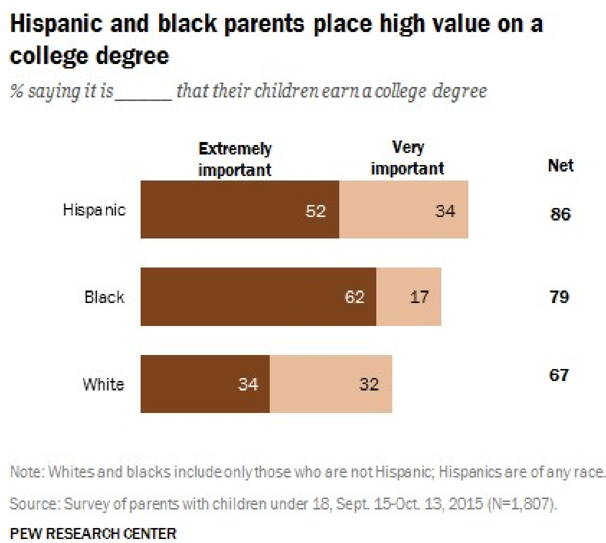
Race/ethnicity	Average hours spent on homework per week by students who did homework outside of school	Percentage distribution of students who do homework outside of school by how frequently they do homework: Less than once per week	1 to 2 days per week	3 to 4 days per week	5 or more days per week	Percentage of students whose parents[1] check that homework is done
Total	**6.8**	**5.4**	**14.8**	**38.0**	**41.9**	**64.6**
White	6.8	4.2	12.9	38.6	44.3	57.2
Black	6.3	‡	20.1	41.0	29.7	83.1
Hispanic	6.4	5.9	17.7	36.6	39.9	75.6
Asian	10.3	#	13.8!	18.5!	67.7	59.0
Native Hawaiian/Pacific Islander	‡	‡	‡	‡	‡	‡
American Indian/Alaska Native	‡	‡	‡	‡	‡	‡
Two or more races	7.1	‡	10.5	32.9	50.5	65.9

The homework time gap exists despite Black and Hispanic parents being more likely than White and Asian parents to check to see that homework is completed [762]:

Race/ethnicity	Average hours spent on homework per week by students who did homework outside of school	Percentage distribution of students who do homework outside of school by how frequently they do homework: Less than once per week	1 to 2 days per week	3 to 4 days per week	5 or more days per week	Percentage of students whose parents[1] check that homework is done
Total	**6.8**	**5.4**	**14.8**	**38.0**	**41.9**	**64.6**
White	6.8	4.2	12.9	38.6	44.3	57.2
Black	6.3	‡	20.1	41.0	29.7	83.1
Hispanic	6.4	5.9	17.7	36.6	39.9	75.6
Asian	10.3	#	13.8!	18.5!	67.7	59.0
Native Hawaiian/Pacific Islander	‡	‡	‡	‡	‡	‡
American Indian/Alaska Native	‡	‡	‡	‡	‡	‡
Two or more races	7.1	‡	10.5	32.9	50.5	65.9

Consistent with this, Black parents place more importance than White parents on their child getting a college degree [761]:

Source 761:



Third, there is a relationship between how non-White a school is and how much violence goes on in the school [735 & 892].

Source 735 - Table 10:

Table 10: Reported Incidents of Serious Violent Criminal Incidents in Public Schools, 1996-97

	% of schools reporting serious violent incidents	Incidents per 1000 students
By minority enrollment of school:		
< 5%	5.8%	0.2
5-19%	10.9%	0.4
20-49%	11.1%	0.5
>50%	14.7%	1.0
By percentage of students participating in the free or reduced-price lunch program:		
<20%	8.6%	0.3
21-34%	11.7%	0.6
35-49%	11.6%	0.5
50-75%	8.9%	0.7
>75%	10.2%	0.8

At the individual level, U.S. Department Of Education data shows Black preschoolers to be 3.6 times as likely as White preschoolers to be suspended [887]. The Black-White gaps in suspension rates also persist as kids grow older and remain after controlling for socioeconomic status [889]. However, this gap does not persist for people with the same previous histories of behavioral problems [890].

Unsurprisingly, Black students are more likely to be bullies than White students, and White students are more likely to be bullied than Black students [891]:

Source 891 - Table 1.1:

Table 1.1: Mean Bullying Involvement, by Demographic Characteristics

	Bullying Outdegree	Proportion with Bullying Outdegree>0	Bullying Indegree	Proportion with Bullying Indegree>0
White	0.56	0.35	0.62	0.33
African-American	0.63	0.36	0.55	0.31
Latino	0.91	0.48	0.86	0.38
Other minorities	0.52	0.31	0.63	0.31
Boys	0.65	0.36	0.59	0.29
Girls	0.70	0.34	0.82	0.36
8th grade	0.85	0.41	0.89	0.37
9th grade	0.61	0.32	0.67	0.31
10th grade	0.56	0.31	0.54	0.29
All	0.60	0.35	0.60	0.32

N=4,567

Family SES, neighborhood SES, physical development, and attachment to friends/parents/school also don't explain racial differences in bullying:

Source 891 - Table 2.3:

Table 2.3: Full Cross-Classified HLM Model of Bullying Outdegree

	β	SE
Intercept	0.50	0.302
Wave 4 bullying	0.28 ***	0.013
Network size	0.00	0.000
Male	0.01	0.035
Black	0.13 ***	0.039
Latino	0.26 **	0.084
Other minority	-0.06	0.070
One parent home	-0.16 **	0.059
Age	-0.02	0.021
Parent attachment	-0.18 *	0.076
School attachment	-0.01	0.014
Sports	0.07 ^	0.036
Service clubs	0.12 **	0.041
DARE	-0.13 ^	0.066
Conventional beliefs	-0.09 **	0.033
Mean bullying of friends	0.05 *	0.024
Family conflict	0.01	0.019
Depression	0.01	0.016
Centrality	0.01 **	0.002
Centrality squared	-0.0001 *	0.000
Bullying indegree	0.05 ***	0.011
Happy with appearance	-0.03	0.025
Friends happy with their appearance	0.01	0.031
Importance of being popular	0.02	0.016
Friends' importance of being popular	0.02	0.039
School Random intercept	0.010	0.006
Neighborhood random intercept	0.007	0.005

N=4,771

^<.05, one-tail test; *p<.05; **p<.01; ***p<.001

With respect to interracial bullying, Black on White bullying is 64% more common than White on Black bullying:

Source 891 - Table 1.5:

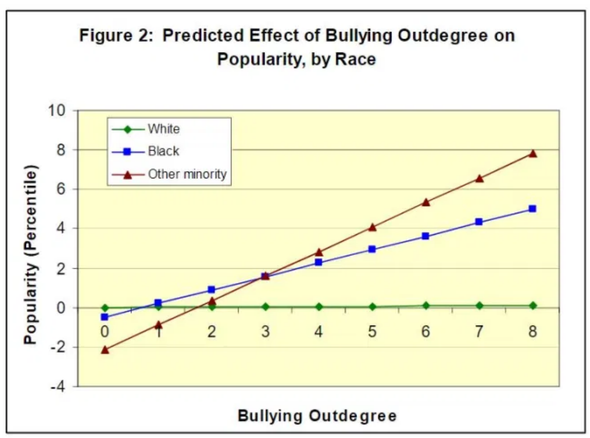
Table 1.5: Bullying Rate Per Thousand Dyads, By Race and Gender

Dyad Type (Sender-Receiver)	Mean	Frequency
Black-Black	3.60	456
Black-Latino	1.57	18
Black-Other	5.86	79
Black-White	2.26	201
Latino-Black	0.87	10
Latino-Latino	22.48	84
Latino-Other	4.64	7
Latino-White	2.34	42
Other-Black	4.30	58
Other-Latino	0.00	0
Other-Other	6.92	23
Other-White	3.95	82
White-Black	1.37	122
White-Latino	2.06	37
White-Other	3.94	76
White-White	4.77	844
Female-Female	4.60	813
Female-Male	2.00	381
Male-Female	2.75	523
Male-Male	3.48	713
Overall	3.19	2430

N=761,558

In part though, the Black-White difference in bullying may partially arise from Black culture being more likely to socially reward bullying:

Source 891 - Table 2:

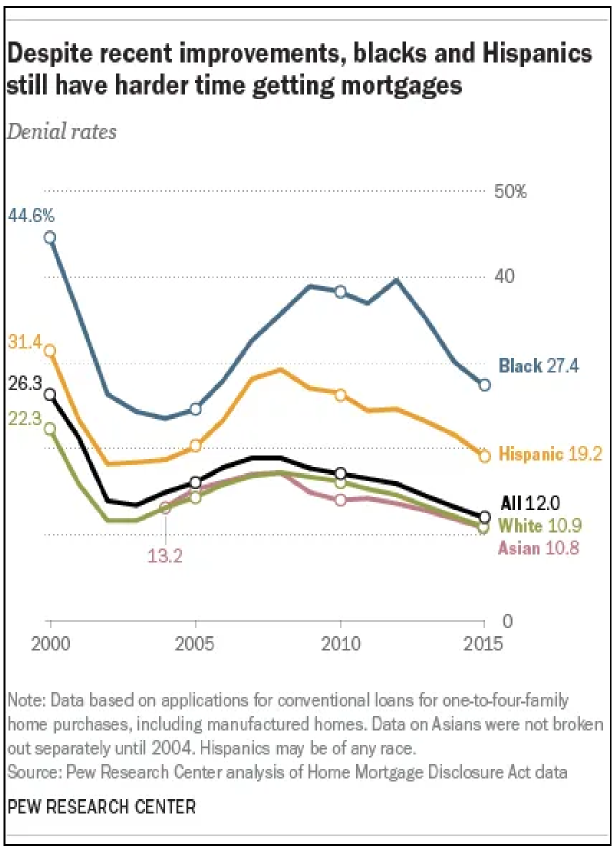


Perhaps most dramatically, Farris finds [891]:

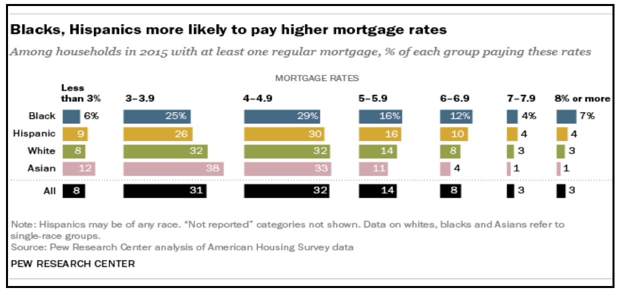
> "For every one percentage point increase in the percent minority in the school, the likelihood of suicide increases by one percent."

Redlining & Bias In Lending:

Racial differences in the ability to acquire a loan are sometimes pointed to as evidence of White privilege or anti-Black bias. These differences are said to lead to racial disparities in home ownership rates, which in turn have a variety of long-term economic and social consequences. Sometimes bias among landlords is also brought up, but White landlords do not 'discriminate against Blacks' in pricing more than Black landlords do [958]. Pew Research Center data [1136] shows that Black people are indeed more likely to be denied for a mortgage loan. However, even among Blacks, the rate of denial is only 27%:



For interest rates, it is true that Black people are more than twice as likely as Whites to get a mortgage interest rate of 8% or more. But this is very rare even among Black mortgage holders. The average interest rate seems to be similar among Whites, Hispanics, and Blacks, though possibly significantly lower for Asians:



The central question to be asked in order to ascertain the existence of racial bias is "why?" Black homes have lower saving rates than White homes even after controlling for differences in income, age, family size, education, and marital status [893]. Thus, if lenders have additional information beyond these variables that lead them to predict the differences in payment ability, it cannot be said that lenders are "racists" who would rather lose money than loan to Black signers. Also worth noting is that [888]:

> "Blacks and Hispanics spend roughly 30 percent more on visible expenditures (cars, clothing, jewelry, and personal care items) than otherwise similar Whites."

Blacks also seem to be lower in self control in general, being less willing to deter short term gains for larger long term gains [more here].

-Credit Scores:

Some point out [895] that racial differences in loan acceptance persist even after adjusting for credit score differences. This is true [894]. It is also true that credit scores don't mean the same things for Blacks and Whites [896]:

> "Consistently, across all three credit scores and all five performance measures, blacks… show consistently higher incidences of bad performance than would be predicted by the credit scores."

On the aggregate, credit scores don't work equally well for Blacks and Whites, but among those with high credit scores, there isn't much of a difference [896]. Consistent with this, there is no racial bias in loan approval rates among those with good credit scores, but a significant "bias" in favor of Whites among those with bad credit scores [897]. Similarly, Black borrowers have a tougher time getting loans, but this is only true among those who don't have mortgage insurance [898].

-Default Rates:

Loans taken on by Black people are more likely to end in default. This result is robust to controlling for the size and type of loan, and characteristics of the borrower such as their age, income, and liquid assets value [899]. If Black people are discriminated against in the loan market, we would expect that Blacks must be more profitable than Whites in order to obtain the same loan, and so they must ensure a lower risk of default than the White default rate in order to get loans. These results show that this is not true and so this is evidence against racial bias.

Perhaps high Black default rates are to be expected because Blacks are charged greater interest rates, but this explanation is not compelling because there is a miniscule gap in interest rates between races once obvious confounds are controlled for. Analyzing data from the U.S. Survey of Consumer Finances from the years 2001, 2004, and 2006 [900], controlling for measures of consumer behavior and debt risk reduces the Black-White average interest rate gap to just 0.29%. This remaining gap is far too small to explain the gap in default rates, and it may itself be explained by variables that are yet to be measured anyways.

-Pay Schedule:

Similarly, in a data set consisting of all FHA-insured mortgages that originated in 2014 and 2015, the Black-White interest gap was 0.03% and the Hispanic-White gap was 0.015% after controlling for lender effects, credit score, and income [901]. The paper included data on discount points, and this revealed a racial difference in favor of non-Whites. Combining this data into a single model, no racial bias in a borrower's expected pay schedule was found. More importantly, it is shown that the expected revenue generated by a loan does not significantly differ by the race of the borrower.

This evidence is hard to reconcile with racial bias. That no bias exists is directly suggested by the fact that races experience the same expected pay schedules once other differences are held constant. The fact that the expected revenue of loans does not differ by race strongly suggests that the differences in the terms of loans given to Blacks and Whites reflect lenders accurately forecasting the terms which will maximize profit within each race of borrowers. It is hard to see how this result could come about if people were acting on the basis of racial animus rather than economic rationality.

-Black-Owned Banks:

This study [902] of several thousand banks finds that Black-owned banks "discriminated" far more harshly against Blacks than did White-owned banks, suggesting that Blacks are more likely to act on economic rationality while Whites try to coddle Blacks. Specifically, at a White owned bank, a Black person was found to have a 78% higher chance of rejection for a loan compared to a White person. At a Black-owned bank, this figure rises to 179%:

Source [902](#) - Table 2:

Table 2. Home purchase application outcomes from the HMDA Loan Application Registers for 1992–1993 by race and bank ownership status (white-owned v. black-owned). The percentage for a particular outcome within a bank type reflects that portion of the row total. The number of observations for the outcome type appears in parentheses beneath the percentages

	White-Owned Banks		Black-Owned Banks	
	Acceptances	Rejections	Acceptances	Rejections
White applicants	90.59% (1984)	9.41% (206)	86.22% (169)	13.78% (27)
Black applicants	83.26% (179)	16.74% (36)	61.56% (458)	38.44% (286)
Total applicants	89.94% (2163)	10.06% (242)	66.70% (627)	33.30% (313)
Disparity ratio[a]		1.78		2.79

[a] The disparity ratio, as referred to in ''Mortgage Gap . . .'' (1992), places the rejection rate for black applicants in the numerator and the rejection rate for white applicants in the denominator.

Thus, racial differences in the riskiness of loans seem to account for why Blacks have a harder time getting loans than White people do, and why their interest rates tend to be slightly higher.

-Redlining:

A narrative related to racial bias in lending concerns the practice of redlining. Essentially, the idea is that in the 1930s, the US government created maps demarcating certain neighborhoods as high risk for investment. One of the variables they utilized when estimating an area's degree of risk was that area's racial composition. Lenders then became less likely to give out loans to people in these communities, and, through public housing and zoning laws, Black people were moved into these same communities making them Blacker than they initially were. Thus, it is said that Black were at a disadvantage in the loan markets because of the neighborhoods they lived in.

Importantly, this bias only impacts race indirectly. The discrimination is directed at neighborhoods and so should apply equally to people of all races who live in these majority Black areas. Accordingly, where investigated, multiple papers have found that the probability of people getting a loan did not relate to the racial composition of their neighborhood once economically relevant confounding variables are controlled for [904, 905, 906, 907, & 908]. The idea that redlining increased racial inequality also seems unlikely in light of the fact that the Black-White home ownership gap today is similar to what it was in the 1920s before redlining began [903]:

Source [903](#) - Figure 1:

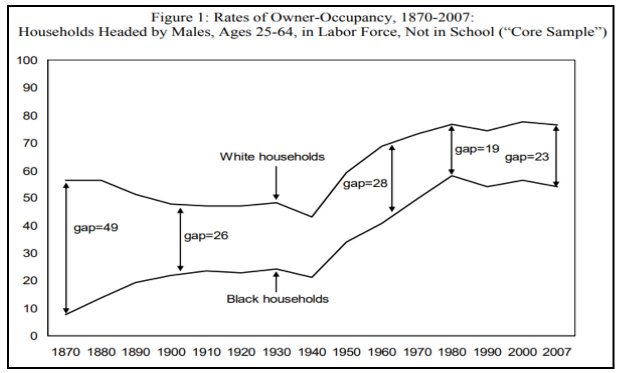


Figure 1: Rates of Owner-Occupancy, 1870-2007:
Households Headed by Males, Ages 25-64, in Labor Force, Not in School ("Core Sample")

Hiring Discrimination:

The strongest case that can be made for the existence of any White privilege is in hiring discrimination. Viral is the story of the Black woman who changed her name to sound more White and started to receive **ten** times the amount of callbacks that she did before [910]. Does this actually happen? Yes, two resumes that are identical aside from one having a Black-sounding name get different callback rates, but the real effect is much more modest than suggested by this outlier story [607]. Moreover, the supplementary materials show the meta-analytic effect size to be inflated by publication bias [606] (warning: direct download link!). Not as exciting.

The question, as always, is **why** this happens. There is evidence that the disparity is due to the socioeconomic connotations of the names rather than Black connotations [962].

However, even if it were the case that there is variance in callback rates that can only be attributed to race and nothing else, it must be recognized that under certain conditions, an employer that selects for ability would be rational if they held Blacks to higher resume standards. Given the following three facts, Whites and Blacks with equal resumes are still different in ability:

1. Qualifications require a threshold of ability.
2. Blacks and Whites differ in distribution of ability.
3. This is enough to create a sizable gap among equally qualified candidates, but Affirmative Action exacerbates this gap.

These points will be argued shortly, but first, there is a good potential objection that needs to be dealt with. Theoretically, if these things are true, a racist employer could discriminate because they dislike Blacks rather than because he is selecting for ability. Such a racist employer could be efficient by accident. That this is not the case is shown by the fact that when criminal records are put on resumes, Blacks and Whites with equal resumes have equal callback rates [912]:

Race	No Crime	Any Crime	Property Crime	Drug Crime
White	14%	8.3%	7.7%	8.9%
Black	13.1%	8.6%	9.1%	8.1%

This is reminiscent of the following famous/infamous cartoon:

It is based on source 913, a criminally (2990) over-cited paper, which does indeed find the advertised result, but does not control for a single resume characteristic. However, with equal credentials and criminal record, the callback gap disappears [912]. That is, employers engage in statistical discrimination, not racial discrimination. The word "discrimination" in "statistical discrimination" does not make statistical discrimination automatically evil either; an employer is no more morally obligated to hire an unskilled Black candidate than a smart, attractive woman is to date a short, fat, weird, highschool dropout.

-Statistical Discrimination Is Rational:

Ideally, employees will be hired based on their ability to perform in their job. On average, Blacks score .35 standard deviations below Whites on measures of job performance [914]:

Source 914 - Table 2:

Table 2

Black–White Differences in Job Performance

Measure	*d*	*K*	N_{Total}	N_{White}	N_{Black}	90% CI	$d_{corrected}$	PVA (%)
Ratings								
Overall ratings	.27	37	84,295	62,073	22,222	25, .29	.35	35
Overall with no military	.31	36	46,183	35,023	11,160	.28, .33	.40	43
Overall no large *N*s	.26	33	9,141	6,693	2,448	.23, .30	.34	76
Quality & quantity measures								
Quality ratings	.21	15	3,613	2,387	1,226	.14, .27	.25	89
Quantity ratings	.21	8	1,268	925	343	.03, .40	.26	69
Other measures								
Job knowledge	.48	12	2,460	1,577	883	.36, .58	54	30
Work samples	.52	10	3,651	2,260	1,391	.39, .66	.59	18
Absenteeism	.19	11	2,376	1630	746	.11, .26	.22	100
On-the-job training	.14	2	132	75	56	−.03, .31	.18	100
Promotion	.31	7	1,404	1,168	236	.20, .42	.38	100

Note. PVA = the percentage of variance accounted for by sampling error.

Worries of racial bias in job performance measures should also be alleviated by the finding that racial differences are larger on the more objective measures:

Source 914 - Table 3:

Table 3
Comparison of Objective and Subjective Measures of Job Performance

Measure	d	K	N_{Total}	N_{White}	N_{Black}	90% CI	$d_{corrected}$	PVA (%)
Quality measures								
Objective	.24	8	2,538	1,632	906	.17, .30	.27	100
Subjective	.20	10	1,811	1,262	549	.12, .28	.26	100
Subjective, no part time	.14	9	1,580	1,063	517	.10, .20	.18	100
Quantity measures								
Objective	.32	3	774	613	161	−.06, .72	.35	84
Subjective	.09	5	494	312	182	−.06, .24	.12	100
Job knowledge								
Objective	.55	10	2,027	1,315	712	.42, .68	.61	34
Subjective	.15	4	1,231	793	438	.08, .23	.19	100
Absenteeism								
Objective	.23	8	1,413	1,005	408	.12, .32	.26	90
Subjective	.13	4	642	377	275	.09, .17	.17	100

Note. Objective measures of performance were corrected for attenuation using the value of .8, whereas subjective measures were corrected by using the value of .6. PVA = the percentage of variance accounted for by sampling error.

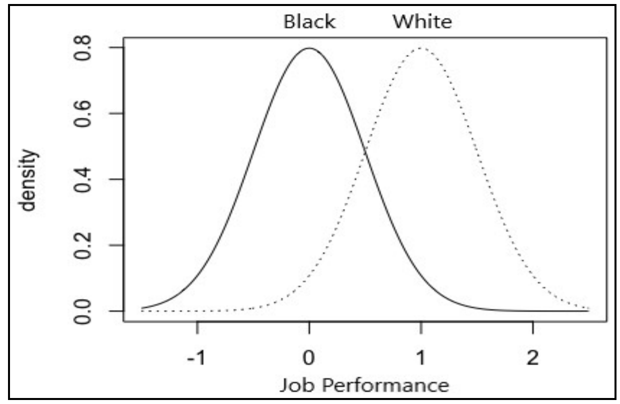
In order for callback studies to be valid measures of racial discrimination rather than statistical discrimination, these differences in job performance must disappear once we control for the sorts of qualifications one finds on a resume. This is unlikely to be true because even assuming completely additive validity (This is guaranteed to be at least partially false because of mediation with IQ), variables which can be found on a resume such as education, job experience, age, and reference checks explain less than 22.1% of variance in job performance [426]:

Source 426 - Table 1:

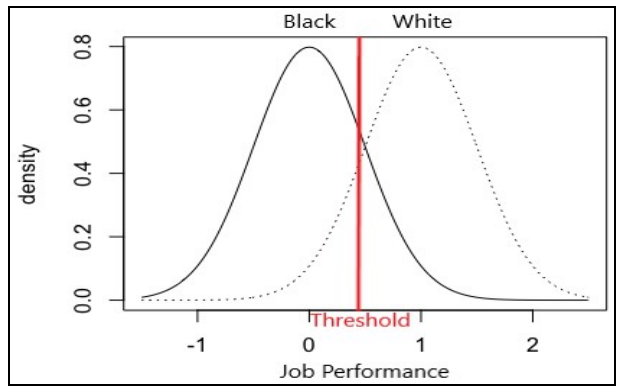
Table 1
Predictive Validity for Overall Job Performance of General Mental Ability (GMA) Scores Combined With a Second Predictor Using (Standardized) Multiple Regression

Personnel measures	Validity (r)	Multiple R	Gain in validity from adding supplement	% increase in validity	Standardized regression weights: GMA	Standardized regression weights: Supplement
GMA tests[a]	.51					
Work sample tests[b]	.54	.63	.12	24%	.36	.41
Integrity tests[c]	.41	.65	.14	27%	.51	.41
Conscientiousness tests[d]	.31	.60	.09	18%	.51	.31
Employment interviews (structured)[e]	.51	.63	.12	24%	.39	.39
Employment interviews (unstructured)[f]	.38	.55	.04	8%	.43	.22
Job knowledge tests[g]	.48	.58	.07	14%	.36	.31
Job tryout procedure[h]	.44	.58	.07	14%	.40	.20
Peer ratings[i]	.49	.58	.07	14%	.35	.31
T & E behavioral consistency method[j]	.45	.58	.07	14%	.39	.31
Reference checks[k]	.26	.57	.06	12%	.51	.26
Job experience (years)[l]	.18	.54	.03	6%	.51	.18
Biographical data measures[m]	.35	.52	.01	2%	.45	.13
Assessment centers[n]	.37	.53	.02	4%	.43	.15
T & E point method[o]	.11	.52	.01	2%	.39	.29
Years of education[p]	.10	.52	.01	2%	.51	.10
Interests[q]	.10	.52	.01	2%	.51	.10
Graphology[r]	.02	.51	.00	0%	.51	.02
Age[s]	−.01	.51	.00	0%	.51	−.01

Moreover, this is actually extremely unlikely to be true for a simple statistical reason: Suppose that the distributions of job performance among Blacks and Whites consist of two overlapping normal distributions, which looks like this:

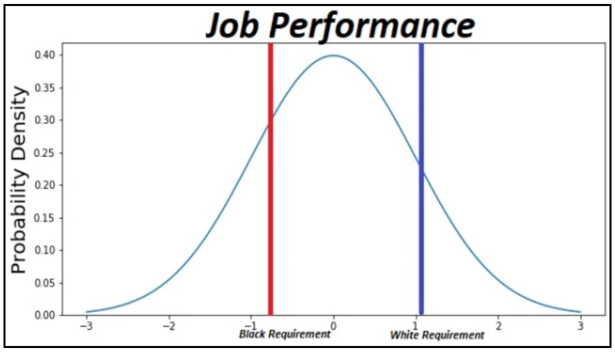


Now, suppose that a given qualification requires a threshold of ability to obtain:



As is hopefully self-evident from the previous example, there is no possible threshold for ability which would cause the average of the Blacks who are above the threshold to be equal to the average of the Whites who are above the threshold. Since variables typically contained within a resume are not direct measures of job performance, but thresholds which can be obtained by anybody with ability greater than or equal to required by the threshold, it is almost certainly the case that a group of Whites with, on paper, equal qualifications to a group of Blacks, would outperform that group of Blacks on the job. Given the well established 1 standard deviation Black-White IQ gap [876], there is no reason for an employer with two equal resumes to assume that a Black and a White applicant are equal in abilities unrecorded by the resumes.

The rationality of statistical discrimination only becomes more extreme when Blacks are given a lower threshold for qualifications than the White threshold:



This is exactly what Affirmative Action in education does [more here]. Being Black is worth the equivalent of 230 SAT points in college admissions [652]. In other words, since the SAT has a standard deviation of 210, and since IQ has a standard deviation of 15, a White applicant's university degree is, on average, worth 16.43 more IQ points on a job resume than a Black applicant's is.

Turning from education to employment, since 1969 with the institution of the Philadelphia Plan under Richard Nixon, all government workers, and many government contractors, have also been required to engage in affirmative action programs aimed at increasing the prevalence of minorities in their work forces.

Some may still object to statistical discrimination, saying that it's still discrimination. Well, think of it this way: A Black Harvard graduate has to send out a few more job applications to get the same job as a White Harvard graduate, but would you rather be the Black Harvard graduate, or the White graduate from Georgia Tech?

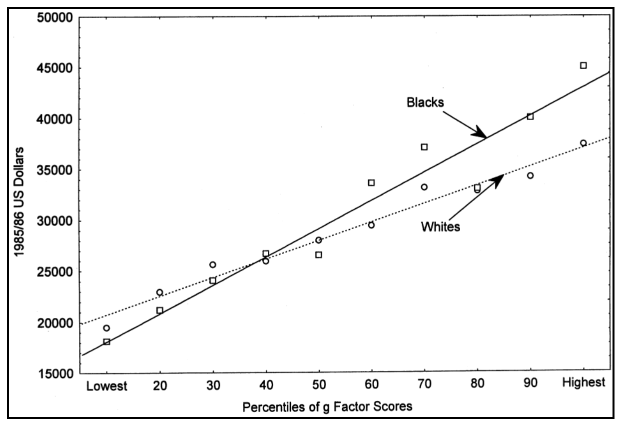
What Of The Gaps?

With mechanism after mechanism of discrimination out of the picture, what are we left with? Why do these gaps still remain?

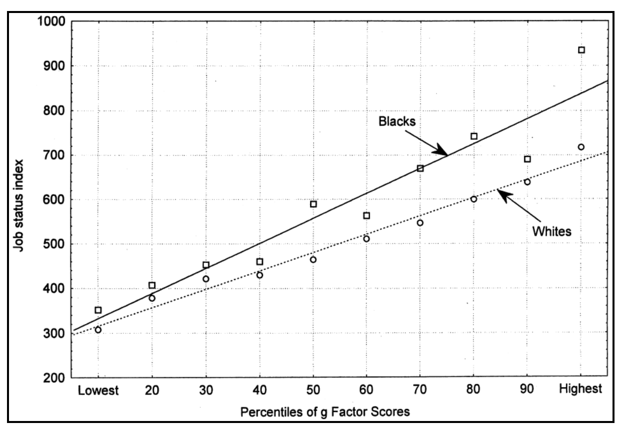
-IQ:

IQ is an absurdly good predictor of life success [more here], and there is a well established 1 standard deviation Black-White IQ gap [876], which we have reason to believe is mostly genetic in origin [see chapter 7]. If the IQ gap were eliminated, Whites would have lower status jobs, and would make less money [703]:

Source 703 - Figure 1:



Source 703 - Figure 2:



Inequalities also reverse, equalize, or reduce substantially in many other domains [666], and these sorts of results have been replicated many times over [706, 704, & 705].

-Self Control:

IQ is negatively, though weakly associated with low self control [871], and this association is genetically mediated [1115]. However, this cannot fully explain the heritability of self control because self control is about 50% heritable [1117, 1118, & 1119]. Self control is important because it has power to predict life success which is independent of IQ and socioeconomic status. IQ is of course important to control for because of its predictive power and its collinearity with self control and success. Socioeconomic status is also an important control variable to include because people under emergency financial pressures may be influenced by said pressures to act in a way which is out of line with their true time preference.

Source 1110:

This paper looked at how well self-control measured in childhood (under the age of 10), based on self and peer reported behavior, predicted life outcomes at age 32 in comparison to childhood IQ and parental socio-economic status in a nationally representative sample. Higher childhood self-control was found to predict better health, more wealth, less criminality, and a lower chance of being a single parent in adulthood even controlling for IQ and parental SES. Particularly interesting is the fact that IQ was not predictive of criminality, drug abuse, or single parenthood when parental SES and self-control were controlled for. However, consistent with the past literature, the paper found IQ to be the best predictor of wealth and adult SES.

Source 1120:

Looking at how childhood self-control, IQ, and class predicted adult unemployment in a sample of 16,780 Brits, this paper finds holding the other two variables constant, high self control was related to lower unemployment while social class was not related to unemployment when the other two variables were held constant.

Source 1121:

This paper finds that self control is a better predictor of GPA than IQ and that self control was related to more time being spent on homework while IQ was related to less time being spent on homework.

Source 1123:

This meta-analysis confirms a correlation between self control and various life outcomes such as love, happiness, getting good grades, speeding, commitment in a relationship and lifetime delinquency, but did not assess the mediating roles of IQ or socioeconomic status.

Black-White Differences In Self Control:

Self control is of course relevant to Black-White inequalities in the things that self control is predictive of because there is evidence that Blacks have lower self control than Whites:

Source 1124:

This paper took advantage of a natural semi-experiment which came about due to the military. In the mid 1990s, the U.S. Government offered sufficiently experienced military personnel two options when they retired: they could take a large lump sum of money now or agree to get a yearly payment from the military for the rest of their lives which, over time, would add up to far more than the lump sum. Data was found on the choices of 66,000 individuals, and Blacks were 15% more likely than non-Blacks to take the lump-sum.

Source 893:

In this paper, the Black homes in a sample of 25,820 households were found to have lower savings rates than White homes even controlling for differences in income, age, family size, education, and marital status.

Source 888:

> "Blacks and Hispanics spend roughly 30 percent more on visible expenditures (cars, clothing, jewelry, and personal care items) than otherwise similar Whites."

Source 1122:

This paper utilized a sample of 5,291 university students from 45 countries and gave participants a chance to choose an immediate monetary reward or a larger long term reward; figure 3 shows the proportion of people from different regions that chose the larger and less immediate reward:

Source 1122 - Figure 3:

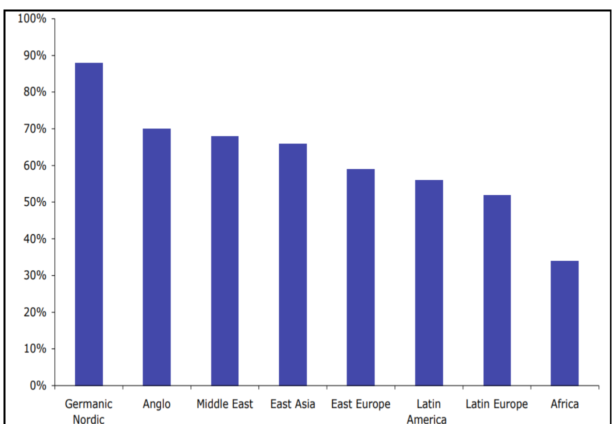


Figure 3: The percentage of choosing to wait grouped by cultural origin

Source 1125:

This paper looked at a sample of 317 individuals with gambling problems and found that White gambling addicts had more self-control than Black gambling addicts even after controlling for education, drug problems, and income.

Source 1126:

The authors of this paper describe their experiment as follows:

> "In our experiment, subjects are asked, orally and in writing, to make twenty decisions in total. For each decision, subjects are asked if they would prefer $49 one month from now or $49+$X seven months from now. The amount of money, $X, is strictly positive and increases over the twenty decisions."

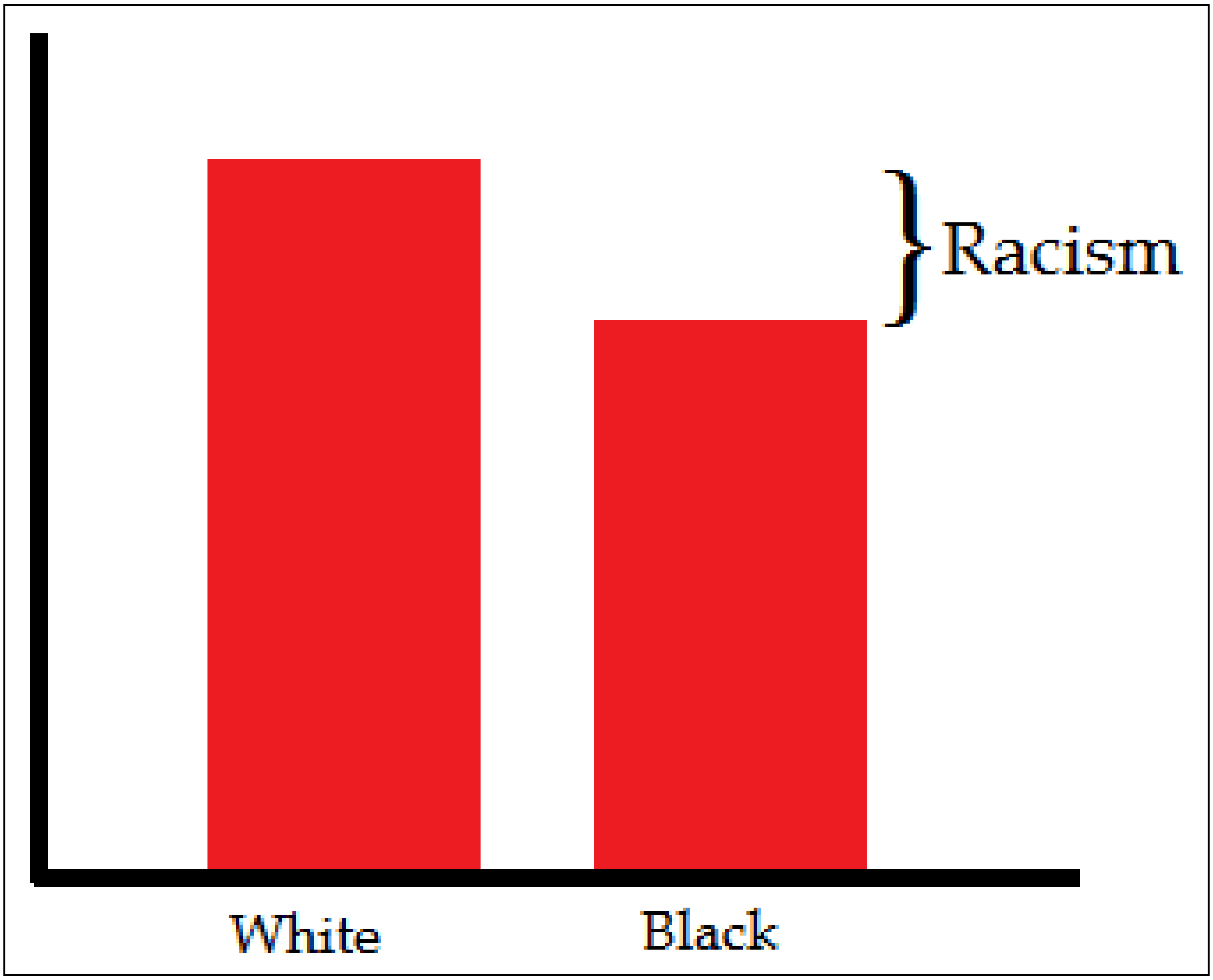
Using this design in a sample consisting of 82% of the student population of 4 middle schools in a poor Georgia school district, the paper was able to measure at what point people began to prefer the later reward and, thus, the strength of their preference for immediate gratification. Blacks were found to have significantly less self-control than Whites.

Source 1127:

This paper looked at a sample of 100 4th grade school children and found that Blacks had lower self control than Whites even after controlling for socio-economic status.

While the within-group heritability of self-control does not necessarily guarantee an above zero between-group heritability of self control, a handful of gene variants which are related to impulsive behavior have also been found to be less common among Blacks than among Whites [1111].

Source Epic - Figure 13.50:



Racism deboonked.

5. Racism:

Navigation:

Previous Chapter   Table Of Contents   Next Chapter

Summary

The terms, "racism" and "racist" are meaningless, dishonest, slander terms used to attack Whiteness. Academia heavily leans to the left, the left is anti-White, and academic publication bias is measurably opposed to hereditarianism. Stereotype threat effects (the idea being that beliefs in group differences cause group differences to become a self-fulfilling prophecy) do not exist. Stereotype threat is pushed by the anti-White left because, if true, it would mean that the mere investigation of group differences is harmful to the groups in question. It is not harmful to investigate group differences, so don't worry about whether or not something is "racist". Instead, worry about whether or not something is correct (not politically correct, but actually correct).

The Word “Racism”:

Circular reasoning:

Is it racist to argue against using the word “racist”? Chad and Stacy are arguing. “Let me dismantle the concept of racism,” Chad begins to explain. “No, only racists question the concept of racism” Stacy dismisses. “But ‘racist’ is the very thing in question!” protests Chad. To convince Stacy to examine the concept of racism, Chad would have to convince Stacy that the “racism” of the topic is not justification to be uncritical of the topic. To do this, Chad needs to convince Stacy to examine the concept of racism; the circle is now complete.

Some would say that the arguers of these arguments are not to be trusted because they are vested interests because they themselves are often called racists. It’s like somebody being hit on the head with a hammer who is not allowed to object because by being hit on the head, he is now a vested interest and thus should not be trusted. If for the sake of argument, the term, “racist”, is a meaningless smear term, who would be more aware of such reality than the people who get hit in the head with a hammer? Moreover, the people using the term, “racist,” would also be vested interests in the argument because if it were accepted that the term, “racist,” was a smear term, such people would be exposed as having been dishonest character assassins since they used the term to smear people. The term racism is used to shut down honest dialogue, and this makes sense when considering what the term really means.

Descriptive Power:

Language is an intersubjective phenomenon which attempts to convey meaning between people. Without language, thoughts go unlabeled and are eventually forgotten, which precludes them from precise use. Take colour as an example, there is no point where one colour ends and another begins, they all gradually blend into each other. However, having the word green and blue creates a distinction in the minds of those who use the words. In cultures which have one word for the colours green and blue put together, they see them as the same broad colour; the imprecision in language leads to imprecision in thought about colour. Sloppy language leads to sloppy thought.

Let someone say “Bob is racist”. What do they mean by that? When the audience asks them why Bob is racist, the audience is asking for more information than how “racist” Bob is and the evidence for that specific amount of “racism”. The accuser may respond that “Bob is racist because Bob is antisemitic”. Nobody knows what the first sentence actually meant because it doesn’t actually mean anything. The first sentence and the term “racist” in the second sentence were only included to give the audience negative perceptions of Bob. If you have to say that somebody is racist because they harbor racial hatred, what information did the term “racist” convey? At least in the middle ages when you were called a heretic, it was widely known that the accusation was of an affront to God.

Often, the term racist is used to give the audience a perception that the defendant is a race-hater without the accuser actually having to take on the burden of making the accusation. Instead, the accuser just piggybacks onto the vagueness of the term. If the defendant tries to ask why he would have Black friends if he hates Black people, the accuser can say that he never accused the defendant of hating Black people, and ask why the defendant preemptively feels the need to defend against such an accusation. The accuser frees himself from having to actually argue for anything and just lets the audience draw whatever preconceived implications they have of the term "racist". The only preconceived implication of the term "racist" that anybody should have is that the accuser is dishonest and that the accusation is an attack, because that's what it is. It is an incredibly powerful attack too, source 463 found that Americans dehumanize "racists" more than they dehumanize groups which are traditionally seen as being dehumanized.

Bob explains his personal definition of pedophile to Bill, that pedophiles like children in a non-sexual, non-predatory way. So, when Bob tells an audience of 12,000 that Bill is a pedophile, Bill should feel safe knowing what Bob's definition of pedophile is, right? If Bill asks Bob what Bob's definition of pedophile is in front of the audience, the audience will just immediately think that Bill is trying to defend pedophillia. Words matter.

One important problem with the term "racism" is a rather obvious one, conflation. James Watson, discoverer of the double helix in DNA, is called a racist for his claims about race and intelligence in order to conflate him with Hitler, who is also called a racist. These are two very different positions. Watson isn't into policy while Hitler was a dictator who was the sole determinant of policy in multiple countries. Watson was focused on descriptive claims about reality while proponents of eugenics are focused on prescriptive ideals or actions that they want carried out.

Eugenics falling out of favor because of perceived ties to Hitler is also ironic because the German speaking and non-German speaking worlds had increasingly divergent schools of thought following the first world war. Hans Eysenck, a psychologist who grew up in The Third Reich, recalls that psychometrically valid intelligence tests were banned under The Third Reich. On page 16 of his 1979 book The Structure & Measurement Of Intelligence, [100] he wrote: "Stalin, as already noted, banned intelligence testing for being 'bourgeois', and Hitler did the same because they were 'Jewish'.“ We, however, don't even have the honesty to declare our target. We reject IQ tests because they're "racist". Source 111, on page 21, notes that those killed by Hiter's eugenics for severe retardation were a small minority and that the killers showed little interest in intelligence testing.

To showcase another incongruity between actual Nazi beliefs and modern day race narrative, as part of the German *Lebensborn* program, 250,000 Jewish children were kidnapped and subjected to propaganda in an attempt to cleanse them of their Jewish heritage [123].

What Is Racist?

Here is a conundrum for "anti-racists", if some racist beliefs are correct, then incorrect things would need to be believed in order to not be racist. On the other hand, if proclaimed that all racist beliefs are incorrect, then suddenly everybody needs to constantly reassess their definitions of racism as new evidence comes to light.

Consider which boxes are racist in the "Is it racist?" table:

Is it racist?

BELIEFS ABOUT WHITES	correct generalization	incorrect generalization
positive generalization		
negative generalization		
BELIEFS ABOUT NON-WHITES	correct generalization	incorrect generalization
positive generalization		
negative generalization		

If all generalizations are racist, whether true or false, whether positive or negative, or whether about Whites or about nonWhites, then recognize how the strictness of this definition contrasts with more relaxed definitions which others may hold. Consider that "racism" being brought up invokes the implications of any definition that anybody in the audience happens to hold. This is a problem because of the wide diversity in definitions recorded in source 600:

Source 600 - Table 13.4:

According to contemporary commentaries in general society and in the social sciences (including applied psychology), you may be accused of "racism" or "being a racist" if:

- You are a human being
- You are White
- Your political interests align with conservative or patriotic principles, policies, or values
- You agree with a highly disliked person (on an issue that has little or nothing to do with race) who has been judged to be "racist"
- You notice social problems that involve racial/ethnic groups and desire to discuss them openly

- You hold views that are different from the mainstream media on multiculturalism, affirmative action, crime, and educational underachievement

Source 600 - Table 13.4 (continued):

- You criticize or disapprove of the negative behavior of individuals from racial/ethnic minority groups
- You fail to combat racism
- You believe that artistic/scientific contributions from Western societies and cultures are superior than contributions from non-Western societies and cultures
- You believe that race is a biologically useful concept for classification of human subgroups
- You attempt to treat persons in a "colour-blind" manner
- You conduct research on racial differences
- You believe that no average differences between racial groups exist beyond superficial differences in skin colour
- You believe that average differences between racial groups exist beyond superficial differences in skin colour
- You believe in a genetic basis for variation in human traits
- You believe in a biological or genetic basis for why certain groups excel on average in certain areas
- You believe that racially/ethnically diverse societies promote greater problems than racially/ethnically homogeneous societies
- You believe that all subgroups must be held to the same standards (e.g., in employment, education, civic behavior)
- You believe that all subgroups should not be held to the same standards (e.g., in employment, education, civic behavior)

Which definition is correct? Is this book racist? My definition of racist is a professional racecar driver, so this book isn't a racist.

Consider also whether or not to fill out the boxes regarding beliefs about Whites differently than the boxes regarding beliefs about non-Whites. Which two boxes out of the entire chart would be most agreed upon by "anti-racists"? The most agreed upon boxes among those who consider themselves anti-racist are probably that:

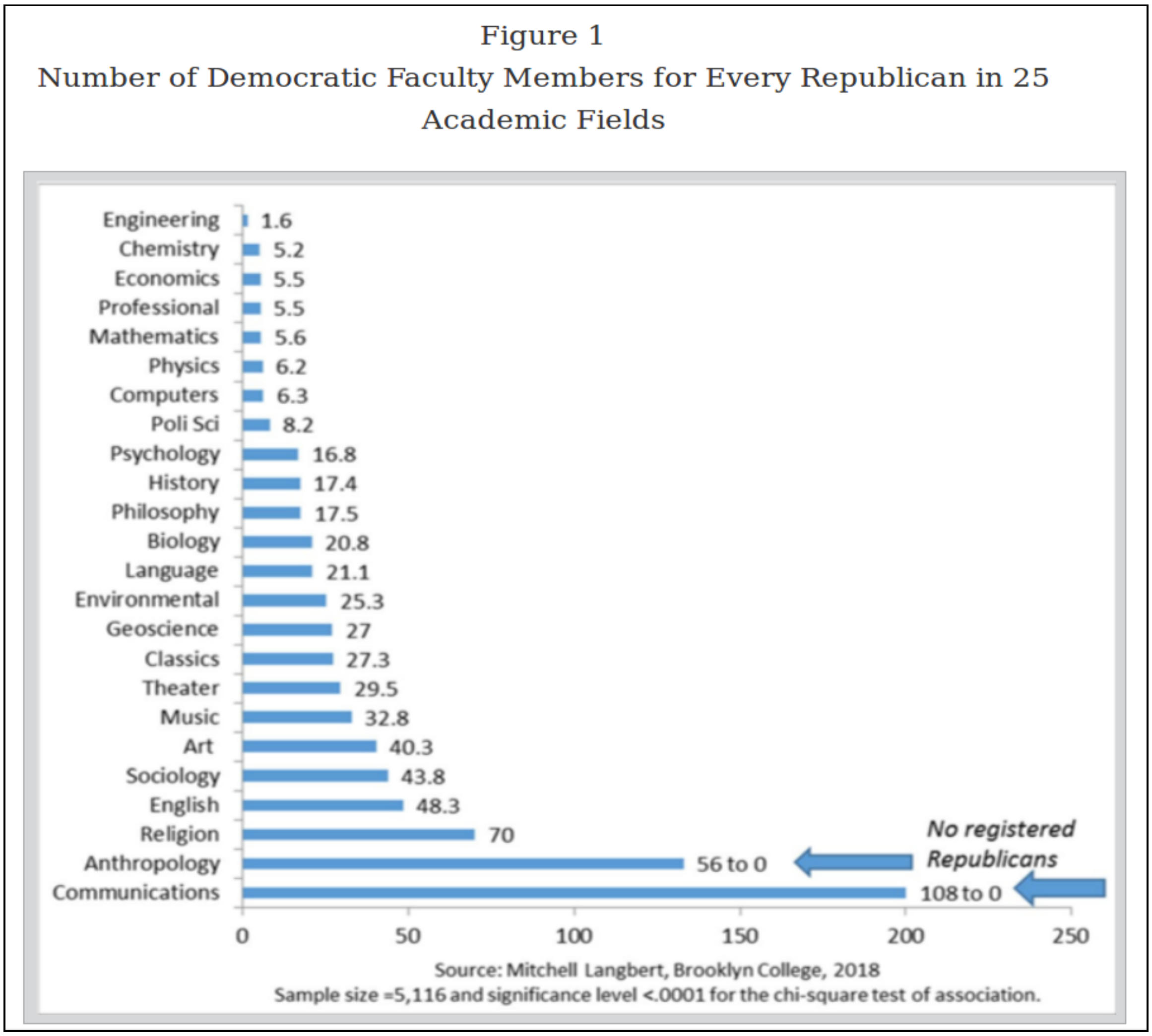
- incorrect, positive generalizations about Whites are racist,
- incorrect, negative generalizations about non-Whites are racist.

Some people may say that calling asians smart is, oddly enough, unintentionally harmful to asians, but rarely do they say that people who believe this are either anti-non-asians or that they are asian supremacists. On the flip side, saying that Whites are smart is called White supremacy and is thought of as being against everybody except for Whites. When asked about situations akin to the classic trolly problem, many are more willing to sacrifice a White for the greater good than they are to sacrifice a Black for the greater good (see more [here]). To the "anti-racists", racism is strongly a synonym for evil; **to "anti-racists", you are evil if you are not anti-White.**

Publication Bias:

An incredible left leaning distribution of political ideology in the university system is well documented [134 & 135]. The trend over time is an increasing leftward skew.

Source 135 - Figure 1:



These ratios may however be somewhat over inflated if registered republicans choose not to openly register for fear of retaliation; anonymous surveys of voting behavior would counteract this problem. This is just what we see from the surveys of source [122] which went over Economics, Political Science, History, Philosophy, Sociology, and Anthropology which allows us to compare it to the method of voter registration records for at least those fields. A comparison between the results of source [122] and the results of source [135] is summarized in the table below. *Anonymous Survey* and *Registered* both give the ratio of democrats to republicans, while ratio gives the ratio of ratios for the two methods.

Field:	Anonymous Survey:	Registered:	Ratio:
anthropology:	30.2	56	1.85
economics:	3	5.5	1.83
history:	9.5	17.4	1.83
philosophy:	13.5	17.5	1.3
political science:	7.7	8.2	1.22
sociology:	28	43.8	1.56

All in all, anonymity seems to multiply the number of registered republicans by about 1.5.

Leftist anti-Whiteness is well documented [more on that here]; the findings of particular interest are that liberals would support censoring research showing White genetic superiority with respect to intelligence more than they would support censoring evidence of Black superiority [460], and that liberals think Black people being genetically superior to White people with respect to intelligence is more plausible than the reverse [143]. Accordingly, publication bias typically seems to lean towards results that left leaning people would want; I'm not sure of any way to systematically demonstrate this other than pointing out how likely this is to be the case based on the findings thus discussed, but I have many documented examples of publication bias which fit with this view. Even a single example is substantial because it takes an enormous amount of evidence to do a single meta-analysis that proves one example.

Stereotype Threat:

Stereotype threat occurs in a situation in which it is plausible that some members of a social group may exhibit behavior which is typical of a stereotype about their respective group. It is thought that belief in one's groups' stereotypes induces feelings of threat that cause the stereotypes to become a self-fulfilling prophecy, and that stereotype threat effects partially contribute to long standing racial and gender gaps in academic performance, intelligence, etc. It is thought that these effects can be tested with so-called "primes" in tests. For an example, let's say two groups are given a test, and for one group the start of their test says that racial groups consistently perform equally on the test, while the control group gets no such prime, or perhaps the prime says that some group performs worse. If the prime group and the control group have different performances, this is supposed to be evidence for stereotype threat.

Or at least that's the theory. The evidence? A bunch of small studies with various p-hacking issues and then some larger studies with null results. Stereotype threat effects do not exist.

Test Settings:

One problem with the evidence in favor of the existence of stereotype threat effects is that it's all small studies in laboratory settings that aren't representative of the real world. The thing is that even in these laboratory experiments, when you introduce an incentive to perform well, stereotype threat effects disappear. For example, source 428 paid men and women money for getting correct answers, and introduced the stereotype threat prime quoted in the top of the right column.

> *"This is a diagnostic test of your mathematical ability. As you may know, there have been some academic findings about gender differences in math ability. The test you are going to take today is one where men have typically outperformed women."*

No stereotype threat effect could be elicited when subjects were paid for correct answers.

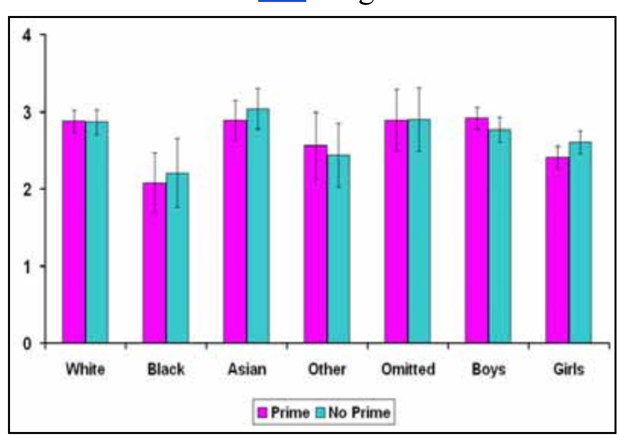
Stereotype threat effects cannot intentionally be tested in real world situations because if stereotype threat were real, it would be unethical to do so. We do however have a few instances in which this accidentally happened. Source 430 used design quirks in 1978-1999 NAEP tests where some, but not others, asked students their gender, and to choose strongly disagree, disagree, undecided, agree, or strongly agree for the following 3 statements:

- Math is more for boys than girls.
- Math is more for girls than boys.
- Fewer men have logical ability than women.

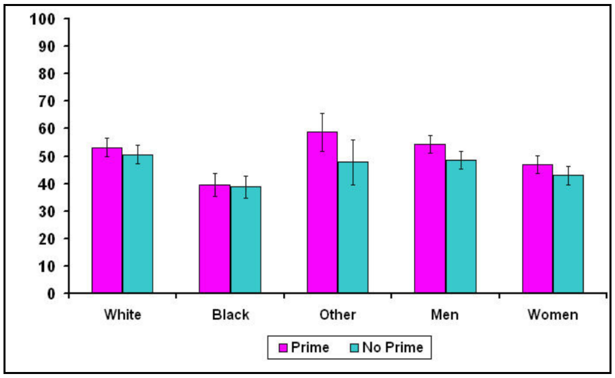
No evidence for stereotype threat was found.

In addition, a little known report [436] has some strong evidence based on two previous papers from the same author [429 & 437]. The figures speak for themselves:

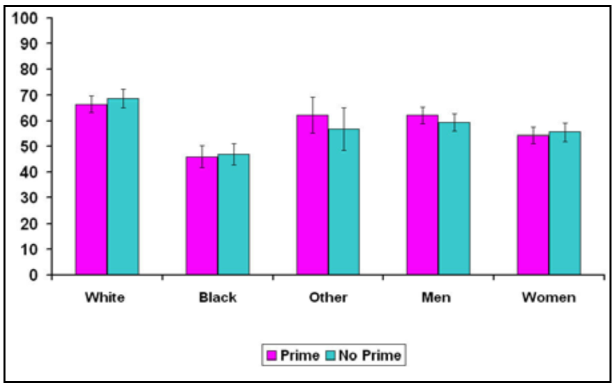
Source 436 - Figure 4:



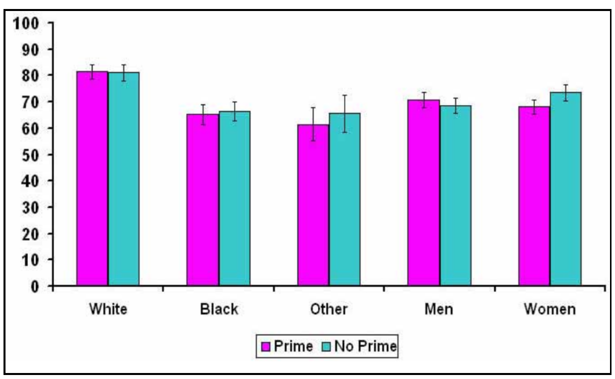
Source 436 - Figure 5:



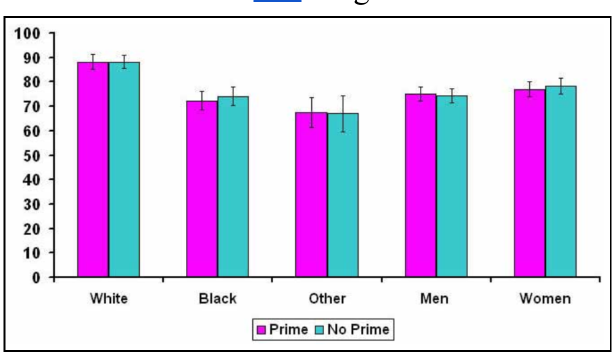
Source 436 - Figure 6:



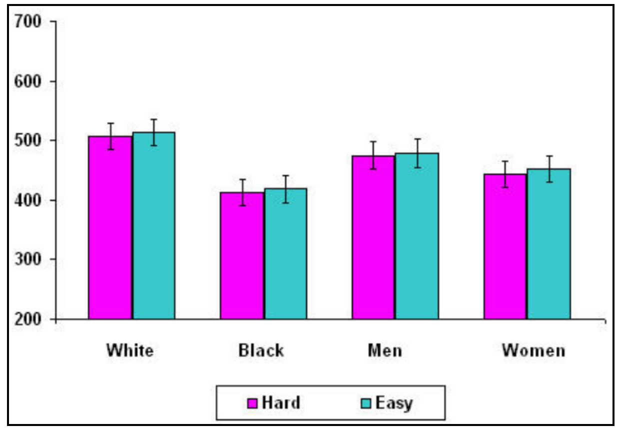
Source 436 - Figure 7:



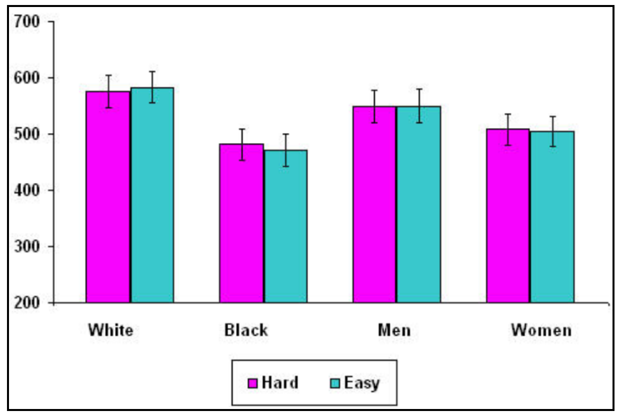
Source 436 - Figure 8:



Source 436 - Figure 20:



Source 436 - Figure 21:



Source 431 meta-analyzed stereotype threat in both the unrealistic and the operationalized testing settings. It found non-trivial evidence for publication bias, and that in the operationalized settings, stereotype threat primes had effect sizes ranging from .00 to -.14 standard deviations.

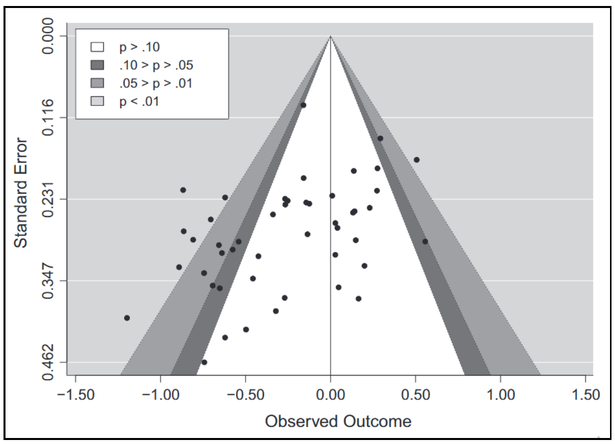
Sex: Females & Math:

For sex differences, women and math is the chosen target because women's relatively worse math performance is a major factor in their lower STEM representation and thus lower wages. Paulette Flore, former PHD student of Jelte Wicherts destroyed the idea of stereotype threat contributing to women's worse math performance with her PHD dissertation [432]. Some of the dissertation has been broken up and published separately as articles. Source [433](#) was her meta-analysis of the influence of stereotype threat on female performance. The mean of the 47 effect sizes was -0.22, however she notes the following:

> *"however, there were several signs for the presence of publication bias. We conclude that publication bias might seriously distort the literature on the effects of stereotype threat among schoolgirls. We propose a large replication study to provide a less biased effect size estimate."*

The funnel plot says it all.

Source 433 - Figure 3:



Source 434 is the replication study that Paulette proposed. The results:

> *"Among the girls, we found neither an overall effect of stereotype threat on math performance, nor any moderated stereotype threat effects."*

That studies go missing due to publication bias is also evidenced by the real unpublished manuscripts which have been found. Paulette's meta-analysis [433] found 2 unpublished manuscripts which supported stereotype threat effects, and 3 which did not support such effects. Source 435 found 4 unpublished manuscripts, and none of them supported stereotype threat. Thus, 7/9ths of unpublished manuscripts go against stereotype threat. The problem with finding unpublished manuscripts is that doing so is inherently difficult by nature of them being unpublished.

Race:

Source 438 looks at Hispanics in the United States and at immigrants in Europe. The funnel plot says it all:

Source 438 - Figure 1:

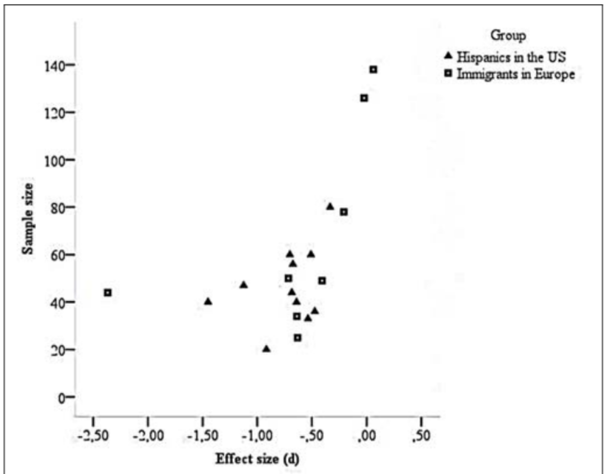


FIGURE 1 | Funnel plot based on effect size (d) and sample size. In studies with negative effect sizes, low stereotype threat groups outperformed high stereotype threat groups.

Worse yet than the simple suppression of valid but undesirable results is the fabrication of desirable results. There is one known instance [866] where the primary author of a paper in support of stereotype threat has admitted to fabricating fake data and requested the retraction of the paper [867].

On page 68 in the program for the 2009 ISIR conference [439], we see something interesting:

> **Stereotype threat and the cognitive test performance of African Americans**
>
> **Jelte M. Wicherts & Cor de Haan**
>
> University of Amsterdam
>
> Numerous laboratory experiments have been conducted to show that African Americans' cognitive test performance suffers under stereotype threat, i.e., the fear of confirming negative stereotypes concerning one's group. A meta-analysis of 55 published and unpublished studies of this effect shows clear signs of publication bias. The effect varies widely across studies, and is generally small. Although elite university undergraduates may underperform on cognitive tests due to stereotype threat, this effect does not generalize to non-adapted standardized tests, high-stakes settings, and less academically gifted test-takers. Stereotype threat cannot explain the difference in mean cognitive test performance between African Americans and European Americans.

If you check Jelte Wicherts' CV on the internet archives like the wayback machine https://archive.org/web/web.php or other sites like https://archive.is , what you will see is that the paper was floating around in review for quite a while before completely disappearing in his 2014 CV [440]. You can also occasionally find references to it in other places [441]. One may surmise that it was "lost in review". Somebody who is aware of what the results are inevitably going to be from reading the other meta-analyses of stereotype threat in other groups may not be so excited to publish a stereotype threat meta-analysis about race which looks at publication bias. Wicherts has been emailed to post a preprint several times to no avail [441].

We have another large stereotype threat replication pertaining to race [443] to look forward to which is similar to Paulette's big replication pertaining to sex [434]. It is pre-registered which means that the procedures are defined prior to publication, and there are certain tests which they will report the results of no matter what the results are which means that the authors can't just selectively report the only results that they find "interesting".

Self-Esteem/Stress/Positive Affect:

Even if we are to just ignore all the evidence and blindly accept stereotype threat theory, we would not expect stereotype to have affect the Black-White IQ gap because Whites have lower self-esteem, higher suicide races, more stress, etc:

Source 758:

This meta-analysis of 354 studies on racial differences in self-esteem finds that Blacks are 0.19 standard deviations higher than Whites in self-esteem. This has been the case for the past 50 years.

Source 840:

In this U.S. nationally representative sample of 38,891, Blacks self reported being less stressed than Whites did.

Source 759:

In this nationally representative sample, Whites are .28σ higher in risk for a panic disorder, .28σ higher in risk for generalized anxiety disorder, .12σ higher in social phobia, and had the exact same rate of PTSD.

Source 760:

In this nationally representative sample of 15-40 year olds, Whites scored .27σ higher than Blacks in major depressive disorder.

Source 786:

In this sample from 11 private, non-profit healthcare organizations constituting the Mental Health Research Network, with a combined 7,523,956, replicates these results finding Whites to universally have more psychological disorders than minorities, aside from Blacks being more likely to have schizophrenia disorders and miscellaneous disorders:

Reproduced from source 786 - Table 2:

Disorder	Asian	Black	Hispanic	Mixed	Native Amer. & Alaska Native	Hawaiian/Pacific Islander
Anxiety disorder	0.43	0.65	0.83	0.68	1.09	0.47
Any psychiatric diagnosis	0.36	0.69	0.72	0.64	1.03	0.47
Bipolar disorder	0.24	0.65	0.44	0.65	1.34	0.33
Depressive disorder	0.32	0.68	0.70	0.66	0.99*	0.46
Schizophrenia spectrum disorder	0.77	1.98	0.72	0.88*	1.18*	0.67
Other psychosis	0.50	1.13	0.61	0.34	0.80	0.51

Odds ratios of mental disorders by US racial groups, compared to the White prevalence scaled as 1.00. * indicated statistical insignificance, all other values differed with p<.001.

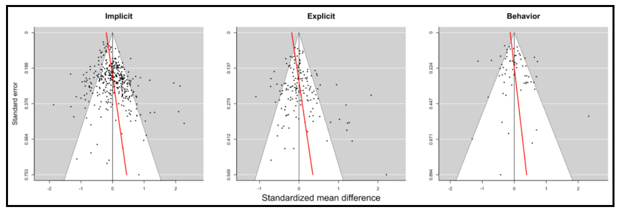
Conclusions:

All in all, stereotype threat doesn't seem to actually exist, and the literature is heavily plagued by publication bias. Remember what the goal of the publication bias is in the stereotype threat literature; if true, stereotype threat would make it harmful to even research group differences. It is not harmful, so do not be concerned about whether or not an argument is "racist". Instead, be concerned about whether or not an argument is correct (not politically correct, but actually correct).

Other Examples:

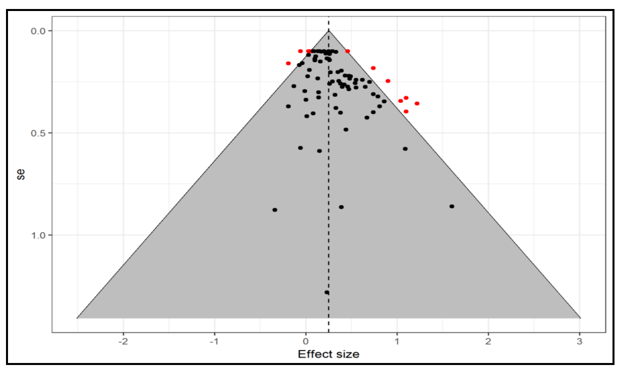
Implicit Associations:

In a similar vein to stereotype threat, implicit associations research aims to expose Whites as terrible evil "racists", but implicit associations tests have no validity for predicting actual behavior. Publication bias also inflates supposed validity even further [479].



Early Intervention Programs:

As gone over [here], publication bias inflates observed IQ gains from head start programs. These gains are not g-loaded, and they fade over time.



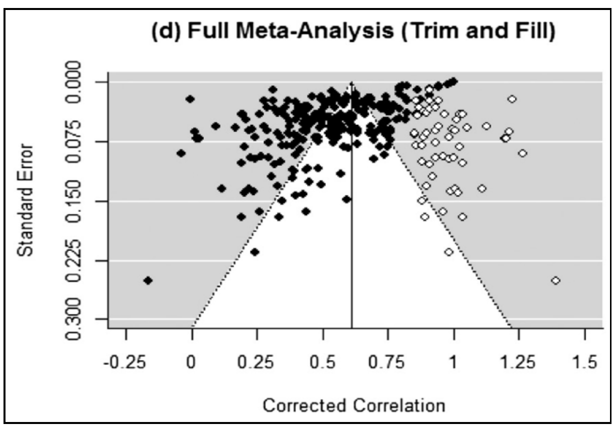
Callback Rates:

Long since pointed to as an example of pro-White discrimination, Whites get more callbacks from hiring employers (why this happens is a separate question). Anyways, the supplementary resources [606] (warning, direct download link!) from source 607 shows that the degree to which this happens is inflated by publication bias.

IQ & Grades:

Sometimes publication bias works to suppress the magnitude of results that really do exist instead of overinflating results that actually don't exist; a sort of "reverse" publication bias. The meta-analysis of the relationship between intelligence and school grades cited earlier [245] had some funnel plots. Trim and fill is basically a method to combat publication bias where you add imaginary studies to the meta-analytic effect until there is no correlation between effect size and standard error in order to see what the effect size would be without publication bias. The White dots are the imaginary studies from Trim and Fill, and the black dots are actual studies.

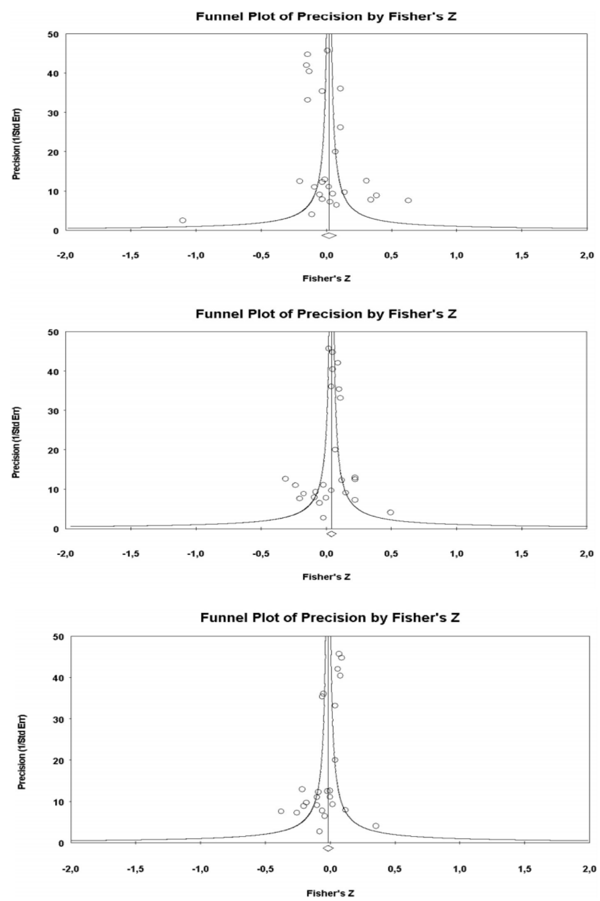
Source 245 - Figure 1:



Scarr-Rowe Effects:

There is actual math where you plug in the heritability of a trait, and the magnitude of group differences in terms of the trait, and it tells you how much worse the poorer-performing group's "environment" has to be in order for the between-group heritability to be 0%. Scarr-Rowe effects (heritability being larger for rich people than for poor people) would mean that the difference between $0 per year and $10,000 per year has a larger impact on intelligence than the difference between $50,000 per year and $60,000 per year, or that more basic environmental improvements matter more than the others even though the magnitude of improvement is the same. If true, this would mean that group differences in intelligence have a smaller genetic component than otherwise assumed. Multiple meta-analyses show that Scarr-Rowe Effects don't exist and that their effect sizes are inflated by publication bias [see more here]. Here are the racial Scarr-Rowe funnel plots:

Source 300 - Figure 2:



Funnel plots of precision by Fisher's Z for A, C, and E, respectively. The x-axes shows Fisher's Z and the y-axes shows precision, measured as the inverse of the standard error.

Race Differences In Personality:

Though racial differences in personality (based on self-report data) are small, there are still signs of reverse publication bias [145]. However, this finding should be taken with a grain of salt because of the reference group effect. Basically, when, for example, somebody says that they are low in neuroticism on a survey, part of the heuristic they are using is that they are low in neuroticism **in comparison** to the people that they regularly interact with [643, 644 & 645]. Evidence on differential item functioning is rare, but it seems that personality fails metric invariance [646].

Racial Bias In Criminal Sentencing:

Source 608 didn't do a funnel plot, but rather analyzed real unpublished studies that they managed to find. Unpublished studies found less bias than published studies.

Video Games & IQ:

A meta-analysis [693] on the experimental effect of video games on intelligence finds that publication bias inflates effect sizes by 30%.

Brain Size & IQ:

There is a well established causal link between brain size and IQ [see more here], however, the size of the association is inflated by publication bias [362]. This is the only example of publication bias I know of that anybody could consider pro-Hereditarian. However, this point isn't of much importance to Hereditarianism as there are many other plausible brain variables, and it would be odd for a Hereditarian to seriously think that a single brain variable would explain so much.

Reading Intervention & Reading Ability:

A meta-analysis on the effect of shared book reading shows it to have a very small effect on language development, that the effect that it does have is inflated by publication bias, and that the fadeout effect for interventions is also replicated [694].

GxE & The EEA:

Most detected gene-environment interaction effects, especially novel effects, fail to replicate [868 & 869]. Failed replications also typically have more statistical power than successful replications, indicating that publication bias is in favor of the existence of gene-environment interaction effects [868].

The Efficacy Of Intelligence Research:

As Neven Sesardić conjectures on page 205 of his book, *Making Sense Of Heritability* [150], double standards in requirements for evidence could strengthen the evidence for Hereditarianism. If Hereditarians have their work picked apart for any potential mistakes that are seen as a sign of malicious political bias rather than human fallibility when discovered, hereditarians would likely take special care in putting extra effort into making sure that their evidence is strong in order to combat such a research environment.

Is there evidence to support this conjecture? Yes. Intelligence and behavioral genetics research, by having a roughly 50-50 political split, is likely the most republican field in academia [151]. Accordingly, Intelligence research, and particularly, intelligence research on group differences, suffers less from problems with statistical power than other fields do.

Discipline:	Mean / Median Statistical Power:	Citation:
Neuroscience	21%	Source 156
Brain Imaging	8%	Source 156
Social and Behavioral Sciences	24%	Source 155
Cognitive Neuroscience	14%	Source 154
Psychology	23%	Source 154
Medical Research	23%	Source 154
Breast Cancer	16%	Source 152
Glaucoma	11%	Source 152
Rheumatoid Arthritis	19%	Source 152
Alzheimer's Disease	9%	Source 152
Epilepsy	24%	Source 152
Multiple Sclerosis	24%	Source 152
Parkinson's Disease	27%	Source 152
Education	23%	Source 153
Intelligence	49%	Source 14
Intelligence	11.9%	Source 647
Intelligence - Group Differences	57%	Source 14

Notes on table creation: Source 14 is the 2018 preprint which is, frankly, superior to the published version. Power to detect median effect was used wherever possible. In some mega-analyses, power to detect median effect was not reported; in these, median effects were small, so power to detect small effects was used.

Note: Source 14 - Table 2 has some more subareas of intelligence research:

Table 2

Descriptive statistics of the primary studies split up in five types of studies and in total.

	# Meta-analyses	# Unique primary studies	Total N	Median N	Range N	Median unweighted Pearson's r	Median meta-analytic effect (r)	Median power
1. Predictive validity & correlational studies	31	779	367,643	65	[7; 116,053]	0.26	0.24	53.3%
2. Group differences (clinical & non-clinical)	59	1,247	19,757,277	59	[6; 1,530,128]	0.26	0.19	59.3%
3. Experiments & interventions	20	188	24,371	49	[10; 1358]	0.18	0.17	26.5%
4. Toxicology	16	169	25,720[a]	60	[6; 1333]	0.15	0.19	23.9%
5. (Behavior) genetics	5	59	30,545	169	[12; 8707]	0.07	0.08	9.3%
Total	**131**	**2,442**	**20,205,556**	**60**	**[6; 1,530,128]**	**0.24**	**0.18**	**51.7%**

Note. "N" indicates number of participants in a primary study. We calculated the meta-analytic effects per subtype by taking the median of the random effects meta-analyses estimates. We calculated the power of each primary study to detect the summary effect in the corresponding meta-analysis. We reported the median of all power estimates per subtype.

[a] One of the meta-analyses reported two studies with non-integer total sample sizes. It seems that the authors wanted to correct their sample sizes to ensure they did not count the same observations twice. Here, we rounded the total sample size.

The only thing that's really surprising is the 9.3% statistical power of "(Behavior) genetics". This seems implausible given my experience of the state of behavioral genetics research, and indeed, an email exchange between Emil Kirkegaard and Michéle [157] reveals that the meta-analyses under the behavior genetics category were mostly useless candidate gene studies. The email:

"

Hi Emil,

We included 5 meta-analyses that we labelled as behavior genetics.

Three of these are candidate gene studies:

Barnett, J. H., Scoriels, L., & Munafo, M. R. (2008). Meta-analysis of the cognitive effects of the catechol-O-methyltransferase gene val158/108Met polymorphism. Biological Psychiatry, 64(2), 137-144. doi:10.1016/j.biopsych.2008.01.005

Yang, L., Zhan, G.-d., Ding, J.-j., Wang, H.-j., Ma, D., Huang, G.-y., & Zhou, W.-h. (2013). Psychiatric Illness and Intellectual Disability in the Prader-Willi Syndrome with Different Molecular Defects – A Meta Analysis. Plos One, 8(8). doi:10.1371/journal.pone.0072640

Zhang, J.-P., Burdick, K. E., Lencz, T., & Malhotra, A. K. (2010). Meta-analysis of genetic variation in DTNBP1 and general cognitive ability. Biological Psychiatry, 68(12), 1126-1133. doi:10.1016/j.biopsych.2010.09.016

One is a candidate gene study involving twins:

Luciano, M., Lind, P. A., Deary, I. J., Payton, A., Posthuma, D., Butcher, L. M., . . . Plomin, R. (2008). Testing replication of a 5-SNP set for general cognitive ability in six population samples. European Journal of Human Genetics, 16(11), 1388-1395. doi:10.1038/ejhg.2008.100

The fifth one studies heritability with twins:

Beaujean, A. A. (2005). Heritability of cognitive abilities as measured by mental chronometric tasks: A meta-analysis. Intelligence, 33(2), 187-201. doi:10.1016/j.intell.2004.08.001

Hope this helps!

Best,

Michèle
"

Here are the sources mentioned in the email given source numbers:

Email:	The Secrets Of The Cakes:
Barnett 2008	Source 158
Yang 2013	Source 159
Zhang 2010	Source 160
Luciano 2008	Source 161
Beaujean 2005	Source 162

Additionally, reanalysis [648] of source 647's intelligence research data [649] with z-curve 2.0 finds no evidence of publication bias or questionable research practices.

For another cross-field comparison, here are replication rates:

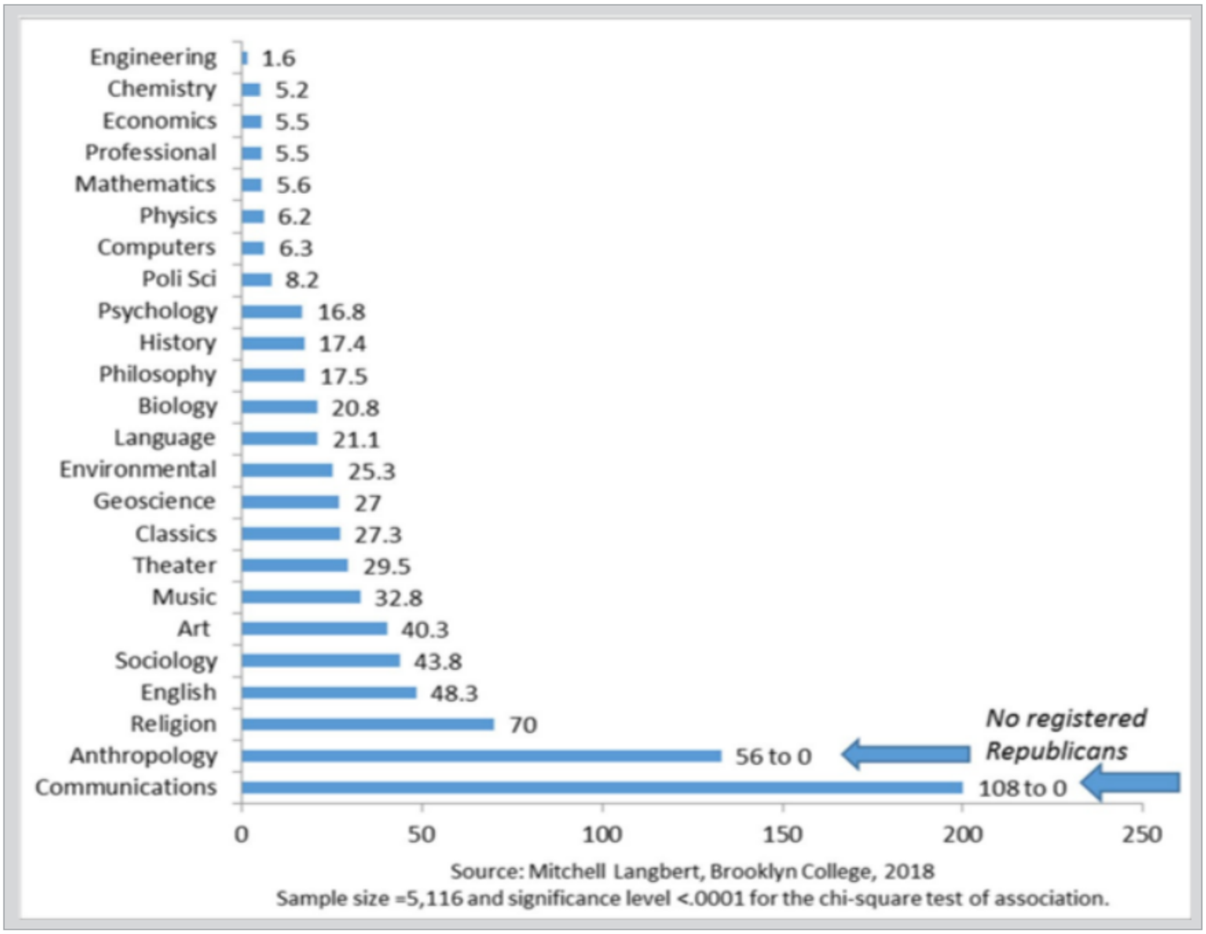
Discipline:	Replication Rate:	Citation:
Differential Psychology	87%	Source 287
Experimental Philosophy	70%	Source 144
Economics	61%	Source 126
Cognitive Psychology	50%	Source 125
Social Psychology	25%	Source 125
Pharmacology	21%	Source 139
Oncology (cancer)	11%	Source 127
Neuroscience	<6%	Source 128
Discipline:	**Self Reported Expectations For Replication Rate:**	**Citation:**
Physics	73%	Source 124
Chemistry	65%	Source 124
Astronomy	65%	Source 124
Material Science	60%	Source 124
Biology	59%	Source 124
Earth and Environmental Science	58%	Source 124
Engineering	55%	Source 124
Medicine	55%	Source 124
Other	52%	Source 124

Finally, here are the ratios of democrats to republicans by field posted once more for the sake of comparison with the previous two tables.

Source 135 - Figure 1:

Figure 1

Number of Democratic Faculty Members for Every Republican in 25 Academic Fields



The Anti-White Media:

In contemporary America, professors openly say things like *"All I want for Christmas is white genocide"* [146] or *"OK, officially, I now hate white people,"* [147]. Teaching assistants claim that *"some white people may need to die"* so that Black people can get what they deserve [146]. Editors at the New York Times assert that *"White men are bullshit"*, use the hashtag *"CancelWhitePeople"*, and complain about *"Dumbass fucking white people marking up the internet with their opinions like dogs pissing on fire hydrants"* [170].

This is the same New York Times which published a piece entitled *"Can my Children be Friends with White people?"* [171], a question which the author answers largely in the negative: *"As against our gauzy national hopes, I will teach my boys to have profound doubts that friendship with white people is possible. When they ask, I will teach my sons that their beautiful hue is a fault line. Spare me platitudes of how we are all the same on the inside. I first have to keep my boys safe, and so I will teach them before the world shows them this particular brand of rending, violent, often fatal betrayal."*

Sometimes, White people don't like this sort of stuff. For instance, a few complained about the New York Times editor previously mentioned, but writers for NBC News explained that *"white people getting mad — or publicly performing anger, at least — about white people jokes is actually white people getting mad about threats to white power. Threats like a woman of colour joining the editorial board of the New York Times after telling smarter and funnier jokes than them on Twitter. Racism is a mechanism of maintaining an imbalance of power — making it literally impossible, by definition, to be racist against white people, or to tell a racist joke about a white person"* [445]. Similarly, The Chicago Tribune has stated that *"American racism is a uniquely white trait"* [446].

USA Today has made this point too, that only white people can be racist [447]. They've also noted that *"A majority of white Americans believe discrimination exists against them in the United States"* [448] but have explained that this is not to be taken seriously [449], arguing that *"America's newest class of victims — i.e., white men — is on the warpath again. They complain that they can't get into college because of affirmative action, can't get a job because of diversity hiring, and can't keep a job because of factories closing due to unfair trade deals. Now we can add to the "whine list" the fact that many white men feel they can no longer get ahead or get an advantage because of identity politics."*

CNN has published material explaining that White people who disagree with non-whites about racism are often engaging in *"Whitesplaining"* [450]. This term was defined as follows: *""Whitesplaining" is an affliction that's triggered when some white people hear a person of colour complain about racism. They will immediately explain in a condescending tone why the person is wrong, "getting too emotional" or "seeing race in everything.""*

The article went on to cite telltale signs of Whitesplaining, such as when White people say things like *"But I'm not a racist"*.

Other times, White people agree with these narratives and devote themselves to fighting White supremacy. This can take an emotional toll on White people as a kind of racial self hatred. The New York Times has noted this in an advice column responding to a woman whose sense of White guilt caused them to have a mental breakdown [451]. As they explain, White suffering is ultimately unimportant: *"You have to relinquish your privilege. And part of learning how to do that is accepting that feelings of shame, anger and the sense that people are perceiving you in ways that you believe aren't accurate or fair are part of the process that you and I and all white people must endure in order to dismantle a toxic system that has perpetuated white supremacy for centuries. That, in fact, those painful and uncomfortable feelings are not the problems to be solved or the wounds to be tended to. Racism is."*

NBC has also acknowledged the psychological toll of their ideology [452], telling White people that, *"you're going to have to take a side. And yes, you have to do it now. It's very likely, and understandable if you feel this is unfair, this is inconvenient, it's frustrating, it's difficult, it's embarrassing, it's going to alienate you from people you know, love, work with, watch the game with. That's privilege. Someone once said, "when you're accustomed to privilege, equality feels like oppression." This is a taste of equality."*

And Forbes too has said that White people need to stop caring so much about their own suffering [453]: *"If you are not Black, your pain and hurt is not the priority right now. This may be an anomaly for you – it is not an anomaly for Black folks who live this life, everyday."*

In the political realm, Joe Biden has talked about how White people becoming a minority is not only not-bad, but in fact a positive good which will improve the country [454].

These news outlets, CNN, the NYT, USA Today, Forbes, and NBC, are not seen as organizations of the radical left. Like Joe Biden, they are seen as center left or moderate, though by all quantifiable evidence, the field of journalism, as a whole, should be seen as being heavily biased leftwards [444]; journalists vote liberal, they say that they are liberal, they reject non-liberal positions, and the general public recognizes them as liberal.

If we looked further to the left, we'd find things like Bernie Sanders saying *"When you're white, you don't know what it's like to be living in a ghetto. You don't know what it's like to be poor."* [456], Buzz Feed running articles like *"37 Things White People Need To Stop Ruining In 2018"* (the first of which, apparently, is America) [457], Vice positively covering vacations non-Whites take just to get away from White people [458], and The Root publishing articles with titles like *"White people are cowards"* [459] which conclude *"I thought white people were evil. I was right."*

The Anti-White Left:

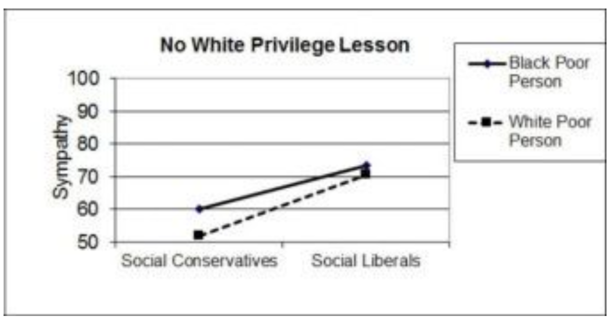
A left leaning media [444] being anti-White is consistent with leftist anti-Whiteness at large:

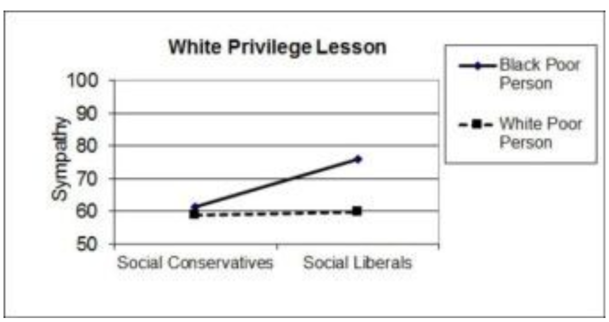
Finding:	Citation:
Liberals are more willing to murder someone for the greater good if that person has a White-sounding name rather than a Black-sounding one.	455
Liberals think that Black people being genetically superior to White people with respect to intelligence is more plausible than the reverse.	143

Continued:

Finding:	Citation:
Hearing about White privilege causes liberals to feel less sympathy for poor White people.	461
Liberals feel non-Whites should not pay more for home insurance due to living in a high-risk area but are neutral about whether or not White people should.	462
Liberals would support censoring research showing White genetic superiority with respect to intelligence more than they would support censoring evidence of Black superiority.	460

It should be noted that to accuse the left of being anti-White is not to accuse the left of being genuinely pro-Black. Source 461 found that exposing people to left wing messages about White privilege caused their sympathy for poor Whites to decrease while their sympathy for poor Blacks remained the same:





Similarly, source 464 finds that:

> *"Across five experiments (total N = 2,157), White participants responded to a Black or White interaction partner… liberals—but not conservatives—presented less competence to Black interaction partners than to White ones… This possibly unintentional but ultimately patronizing competence downshift suggests that well-intentioned liberal Whites may draw on low-status/competence stereotypes to affiliate with minorities,"*

In other words, White liberals talk to Black people like they are children or pets who need maternal protection from White people. The boomer-conservative talking point was true.

-White Guilt:

Liberals have, on average, lower self esteem than conservatives [465].

Identifying with one's own race is positively correlated with self esteem:

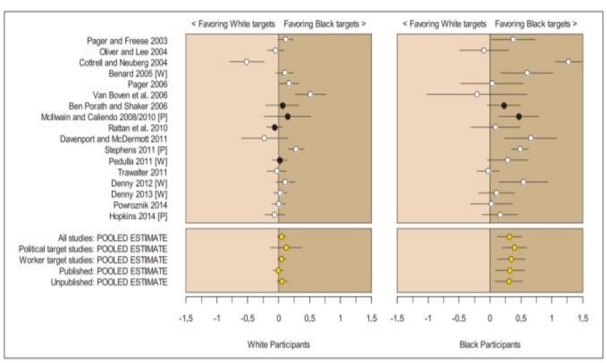
N	Whites	Blacks	Hispanics	Citation:
669	.44	.17	.27	476
5423	.24	.11	.14	473
898	.27	.39	.27	474

There is even evidence that making people feel more physically attractive causes them to lean more right wing [466]. This may explain why more attractive people and politicians are more right leaning [467 & 468].

Whites also have the lowest level of racial identification of any ethnic group in America:

Finding:	Citation:
One in seven (15%) of Whites, 56% of Asians, 59% of Hispanics, and 74% of Blacks say that their race/ethnicity is central to their identity	475
On a measure of ethnic identity, Black Americans scored higher than Latinos who scored higher than Whites.	476
Across ten ethnic groups, Black Americans had the highest score on a measure of ethnic identity while White Americans had the lowest.	473
Across five ethnic groups, Black Americans had the highest score on an ethnic identity measure while White Americans had the lowest.	477
On a measure of ethnic identity, Black Americans scored the highest followed by Hispanic Americans who scored higher than White Americans.	474

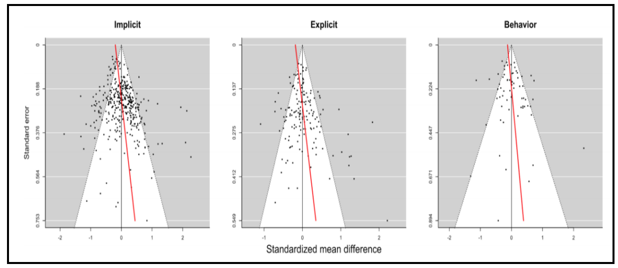
Unlike Black Americans, White Americans generally don't exhibit any racial bias in formal experiments measuring "racism" [478]:



Leftists sometimes deny this based on the results of implicit association tests which are supposed to measure subconscious biases which people may be totally unaware of.

In addition, a huge meta-analysis [479] with 92 studies and 87,418 participants finds that changing implicit bias measures has no effect on explicit bias or actual behavior. It also finds significant evidence that publication bias inflates its supposed validity.

Source 479 - Figure 9:



There is also more direct evidence that feelings of White guilt have gone up over time, and that leftist ideology has a direct impact on White guilt. Research on the average level of White guilt seems to have started in the 1970s [469]. Guilt was measured on a 5 point scale (5 = maximum guilt) with questions like "Do you feel personally guilty about the American Negro's present social inequality?" The results:

1. Personal guilt about past, 1.70.
2. Personal guilt about present, 2.10.
3. Guilt of immediate family, 1.71.
4. Guilt of white friends, 2.02.
5. Guilt of white society, 1.82.

The next known paper comes from 1999. Agreement with the same sorts of statements as before was rated on a 5 point scale, and the average response was 2.12, implying only slight guilt and that the mean level of guilt had not changed much since the 1970s [470]. It should be noted that the vast majority had at least some guilt with only 6% saying that they strongly disagreed with all 5. The same scale was administered to a sample of college kids in 2007 [471]. This time, the mean response was 3.64. After these students took a diversity course, the mean score increased to 3.94, implying a good deal of guilt, and implying that leftism causes such guilt. Similarly, source 472 reported the following:

> *"In Experiment 1 (N = 110), White American participants assessed 24 statements about racial inequality framed as either White...*

Continued:

> *...privileges or Black disadvantages. In Experiment 2 (N = 122), White participants generated examples of White privileges or Black disadvantages. In both experiments, a White privilege framing resulted in greater collective guilt"*.

6. The Existence Of Race

Navigation:

Previous Chapter   Table Of Contents   Next Chapter

Summary:

Many boldly insist that race does not exist. When you dig below the surface, this seems to be a semantic game. Taxonomy is subjective, you can call things what you like, and 2+2=5 if you define the symbol “5” as the concept of “four”. Where I take issue is when people hear the statement “race is a social construct / more variation within than between” and think this implies the truth of statements shown to be falsehoods such as “there are no genetic differences between the races” or “an individual of one race can be more genetically similar to another race than his own race”. Ideally, we should just treat human variation like any other animal and apply the same standards. Doing this, we see that the human “races” seem to hit similar markers to those of the subspecies of many other animals.

Definitions:

Heterozygosity:

At a given gene locus, there are variants, if two people have a different gene-variant at a locus, they are heterozygous at that locus. Heterozygosity for a locus is the percentage of the population which is heterozygous on that locus.

F_{ST} (a.k.a. Fixation Index):

A species may have subspecies. You can calculate heterozygosity for the entire species for a locus, let’s call this total heterozygosity (H_T). Alternatively, you can calculate heterozygosity for a specific subspecies on that locus, let’s call this subpopulation heterozygosity (H_S). Average together every H_S figure on that locus and we’ll call that H_S’. Subtract H_S’ from H_T (H_T - H_S’), and we’ll call the result D_{ST}. What percentage of H_T is D_{ST}? (D_{ST} / H_T)? D_{ST} / H_T = F_{ST}. If the loci of an F_{ST} value isn’t specified, assume this refers to the average of F_{ST} values for all recorded loci. An F_{ST} can also be a genetic distance between two specific subspecies where H_T is heterozygosity of the two subspecies pooled, and H_S’ is heterozygosity of the two subspecies.

“Evolution Takes A Long Time”

“Out of Africa II” refers to the migration of modern humans out of Sub-Saharan Africa after our emergence about 200,000 to 300,000 years ago. Source 537 suggests that early Homo sapiens, or "another species in Africa closely related to us," might have first migrated out of Africa around 270,000 years ago. Finds at Misliya cave, which include a partial jawbone with eight teeth, have been dated to around 185,000 years ago [538].

By comparison, here is how long ago the subspecies of various other animals diverged from each other:

Subspecies:	Subspecies’ Time Of Divergence:	Source:
European Moose - American Moose	165,000	540
Polar Bears - Brown Bears	152,000	539
Humans	~200,000	538
(birds) Cyanoptera – Septentrionalium	95,000	541
(birds) Discors – Septentrionalium	70,000	541
(birds) Cyanoptera – Discors	65,000	541
Tigers: 8 subspecies	72,000	542
Eastern & Western Wood Ducks	34,000	543
2 Lizard Subspecies	12,000	544

Note: Although the split from Brown Bears happened 152,000 years ago, Polar Bears are estimated to have genetically adapted to their new environment within only 10,000-30,000 years.

This isn’t necessarily to say that genetic changes between populations have to take these spans of time. If all of humanity’s tall people were genocided, then the very next generation would instantaneously be genetically predisposed to be shorter than the previous generation was. Further, in a famous Soviet experiment, a group of silver foxes, were domesticated via selective breeding within just 10 generations [489]. In addition, the selection, intended exclusively for this behavioral trait, led to population changes in physical traits such as floppy ears.

-The Implausibility Of Equality:

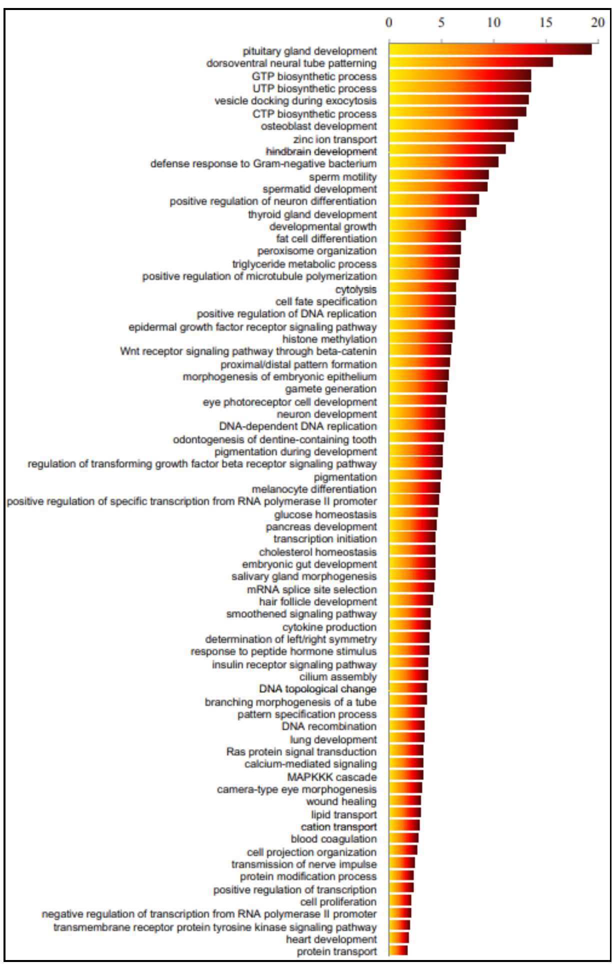
The argument for Hereditarianism which people rate to be the most effective and which convinces most people is not technically the best, most comprehensive one. It is a rather simple question:

Given that different people evolved in different places with different climates, different diseases, different challenges, different plants & wildlife, etc, what is the chance that evolution stopped at the neck?

What is the chance that there happens to be zero difference in parts of the genome related to cognition despite ~40% of the genome influencing cognition [672 & 673]? What is the chance that all the different people groups of the world evolved to have the exact same amount of all the different intelligences despite the population differences in Neanderthal ancestry [636] which has associations with skull shape [671]? If IQ gaps are due to oppression, why do Blacks score better on the long term memory factor [670]?

Especially given that human evolution has sped up by a factor of 100 in the past 5000 years [674], and genes involved in the brain are overrepresented among those having recently undergone selection [611], we should not be surprised that as it turns out, racial differences in terms of genes involved in the brain are larger than the racial differences in terms of genes involved in physical traits like skin colour or hair texture [610].

Source 610 - Figure 1:



λ values of GO categories in biological processes enriched for higher F_{ST} SNPs with P-value lower than 10^{-10}

-"We're 99.9% the same!":

For the same reason that it is claimed that the existence of race is implausible because of how long evolution takes, many claim that human genetic variation is too small to permit races, or really much variation at all. One thing many have probably heard is the famous phrase that "We're all 99.9% the same!". It comes from Craig Venter, and in 2007, he was involved in a second analysis which revised the number down to 99.5% [545]. Heres what the 99.5% number means. We get 23 chromosomes from each parent, and all of them aside from the y chromosome have a counterpart copy coming from the other parent. 99.5% similarity is just the sequence similarity between the two chromosome copies that an individual person has. It is assumed to be representative of all of humanity, but genetic assortative mating exists, and between-race similarity on a given chromosome is probably smaller than that. Keep in mind that by the same scale, Humans are 98.76% similar to chimpanzees [555]. Using a more appropriate measure, within species heterozygosity, it is also clear that heterozygosity within humans is well within the normal bounds for other species. Human heterozygosity seems to be even higher than many other species (see the following table).

Heterozygosity By Species:

Species:	Heterozygosity:	# Of Subspecies:	Source:
Humans	.776	?	547
Humans	.70-.76	?	548
Humans	.698	?	549
Chimpanzees	.63	4	547
Chimpanzees	.765	4	550
South African Buffalo	.729	5	551
Leopards	.58	13	552
Jaguars	.739	9	553
Pumas	.52	6	554
Canadian Lynx	.66	3	556
North American Brown Bears	.5275	19	557
Scandinavian Brown Bears	.678	19	558
Coyotes	.629	19	559
Gray Wolves	.574	37	559
Domestic Dogs	.5085	?	559
African Wild Dogs	.643	5	560
North American Wolverines	.55	2-3	561
Scandinavian Wolverines	.325	2-3	562
Elk	.395	7-8	563
Bighorn Sheep	.6235	3	564
Bonobos	.535	1	565
Polar Bears	.68	1	566
Australian Dingoes	.445	1	567

Average Of Human Estimates: .733 Average Of Non-human Estimates: .58694

[“More Variation Within Than Between”]

Though human subspecies are plausible, some would claim that the proposed races happen to not be genetically distinct enough to warrant the label. F_{ST} (Fixation Index) is the proportion of total variation at a gene loci that exists between two populations compared to the total variation within both populations.

Richard Lewontin became the first to measure human F_{ST} in 1972 [612], and he found it to be .063. Based on this finding, Lewontin declared that categorizing humans racially has no “genetic or taxonomic significance”. He never explains why this number is too low, he just says that race is meaningless since the difference is 6.3%.

The first important thing to point out about F_{ST} statistics is that when only one is given for an entire group difference, that is probably the average F_{ST} for all tested loci. Pointing the average F_{ST} out and saying we can’t predict race based on genes is to be ignorant of the concept of binomial probability:

Let’s plug Lewontin’s 6% into a binomial probability calculator and say we’re trying to predict a person’s race in a 2 race category scheme. We know person A’s race is race 1 (R1) rather than R2, and that the F_{ST} between R1 and R2 is 6%. If average F_{ST} was 0%, then somebody’s loci would tell us nothing about their race and we would have a 50% chance of successful prediction. With an F_{ST} of 6%, a single loci will give us a 56% chance of successful prediction. With 2 gene loci, the probability of person A having less in common with R1 than R2 in terms of those two loci is 19.36%, the probability of having the same amount in common with both races is 49.28%, and the probability of having more in common with R1 than R2 (Let’s call this the probability of outcome 1, or O!) is 31.36%. With 4 loci, the probability of O1 is 40.7%. With 100 loci the probability of O1 is 86.6%. With 1,000 loci the probability of O1 is over 99.99%.

This theoretical demonstration of binomial probability is experimentally borne out by the [clustering studies], though I’m not sure why citations should be needed for common sense..

There is another important thing to point out about F_{ST}, what is human F_{ST} 6.3% of? F_{ST} =/= D_{ST}. To put it differently, let H_T be the total amount of heterozygosity within an entire species, and let H_S’ be the amount of heterozygosity within the subspecies. F_{ST} is the difference between H_T and H_S’ expressed as a percentage of H_T [(H_T - H_S’) / H_T]. The same absolute difference in heterozygosity can produce wildly different F_{ST} values, and wildly varying differences in heterozygosity can produce the same F_{ST} value:

HT	**HS**	**F_{ST}**	**-**	**HT**	**HS**	**F_{ST}**
.9	.8	1/9	-	.9	.45	1/2
.8	.7	1/8	-	.8	.40	1/2
.7	.6	1/7	-	.7	.35	1/2
.6	.5	1/6	-	.6	.30	1/2
.5	.4	1/5	-	.5	.25	1/2
.4	.3	1/4	-	.4	.20	1/2
.3	.2	1/3	-	.3	.15	1/2
.2	.1	1/2	-	.2	.10	1/2
.1	0	1/1	-	.1	.05	1/2

F_{ST} by Species:

Species	F_{ST} Distances:	# Of Subspecies/Groups:	Source:
Humans (k=10)	11.7%	10	568
Humans (k=8)	12%	8	569
Humans (k=8)	11.9%	8	570
Humans (k=3)	16.3%	3	571
Jaguars	6.5%	9	553
Canadian Lynx	3.3%	3	556
Asian Dogs	15.35%	11	572
African Buffalo	5.9%	5	573
North American Coyotes	10.7%	19	574
North American Wolverines	7.6%	2-3	575
Gray Wolves	16.7%	37	574
Humpback Whales	12%	3	576
Plains Zebras	11%	3	577
Kob Antelope	11%	2-3	578
South West European Cow	6.8%	18	579
Red Winged Black Bird	0.9%	5	580

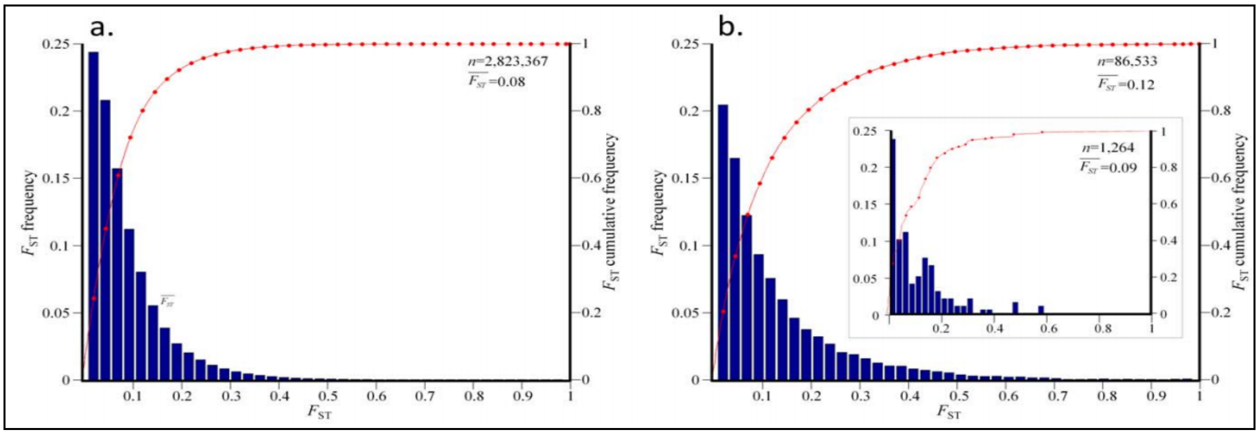
Note: Successful prediction of race from genes is confirmed by the [clustering studies].

And so we can clearly see that the ignorance of binomial probability would lead to subspecies denial if "more variation within than between" "logic" were applied to other species. Humans are just an animal like any other, we only treat ourselves differently because of political considerations.

For an even more outrageous example, source 570 calculated F_{ST} for humans resulting in a value of 11.9%, and then when it added a population of Chimpanzees, F_{ST} only went up to 18.3%. It would seem that not only are Humans and Chimpanzees the same species, they can't even be considered subspecies because there's more variation within than between. Obviously, the only reason Chimpanzees can't speak proper English is because of their poor school funding.

By far, the largest human sample was source 569, which recorded millions of SNPs. Something important to note is that source 569 also looked at the distribution of F_{ST}s, and the median F_{ST} value is much smaller than the mean F_{ST} value. Most F_{ST} values are pretty small, but a somewhat small number of loci have F_{ST} values that are much larger than median thereby dragging the average upwards. So, smaller samples like Lewontin's are likely to underestimate the true average by missing the SNPs with larger F_{ST} values.

Source 569 - Figure 4:



Additionally, with an average F_{ST} larger than 6%, not as many SNPs are needed to predict race.

Racial Differences Compared To Family:

Henry Harpending's paper, "Kinship and Population Subdivision" [613], explains why F_{ST} functions as an inverse kinship coefficient divided by 2. The F_{ST} distance between the races is .12 [569], which can be modeled as a -.24 kinship coefficient:

Relationship:	Kinship:
MZ Twins	1.0
Parent-Child, DZ Twins	0.5
Aunt, Uncle, Niece, Nephew, Grandparent	0.25
First Cousin	0.125
Racial Differences	-0.24

Racial Differences Compared To Sex:

Humans mostly all share the same 46 chromosomes, except men have a y chromosome instead of a second x chromosome. If we treat this as a 100% FST for 1/46th of the genome and a 0% FST for the other 45/46ths of the genome, this averages out to a Male-Female FST of 2.17%.

Apparently even this is enough for things like breasts or differential genitalia. Additionally, biological sex appears to somewhat affect gene expression in chromosomes other than the sex chromosomes [614]. Genes don't just evolve in isolation, they're passed on in sets. Genes sometimes have different effects depending on what other genes they interact with, these are known as non-additive effects. Does this happen with race? Yes, somewhat. The ApoE4 allele confers less risk of Alzheimer's disease in Blacks than in Whites [615 & 616].

For another example, HapK is very rare in Africa, and only present in African-Americans due to European admixture. It carries a modest risk of myocardial infarction for Europeans, but a threefold larger risk for Africans [617].

It has also been demonstrated that race/ethnicity information enhances the ability to understand population-specific genetic architecture [688].

Are The Races Subspecies?

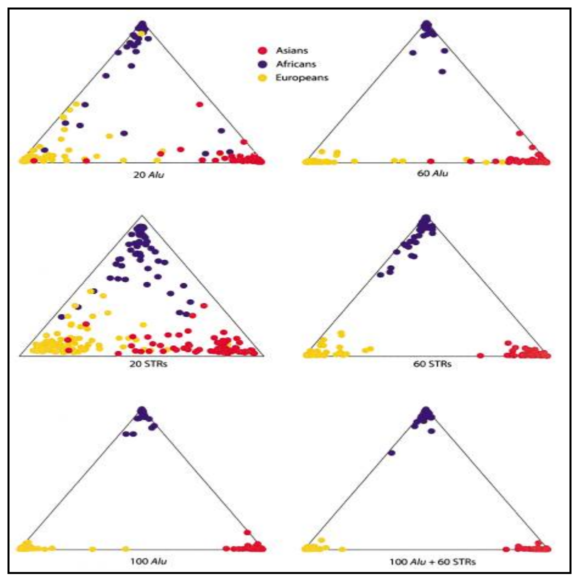
Given all of this, are the races subspecies? Perhaps not. The concept of subspecies is generally not based on genetic measurements like F_{ST}. Moreover, subspecies is a very poorly defined taxonomic rank, which has led some taxonomists to evade it, especially after the famous critique from source 653. Whether or not one would like to call races "subspecies" simply boils down to semantics.

Genetic Clusteredness:

The theoretical demonstration that race can accurately be predicted by SNPs when accounting for binomial probability is experimentally confirmed by studies of best fit genetic clusters. In these, a computer algorithm takes a bunch of people and their genetic data and sorts it into best fit genetic clusters such that within group differences are minimized and between group differences are maximized. From source 581, the correspondence between best fit cluster and geography of origin, by number of SNPs used, is shown in the figure on the right. Which triangle would be shown by somebody who wants to prove that race doesn't exist? The one with the fewest loci of course. This sort of result also works for self identified race/ethnicity (SIRE) [582, 586, 685, 687, 688, & 689], at least for people who self identify as a single race; for things like the one drop rule where somebody who is 7/8ths European is classified as Black because they're 1/8th African, identity doesn't correspond well to genetic clusters. Though to be fair, even Hispanics, (who, in the SouthEastern USA, are on average about 46% Amerindian, 46% White, and 8% Black [623]) cluster as a group much better than expected.

Clustering also works when using a random selection of SNPs [583], when using short tandem repeats rather than SNPs [584], and using methods other than STRUCTURE, PCA, or K-means [686 & 689]. These sorts of clustering results have been replicated further [585, 587, 588, & 589].

Source 581 - Figure 5:



These sorts of results have been around since 1977 when it was shown that simultaneous analysis of multiple blood group loci allowed for clear racial differentiation [684].

In addition to the results from studies of best fit genetic clusters, it has recently been shown that somebody's biogeographic ancestry can successfully be predicted based on the shape of their brain [618].

Clines Or Clusters?

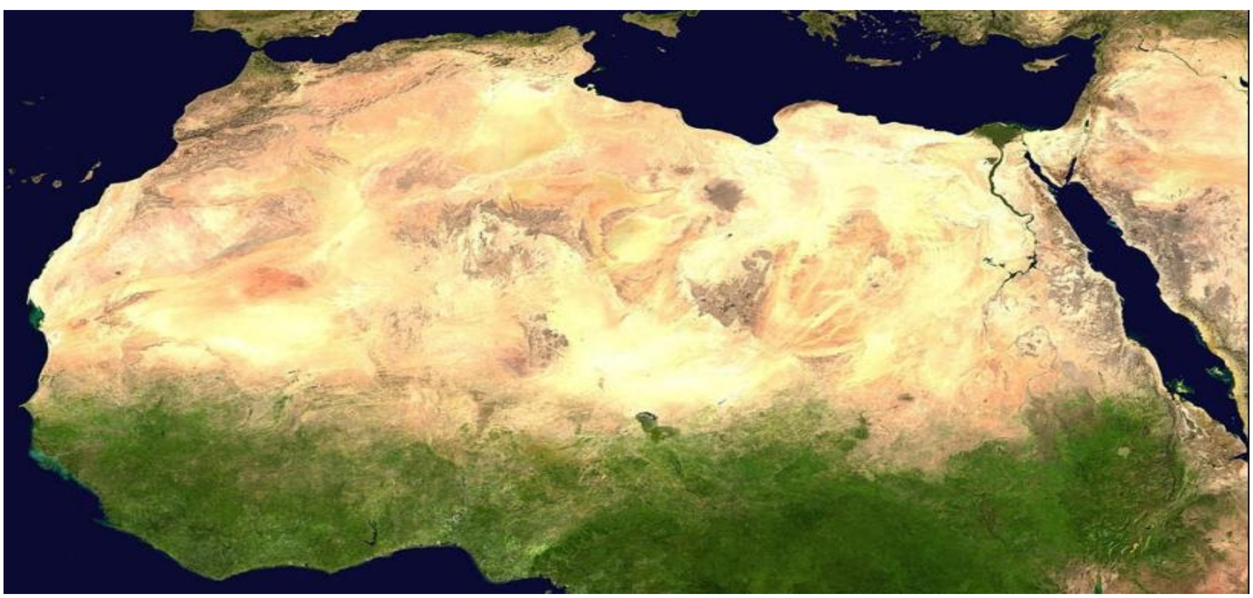
Another argument is that human variation is continuous, and so not discrete along racial boundaries.

Before examining further, let's look at the implications were it to be true. Think of colour, there is no hard boundary between blue and light blue, do blue and light blue not exist? Or think of red and light red, there is no hard boundary between the two, so does red exist? Why call it light red? Isn't it called pink? The Russians think the same thing about our colour scheme, they consider blue and light blue to be different colours. To them, lumping the two into the same name is as strange to us as lumping red and pink into the same name. Contrast to Japan, and to us, their choices are even harder to understand; they consider blue and green to be broadly the same colour. Clearly, colour is a social construct, but does that mean colour doesn't exist? Obviously not.

Or take plains and forests. How many trees need to be planted in plains for it to become a forest? 1? 10? 69,420? There is no hard boundary. When do plains become hills? When do hills become mountains? There are no hard boundaries. Do forests exist? Do mountains exist? Are there no meaningful differences between these social constructs? If race is a social construct and therefore does not exist because there are no hard boundaries, then probably not.

That being said, there is conceptual reason to expect there to be soft boundaries, and evidence that variation is indeed, not perfectly clinal [585].

Geography, oceans, deserts, and mountains could be real practical barriers which are really difficult to cross, which may be crossable if humans really wanted to do it, but would be difficult enough that regular trade would not be frequent. For example, the Saharan desert separates Sub-Saharan Africa from North Africa and keeps Sub-Saharan Africa isolated away from the rest of Humanity:



The Mediterranean Sea separate North Africa from Europe:

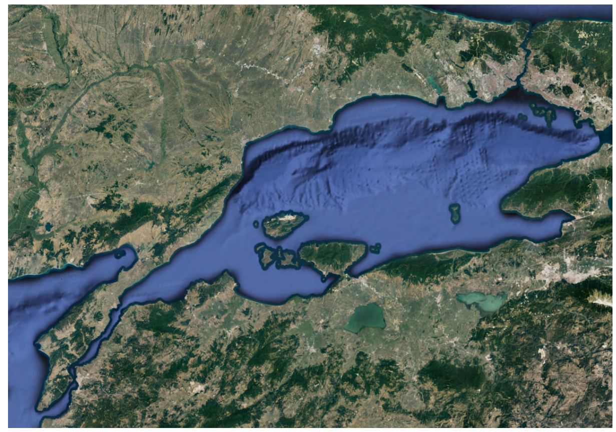
Some suggest that the proximity of the Iberian Peninsula to Morocco, and the Moorish Invasion of Southern Spain should mean substantial North African genetic admixture, but more detailed knowledge of Spanish history makes it unsurprising that there is very little North African admixture in Southern Spain [619]. Furthermore, the Strait Of Gibraltar was more of a barrier than a bridge during prehistoric times [620, 621 & 622].

'The surrender of Granada' (1882). Boabdil, the last Muslim king, surrenders Granada to the Catholic Monarchs

The Caucasus Mountains separate Georgia in the Middle East from Russia in Europe:

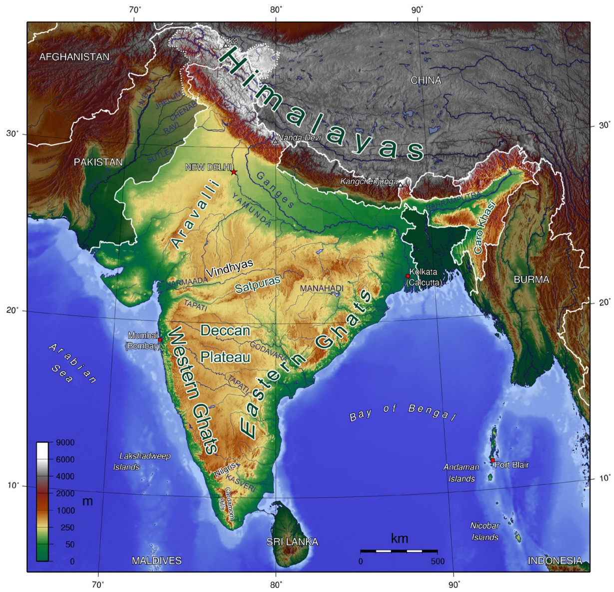
Turkey is a fairly dry and mountainous barrier to migration:

And any immigration through Turkey has to be filtered through either of two narrow straits:

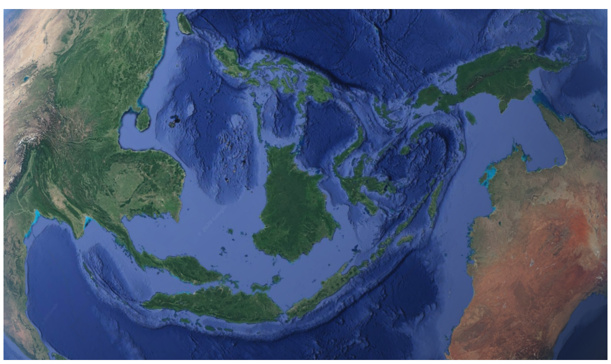


The Ural Mountains also separate European Russia from Asian Russia:

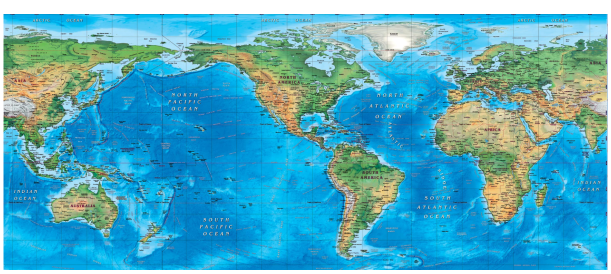
East of India, the Himalayan and Caro Khasi Mountains separate East Asia from India and the Middle East:



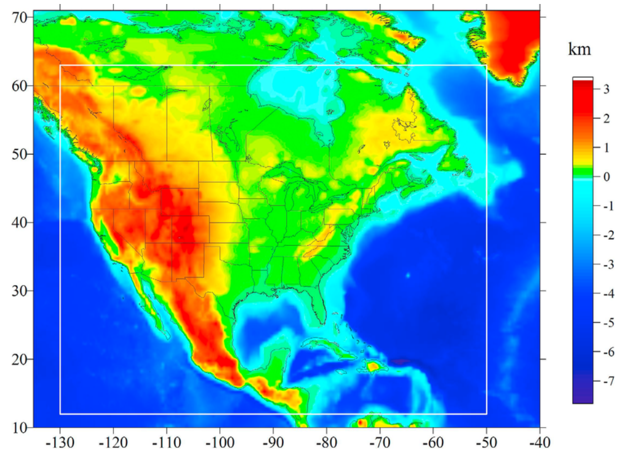
Large bodies of water separate Oceania and Australia from Asia:



Finally, the two largest oceans on the planet separate the Americas from the rest of the world:

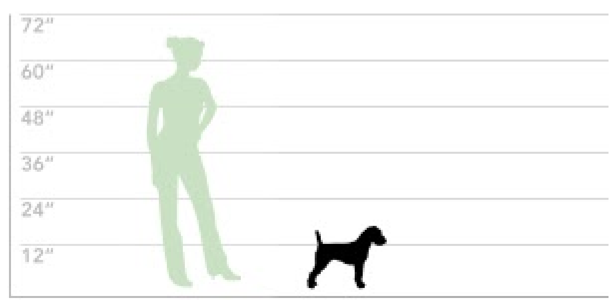


Though even hard physical barriers are not necessarily required to prevent migration and mixing. For example, in North America, there are no mountain ranges which separate North from South:



Yet despite this, adaptation to climate alone is enough to separate Polar Bears from Brown Bears [566]. As a side note, Polar Bears and Brown Bears can breed to produce fertile offspring, does that mean that they aren't distinct enough to be considered different? How about the Parson Russell Terrier, which cannot breed with the Irish Wolfhound despite both being considered to be the same species?

Parson Russell Terrier:



Irish Wolfhound:

Geography and natural selection also aren't the only things which can keep people separate. For example, even after hundreds of years of Whites and Blacks living together in the USA, over 95% of Whites have less than 1% African admixture [590]. People tend to like those of genetic similarity (see [assortative mating]). Additionally, the bias of White women against Black men also increases during the part of the menstrual cycle where sex is most likely to result in a pregnancy [650].

In conclusion, that human genetic variation is not perfectly continuous across racial lines was shown by source 585 which found that two populations of the same race are, on average, more genetically similar than two populations of different races, even when both population pairs are equally far from one another geographically.

Does this mean that human variation is perfectly discrete? No, there is still some migration across even the overwhelming geographic barriers, but these borders still give the human population some structure because migration is less frequent than it is when there are no barriers. Figure 1 of source 959 shows a map of the globe overlaid with where migration is most frequent (blue), and where migration is least frequent (brown). The borders aren't hard boundaries, but they are still important, and they correspond to the racial boundaries:

Source 959 - Figure 1:

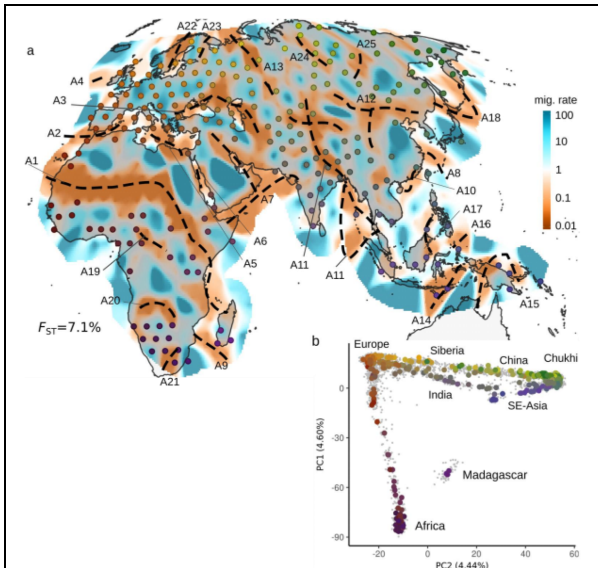
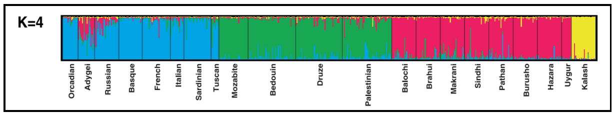


Figure 1: Large-scale patterns of population structure. a: EEMS posterior mean effective migration surface for Afro-Eurasia (AEA) panel. Regions and features discussed in the main text are labeled. Approximate location of troughs are annotated with dashed lines (see Supplemental Figure 2). **b:** PCA plot of AEA panel: Individuals are displayed as grey dots, Colored dots reflect median of sample locations; with colors reflecting geography and matching with the EEMS plot. Locations displayed in the EEMS plot reflect the position of populations after alignment to grid vertices used in the model (see methods). For exact locations, see annotated Supplemental Figure 2 and Table S1. The displayed value of F_{ST} emphasizes the low absolute level of differentiation in human SNP data.

The existence of Whites (Europeans) specifically is shown by source 589 which finds that when splitting all of Eurasia into 4 clusters, Whites become clearly distinguished as being the blue cluster.

Source 589 - Figure 2 - Eurasia - K=4:



For more specifics on genetic distances between specific Human groups, here is a good table from page 64 of source 591:

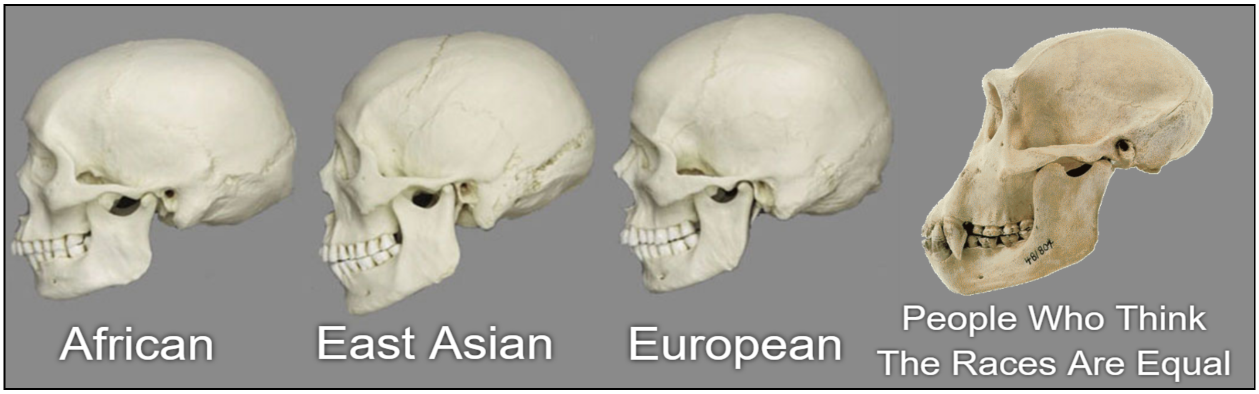
Source 591 - Table 3.1:

	BAN	EAF	WAF	SAN	MBU	IND	IRA	NEA	JPN	KOR	MNK	THA	MNG	MAL	FIL	NTU	SCH	BAS	DAN	ENG	GRK	ITA	CAM	ESK	PLY	AUS
Bantu	0																									
E. African	658	0																								
W. African	188	697	0																							
San	94	776	885	0																						
Mbuti	714	1232	801	1495	0																					
Indian	2202	1078	1748	1246	2663	0																				
Iranian	2241	1060	1796	1267	2588	154	0																			
Near Eastern	1779	709	1454	880	2138	229	158	0																		
Japanese	2361	1345	2252	1905	3089	718	1059	1056	0																	
Korean	2668	1475	1807	1950	2996	681	905	933	137	0																
Mon Khmer	2446	1538	1951	1977	2766	866	1282	987	961	946	0															
Thai	3364	1602	2480	2064	3872	852	1155	1023	743	814	99	0														
Mongol-Tungus	2882	1423	1733	1398	2568	509	681	827	218	170	1093	957	0													
Malaysian	1658	1216	1365	1434	1743	1130	1489	1173	1175	1001	264	455	1251	0												
Filipino	2913	1770	2299	1922	3776	872	908	909	1020	1218	552	625	737	485	0											
N. Turkic	2486	1386	2163	1448	2989	638	821	710	627	732	1259	1225	728	1189	1044	0										
S. Chinese	2963	1664	1958	2231	3384	847	1092	983	541	498	254	105	705	635	315	1109	0									
Basque	1474	922	1299	1307	1965	418	285	246	1481	1063	1831	1726	1049	1784	1634	903	1675	0								
Danish	1708	909	1458	1025	1462	293	179	238	1176	947	1463	1390	680	1628	1279	820	1306	184	0							
English	2288	1163	1487	1197	2373	280	197	236	1244	982	1100	1143	896	1275	1117	866	1152	119	21	0						
Greek	1479	892	1356	1068	1735	272	70	129	1175	904	1482	1355	735	1482	1109	794	1095	231	191	204	0					
Italian	2292	1234	1794	1181	2931	261	133	208	1145	936	1446	1382	905	1599	1136	949	1236	141	72	51	77	0				
C. Amerind	2237	1475	2293	2143	3499	1089	1199	1037	658	790	1522	1323	970	1731	1527	859	1192	1539	1266	1246	1271	1198	0			
Eskimo	3251	2116	2693	2217	3329	940	1234	1225	791	843	1595	1417	545	1617	1597	796	1304	1637	1180	1185	1254	1135	903	0		
Polynesian	2649	1414	1992	1940	3136	927	1142	869	823	890	860	589	969	849	650	1147	508	1406	1210	991	1096	1215	1312	1415	0	
Australian	3272	2131	2694	2705	4287	1176	1546	1408	321	850	1699	1314	781	1665	1300	1580	1081	1949	1400	1534	1498	1413	1360	1230	1145	0
	BAN	EAF	WAF	SAN	MBU	IND	IRA	NEA	JPN	KOR	MNK	THA	MNG	MAL	FIL	NTU	SCH	BAS	DAN	ENG	GRK	ITA	CAM	ESK	PLY	AUS

Some group F_{ST}s are higher than 30%! Some of the lowest F_{ST}s are between Italians, Greeks, the English, and the Danish.

Miscellaneous Differences:

- Illustrating the importance of including multiple races in molecular genetic studies of various traits, source 667 does a GWAS on 15 blood cell traits with a sample of 746,667 participants, including 184,535 non-Europeans, and identified 71 novel genetic associations that aren't present in Europeans.
- The Allele CCR5-Δ32 confers greater resistance to HIV-1, and is present almost exclusively in Europeans [668].
- There is racial variation in humoral and cellular immune responses to measles vaccination [669], as well as racial differences in Pharmacogenomic variants which mediate how individuals respond to medication [687].
- Vitamin D is not an actual vitamin, but rather a hormone produced in the skin during exposure to sunlight, and darker skin reduces absorption capacity [691]. Accordingly, Vitamin D deficiency is present in 1/3 Blacks but only 1/33 Whites [692].
- The ApoE4 allele confers less risk of Alzheimer's disease in Blacks than in Whites [615 & 616].
- Black soldiers are significantly more likely to suffer frostbite injury [675].
- The races differ in traits such as skin colour, hair colour and hair type, the length and density of various bones, muscle composition, etc [677, 678, & 679].
- HapK is very rare in Africa, and only present in African-Americans due to European admixture. It carries a modest risk of myocardial infarction for Europeans, but a threefold larger risk for Africans [617].
- There are population differences in Neanderthal ancestry [636] which has associations with skull shape [671], as well as depression and skin conditions [676].
- Racial groups differ in the rate at which they possess various diseases, including genetic diseases [680, 681, 682, & 690].


African
East Asian
European
People Who Think
The Races Are Equal

7. Race & g

Navigation:

Previous Chapter   Table Of Contents   Next Chapter

Summary:

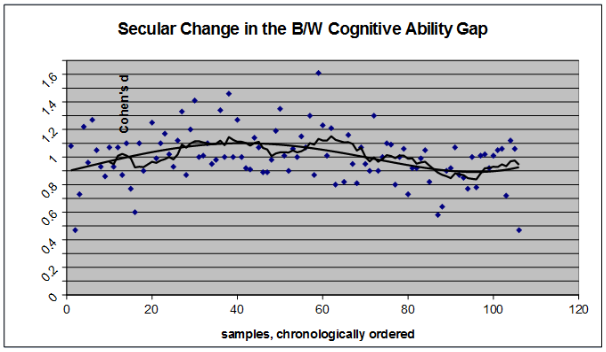
The question of the between-group heritability is like attempting to engineer the tastiest possible cake. We can experiment with the baking to see if cooking at lower temperatures for different times in different ovens has effects, or we can experiment with the effects of different ingredients until we've come up with the optimum cake. With what we've learned, we can then record the ingredients and baking that have gone into the two different cakes to infer the extent to which the differences in ingredients or baking is responsible for the difference in tastiness. Similarly, we can do the same thing to infer the between-group heritability of the Black-White gap in the general intelligence factor (g). This is important because when controlling for IQ, Black-White inequities across a variety of domains either flip in the other direction, equalize, or are substantially reduced [see more here].

The effect size of European ancestry on IQ, taken with the Black-White difference in European ancestry, implies a between-group heritability of 50%-70% [see more here]. However, this is based on a ~13.7 year old sample; we should expect heritability to rise towards ~80% with age [see more here]. The relationship between European ancestry and IQ is also mediated by genetic variants known to influence IQ, and 20%-25% of the Black-White IQ gap can already be naïvely explained by racial differences in polygenic scores derived from the current Genome-Wide Association Studies [see more here]. 30% of the Black-White IQ gap can also be explained by the well established Black-White gap in brain size, which is confirmed to be at least partially genetic in origin [see more here].

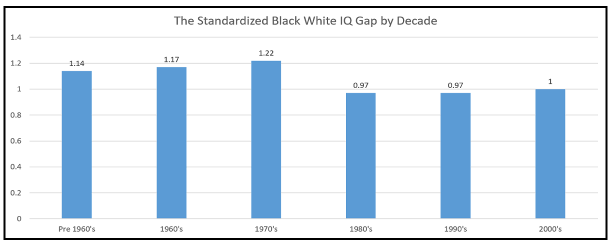
But the Black-White IQ gap isn't completely heritable, right? Surely, given the magnitude of the between-group heritability, we should be able to fix whatever inequalities do exist in order to get rid of 30%-50% of the IQ gap, right? No, not necessarily; "not heritable" doesn't necessarily mean easily/possibly malleable, or even necessarily mean that anybody even knows what environmental variables are etiologically relevant. By my most generous calculations, ~91.69% of the Black-White IQ gap is unexplainable by all of the environmental variables I could think of [see more here], including [nutrition], [lead exposure], [education], [race-unique culture/home-environment], [income], [the Flynn Effect], [racial discrimination], [racial IQ test bias], [stereotype threat], and [x-factors in general]. In order for the between-group heritability of the Black-White IQ gap to be 0%, Blacks must have an environment at the ~0.0111 percentile of White environment [see more here], which is implausible on its face.

Introduction:

There is a well established 1 SD gap in IQ between Blacks and Whites [876] which has remained essentially the same size for as long as it has been recorded [701, 702, & 956]:



X-axis: # of years after 1900; Y-axis: B-W IQ △ (in σ).



However, raw score differences aren't all that important. What matters is the Black-White difference in terms of the general intelligence factor, g. It is a well replicated finding that Black-White differences are larger in terms of more g-loaded tests [see more here], specifically, the Black-White gap in g is ~1.16 Standard Deviations.

This is important because IQ is an absurdly good, causal predictor of life success across a variety of domains [see more here]. Accordingly, when accounting for IQ, a variety of Black-White inequalities are reversed, equalized, or substantially reduced [703, 706, 704, 705, & 666 - ch. 14]. Thus, the question of the cause of the inequalities is the question of the cause of the IQ gap.

- On the validity of race, [see chapter 6].
- On the validity of g/IQ, [see chapter 3].

The Baking:

The Plausibility Of Equality:

Assuming no X-factors and no Scarr-Rowe effects (which will be argued for shortly), we can infer how bad Black environment has to be in order for the between-group heritability of g to be 0% by using the within-group heritability of g and the group differences in g:

- The Black-White difference in g is ~1.16 standard deviations [707 & 708].
- The heritability of IQ, that twin-based heritability means the degree of genetic causality, etc, is dealt with [here]. The direct heritability of g is .91 [493 & 843].

Heritability is akin to an r^2 statistic; 0.91 heritability means that 91% of variance in IQ is explained by genotype. Since r^2 is a squared correlation coefficient [141], the correlation between environment and phenotype (IQ/g) is ~.3, and the causal correlation between genotype and phenotype is ~.954. A correlation of 0.5 means that a 1.0 standard deviation increase in variable-A is associated with a 0.5 standard deviation increase in variable-B. Thus, if we take the ~1.16 standard deviation group difference in g and divide by the ~0.3 correlation coefficient between environment and phenotype, we see that Black environment has to be ~3.867 standard deviations worse than White environment in order for the between-group heritability of g to be 0%. How large is that? If we set White environment to be 0.0, Black environment must have a z-score of -3.867 in order for between-group heritability to be 0%. With this z-score calculator [709], we can see that Blacks must be at the ~0.0111 percentile of White environment to have equal genetic potential for general cognition.

How Bad Is It?

Thus, between-group heritability can be solved via one approach of figuring out how much various environmental variables impact the Black-White gap. Obviously, complete equality with a 3.1 standard deviation difference in environment is implausible on its face, but let's precisely calculate the gap.

First however, we must justify the assumptions of no X-factors and no Scarr-Rowe effects:

- X-factors are dealt with [here].
- Scarr-Rowe effects are dealt with [here].

With analysis of all of the environmental variables that I could think of and find the appropriate evidence for, My most generous estimate, given my criteria, is that controlling for all environmental factors reduces the group difference in g from ~1.16σ to ~1.064σ. If you think that you've thought of an environmental variable that I haven't covered, it may be covered under the section on [X-factors].

The methodology used is straightforward; in order for a variable to be considered a contributor to racial differences, it must:

- Have causal influence on g into adulthood; this means either experimental evidence, or otherwise evidence without genetic confounding. If found, an anti-Jensen effect is also a disqualifier because group differences are on g, and thus an environmental influence must be too. Effects will however be (frankly, charitably) assumed to be on g if no evidence is found for or against Jensen effects.
- Differ in its racial distribution. With the causal effect size for a variable found, and the standardized group difference in a variable found, we can precisely calculate how much of an impact said variable has on the IQ gap.

The reduction in the group difference in g from ~1.16σ to ~1.064σ comes from totalling up the effects of the following environmental variables on the gap:

- [Nutrition]
- [Lead Exposure]
- [Education]
- [Income]
- [The Flynn Effect]

Even though experimental evidence is used, this may be an overestimate of the impact of the environmental variables for two reasons:

1. If two environmental variables covary, the additivity of environmental effects on the gap is dubious; e.g. if controlling for income controls for nutrition, adjustment should be for income, not income + nutrition.
2. [Environment is partially heritable]

In addition, generally relevant is Spearman's hypothesis [see more here].

-Nutrition:

Percent Deficient By Race:

#	Nutrient	White	Black	Hispanic
692	Vitamin A	0.2	0.5	-
692	Vit. B-12	2.2	1.2	1.0
692	Vit. C	7.2	4.3	3.1
692	Vit. D	3.6	31.1	11.3
692	Vit. E	0.6	1.2	-
692	Iron	9.8	4.9	11.4
710	Iodine	25.7	31.2	21.9

While effect sizes of these nutrients are based on randomized placebo control trials (RCTs) where possible, it is possible that some inequities are genetically attenuated because diet is partially heritable [351]. Most environmental explanations would predict that nutritional deficiencies would be highest in Blacks, followed by Hispanics, followed by Whites, but this is clearly not a pattern we consistently see. Contributors to the Black-White gap are narrowed down to Iodine and down to Vitamins A, D, and E; though for Iodine, Hispanics have less deficiency than Whites which is a problem for Environmentalist explanations.

Vitamin D:

Vitamin D is not an actual vitamin, but rather a hormone produced in the skin during exposure to sunlight, and darker skin reduces absorption capacity [691]. Accordingly, Vitamin D deficiency is present in 1/3 Blacks but only 1/33 Whites [692]. If Vitamin D deficiency contributed to the gaps, this would support the Hereditarian view. However, RCTs show that Vitamin D does not impact IQ [712].

Vitamin A:

I was only able to find one RCT [713] for the impact of Vitamin A deficiency on IQ. It assessed the effect of IQ along with two other nutrients. It tested a placebo group, and 7 groups with every combination of nutrient supplementation. There were 4 apples to apples comparisons:

- Male Placebo VS Male Vitamin A: **+6.2**
- Female Placebo VS Female Vitamin A: **-1.5**
- Male Glutamine+Zinc VS Female Vitamin-A + Glutamine + Zinc: **+7.7**
- Female Glutamine+Zinc VS Female Vitamin-A + Glutamine + Zinc: **-2.6**

N-Weighted Average Of Effects: **~+2.75 IQ**

However, this RCT does not inspire much confidence because sample sizes are very low, which may be responsible for the heterogeneity of results; if we had ignored the male sample, we'd say the nutritional deficiency gives Blacks an IQ advantage. This said, if we take the effect size and account for the magnitude of difference in Vitamin A deficiency, **~0.00825** points of the IQ gap is accounted for.

Vitamin E:

An RCT of a sample of over 6,000 women found no effect of Vitamin E supplementation on IQ; review of 3 previous trials also found no effects [711]. Thus, the 0.6% racial gap in Vitamin E deficiency cannot account for any of the Black-White IQ gap.

Iodine:

A meta-analysis of 36 RCTs [714] finds Iodine deficiency decreases g by .53 standard deviations, or 7.95 IQ points. Accounting for the racial difference in Iodine deficiency, **~0.43725** IQ points are accounted for.

Vitamin B:

A review of 14 RCTs [715] on the effect of Vitamin B and folate supplements found no effect on cognitive ability

Vitamin C:

I don't have any RCT evidence for Vitamin C, but a review of cross-sectional and longitudinal data shows that evidence for Vitamin C impacting IQ is very weak [716].

Zinc:

A meta-analysis of 8 RCTs [717] finds no effect of Zinc supplementation on IQ. Sources 718 and 719 reviewed 5 RCTs not reviewed by source 717; three of them found no effect. Source 720 also found evidence of an effect while source 721 did not. Overall, the evidence is not in favor of Zinc impacting IQ. This said, Whites, Blacks, and Hispanics matched for age/sex do not significantly differ in zinc intake [722].

Conclusion:

Totalling up all effect sizes, the most charitable possible estimate of the influence of nutrition on the IQ gap is **0.4455/15ths** of the Black-White IQ gap accounted for. This is certainly an overestimate because in addition to the problems thus discussed, iron also somewhat impacts IQ [723 & 714]. The negative g-loading of effects [850] also raises concern for etiological relevance.

-Lead Exposure:

RCTs on the effect of lead exposure on IQ do not exist because giving people lead poisoning is obviously an unethical research practice. However, longitudinal data for the effect of lead exposure on IQ from 7 studies controlling for potentially confounding variables including race, sex, birth weight, birth order, maternal education, maternal IQ, maternal age, marital status of parents, prenatal smoking status, prenatal alcohol use, and HOME inventory score link Lead Exposure to lower IQ [724]:

Changes in Blood Lead Levels and IQ

Change in BLL (ug/dl)	Predicted Change in IQ
2.4 → 10	-3.9
10 → 20	-1.9
20 → 30	-1.1

This is not experimental, but it's the best we have. One potential concern is that the relationship between race or lead exposure and IQ have different etiologies; the Black-White gap has a g-loading of ~0.5 [see more here] while lead exposure effects have a g-loading of only ~0.1 [725]. A review of 5 national samples from 1988 to 2004 found that Blacks had a mean BLL that was ~1.4 ug/dl higher than Whites [726]. However, this gap has since disappeared [727]. Using the most recent data available, Blacks have a mean BLL about 0.5 ug/dl higher than Whites which is 6.57% as large as a 7.6 ug/dl difference (10 - 2.4). Assuming linearity with the longitudinal data, this should have an effect on IQ about 6.57% as large as -3.9 points, which is 0.26 points. Though assuming non-linearity, even back at the peak of the gap, it would be hard to imagine lead having more than a 1 point effect on the Black-White IQ gap, so to be charitable I'll say that the effect on the gap is **1 point.** However, since there is no evidence for racial Scarr-Rowe effects, we would assume that the totality of all environmental effects affect the gap linearly [see more here].

-Education:

Educational Quantity:

The raw correlation between educational attainment and IQ is partially genetically attenuated [330]. However, the most recent meta-analysis on the experimental effect of an extra year of education on IQ [630] finds an increase of up to 5 IQ points per year, though some possible evidence for the fadeout effect is recorded. Since the meta-analysis' recorded effects are experimental, there is little possibility of genetic attenuation. The other criteria also seems to be met: There is a Black-White gap in years of schooling [728]:

Educational Attainment of the Population Aged 25 and Older by Age, Sex, Race and Hispanic Origin, and Other Selected Characteristics
(Numbers in thousands)

Characteristic	Total	High school graduate or more		Some college or more		Associate's degree or more		Bachelor's degree or more		Advanced degree	
		Percent	Margin of error[1] (±)	Percent	Margin of error[1] (±)	Percent	Margin of error[1] (±)	Percent	Margin of error[1] (±)	Percent	Margin of error[1] (±)
Population 25 and older	**212,132**	**88.4**	**0.3**	**58.9**	**0.5**	**42.3**	**0.5**	**32.5**	**0.5**	**12.0**	**0.3**
Age											
25 to 34	43,006	90.5	0.6	65.0	0.9	46.5	0.9	36.1	1.0	10.9	0.6
35 to 44	39,919	88.7	0.5	62.8	0.9	46.7	1.0	36.3	1.0	13.8	0.7
45 to 64	83,213	89.4	0.4	59.0	0.7	42.6	0.7	32.0	0.7	12.1	0.5
65 and older	45,994	84.3	0.7	49.7	0.9	34.1	0.9	26.7	0.8	11.3	0.7
Sex											
Male	101,888	88.0	0.4	57.6	0.7	41.2	0.7	32.3	0.6	12.0	0.4
Female	110,245	88.8	0.3	60.1	0.6	43.4	0.6	32.7	0.6	12.0	0.4
Race and Hispanic origin											
White alone	168,420	88.8	0.3	59.2	0.6	42.8	0.6	32.8	0.6	12.1	0.3
Non-Hispanic White alone	140,638	93.3	0.3	63.8	0.6	46.9	0.7	36.2	0.7	13.5	0.4
Black alone	25,420	87.0	0.9	52.9	1.4	32.4	1.4	22.5	1.2	8.2	0.7
Asian alone	12,331	89.1	1.2	70.0	1.9	60.4	2.0	53.9	2.0	21.4	1.5
Hispanic (of any race)	31,020	66.7	1.1	36.8	1.0	22.7	0.9	15.5	0.7	4.7	0.4
Nativity Status											
Native born	175,519	91.8	0.3	61.3	0.5	43.3	0.6	32.7	0.6	11.9	0.3
Foreign born	36,613	72.0	1.0	47.6	1.1	37.6	1.1	31.4	1.1	12.5	0.7
Disability Status											
With a disability	28,052	78.6	0.9	41.6	1.2	24.9	1.0	16.7	0.9	5.7	0.5
Without a disability	183,351	89.9	0.3	61.5	0.5	45.0	0.6	34.9	0.5	12.9	0.3

Though we don't know the degree to which the Black-White educational attainment gap in particular is genetically attenuated, we may still expect some environmental effect given that experimental evidence does show educational attainment having a genuine effect. Given this, we may expect that some of the Black-White gap is explained by educational attainment. However, the full 1 standard deviation IQ gap is already present in highschool students and in college applicants before the gap in educational attainment has had time to form [729]:

Source 729 - Table 5:

TABLE 5

Analysis of Black–White Samples for Educational Level

Sample	d	K	N	95% Conf. int.	Observed variance	Sampling error
High School	.95	5	18,104	.86 – 2.05	.0075	.0001
College applicants[1]	.98	13	2,911,312	.95 – .99	.0000	.0000
College students	.69	7	1,953	.55 – .85	.0066	.0034
Graduate school applicants[2]	1.34	10	2,371,255	1.32 – 1.36	.0000	0000
Other graduate applicant samples	1.17	13	11,604	.72 – 1.34	.0097	.0007

How are we to explain these two seemingly contradictory lines of evidence? One solution, as mentioned earlier, is that perhaps the Black-White educational attainment gap is genetically attenuated. Another however is that perhaps the IQ gap and the experimental effect of educational attainment are etiologically different. This appears to be the case; the Black-White IQ gap has a much higher g-loading [see more here] than the effect education [see more here]. In addition to the [evidence] on the g-loading of gains from a year of schooling, we know that IQ gains from cognitive training [276], retesting [275], head start programs [142], adoption [306], and the Flynn Effect [274] are not on g. Are more g-loaded tests more resistant to training gains? Not necessarily, IQ gains decrease the g-loadings of those IQ tests [275 & 416].

Educational Quality:

The first thing we should note is that even if there was Black-White gap in pre-college educational quality, we would not expect this to matter because voucher studies where a random selection of poor kids are sent to prestigious schools to be compared to poor kids who happened to not receive a voucher (thus an apples to apples comparison), find that school quality has barely any effect:

The Cleveland Voucher Program [730]:

Grade:	Voucher:	No Voucher	Non-Applicant
1	555	546	548
2	587	577	580
3	615	605	607
4	632	620	624
5	643	636	636
6	654	639	638

The Milwaukee Voucher Program [731]:

Grade/Subject:	G1 - 2006:	G2 - 2006:	G1 - 2010:	G2 - 2010:
7 - R	432.2	435.3	492.2	485.4
8 - R	446.5	436.9	505.1	486.1
9 - R	458.0	472/9	593.5	492.0
7 - M	388.2	395.7	501.6	500.0
8 - M	426.3	424.4	504.2	493.3
9 - M	462.9	478.7	515.5	524.2

G1: Received Voucher; G2: Denied Voucher; M = Math; R = Reading.

The Washington DC Voucher Program [732]:

Group:	Math:	Reading:
Voucher:	541.00	645.92
Applicant:	543.36	645.24

Voucher given at the beginning of high school, test scores from the end of high school.

This stated, Black students in grade school now receive more funding. Black school districts receive less funding, but the Blacker schools within the Blacker districts get more funding than the Whiter schools in the Blacker districts [874]. Accounting for this, in 1972, Black students received $0.98 for every dollar spent on White students, and in 1982 this trend reversed such that Black students now receive more funding than White students [733]. This result has achieved replication [734]. One more replication [875] comes to the same finding, but interprets it in a bizarre fashion, the authors take issue with the fact that this figure is expressed as a nationwide average, writing:

> "But racial disparities in education spending clearly exist in a host of other states. In Illinois, New York, and Pennsylvania, per pupil expenditures for black and Hispanic students hover around 90 percent of those for white students. This finding is a reflection of these states' regressive funding tendencies, and the fact that people of color tend to be more concentrated in high-poverty districts. The flip side of this disturbing evidence comes from states such as Massachusetts and New Jersey in which high-poverty districts receive greater support from state and local sources than low-poverty districts."

They express dismay at the fact that, in some states, Black children receive 10% less funding than White children, but seem relieved that in others Black children receive as much as 18% more funding than White children. Their language seems to imply a sort of anti-White bias on the part of the authors. In any case, if we are trying to explain why, on average, Black life outcomes differ from White life outcomes, and we are talking about national populations, then average spending per pupil across the nation is obviously the correct statistic to look at.

Source 875 - Table 2:

Per pupil expenditures for each racial group expressed as a percentage of per pupil expenditures for white students, by state

State	Asian	Black	Hispanic	Native American
Alabama	103	97	101	98
Alaska	101	96	100	120
Arizona	98	98	99	104
Arkansas	100	106	99	99
California	94	97	99	109
Colorado	98	103	102	103
Connecticut	100	103	101	102
Delaware	97	99	99	105
Florida	97	99	100	100
Georgia	98	102	100	100
Idaho	100	99	97	101
Illinois	104	93	91	98
Indiana	101	112	108	103
Iowa	99	99	100	100
Kansas	94	95	98	102
Kentucky	99	102	99	99
Louisiana	105	104	103	98
Maryland	106	97	101	97
Massachusetts	108	118	113	108
Michigan	107	107	104	105
Minnesota	104	107	106	116
Mississippi	101	103	99	100
Missouri	106	110	107	101
Montana	97	94	95	100
Nebraska	90	83	93	106
New Hampshire	96	89	83	97
New Jersey	100	117	110	109
New Mexico	94	96	102	86
New York	91	91	90	99
North Carolina	98	100	98	101
North Dakota	99	99	98	98
Ohio	109	111	104	102
Oklahoma	95	97	99	102
Oregon	97	101	99	105
Pennsylvania	96	89	85	98
Rhode Island	100	99	100	105
South Carolina	101	105	101	98
South Dakota	96	93	97	99
Tennessee	100	100	101	98
Texas	90	93	99	100
Utah	95	95	97	117
Vermont	103	101	102	92
Virginia	107	101	105	99
Washington	97	98	98	103
West Virginia	100	99	95	99
Wisconsin	102	98	98	105
Wyoming	98	97	98	97

Turning to more specific measures of school quality, racial differences in class size were non-existent by the early 1970s [735]:

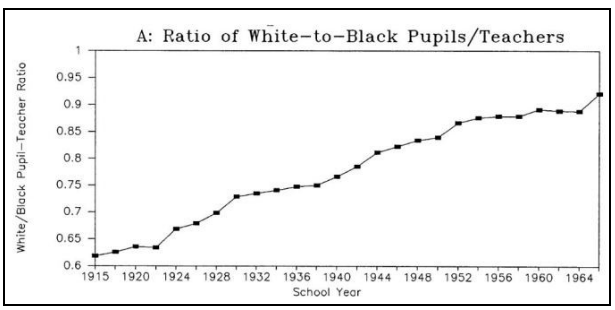
Source 735 - Table 6:

Table 6: Schooling Inputs by Demographic Characteristics 1972-1992

Category	Expenditures/Pupil (1992$) 1972	1982	1992	Pupils/Teacher 1972	1982	1992
By average white and non-white student in the district:						
(1) White	2,856	3,414	4,661	19.32	15.13	13.09
(2) Nonwhite	2,800	3,460	4,796	19.58	14.58	12.52
Ratio (1)/(2)	1.02	0.99	0.97	0.99	1.04	1.05
By median household income in the district:						
1st quartile	2,212	3,040	4,214	19.22	14.24	11.93
2nd quartile	2,388	3,381	4,324	19.24	14.56	12.56
3rd quartile	2,970	3,359	4,686	18.82	15.25	13.20
4th quartile	3,095	3,667	5,047	19.82	15.70	13.53
Ratio (4th)/(1st)	1.40	1.21	1.20	1.03	1.10	1.13
By poverty status:						
(1) Out of poverty	2,881	3,432	4,700	19.34	15.11	13.06
(2) In poverty	2,660	3,331	4,531	19.42	14.81	12.81
Ratio (1)/(2)	1.08	1.03	1.04	1.00	1.02	1.02

In fact, class size differences had been quickly equalizing, even during segregation in the south 1940's [736]:

Source 736 - Figure 1-A:



Class size is of course relevant because it has small to moderate effects on school achievement test scores [877, 878, 879, 880, 881, 882, & 883].

Moreover, Blacker schools have more experienced teachers with more formal education and more pay [735]:

Source 735 - Table 12:

Table 12: Characteristics of Newly Hired Teachers by Race and Income Composition of School Schools and Staffing Survey 1993-94

Percent of School Enrollment that is Black:	All	0-10%	10-50%	50-90%	90+%
N	3,643	2,656	696	181	110
Mean Years of Experience	1.48	1.48	1.49	1.49	1.51
Fraction Certified in Primary Teaching Field	91.4	93.8	88.8	87.3	86.8
Fraction with Bachelors Degree or Higher	99.5	99.4	99.7	99.8	99.7
Fraction with Masters Degree or Higher	16.7	15.6	15.1	26.2	28.4
Fraction Teaching Full-Time	86.0	83.6	88.1	94.7	94.2
Fraction Who Say They Would Teach Again	77.3	81.3	73.1	66.3	60.7
Fraction Who Plan to Exit Teaching as Soon as Possible	2.5	1.6	2.2	8.2	9.1
Fraction Who Plan to Exit Teaching at First Opportunity	14.3	13.1	12.9	27.2	21.7
Mean Academic Base Year Salary	23,083	22,741	23,509	23,943	24,209
Percent of School Enrollment Qualified for Free or Reduced-Price Lunch:	All	0-10%	10-50%	50-90%	90+%
N	3,643	834	1,878	729	202
Mean Years of Experience	-	1.47	1.47	1.49	1.58
Fraction Certified in Primary Teaching Field	-	95.6	93.1	86.7	80.9
Fraction with Bachelors Degree or Higher	-	99.3	99.6	99.5	99.6
Fraction with Masters Degree or Higher	-	22.9	14.3	16.3	14.7
Fraction Teaching Full-Time	-	82.6	84.4	91.1	90.5
Fraction Who Say They Would Teach Again	-	79.9	78.1	74.5	72.5
Fraction Who Plan to Exit Teaching as Soon as Possible	-	1.6	1.5	5.0	3.9
Fraction Who Plan to Exit Teaching at First Opportunity	-	13.1	13.5	17.8	11.2
Mean Academic Base Year Salary	-	24,282	22,331	23,232	24,268

This is not a recent development either; even during segregation in the South, teacher pay equalized in the 1950's [736]:

Source 736 - Figure 1-C:

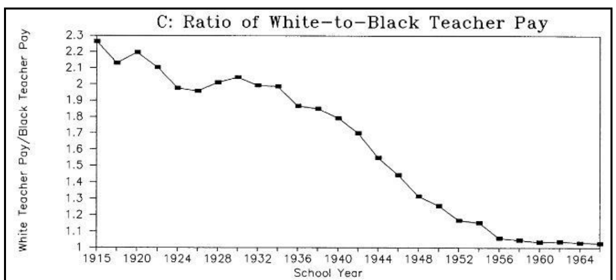


FIGURE I
Relative School Quality in Eighteen Segregated States, 1915–1966

Thus, Black students are advantaged relative to White students in their pre-college education, not that it actually matters.

As for college quality, first it should be noted that the higher incomes of students of better colleges is a result of selective admissions; students who don't go to selective schools despite being good enough to do so make about as much as students who do go:

Source 57:

> "*The difference in R^2's indicated that for men selectivity explained only 0.21% (for women 0.4%) of the total variability in income above and beyond the controls. Note that the zero-order correlation between selectivity and income was only 0.07 for males and 0.11 for females.*"

Source 58:

> "*Holding all student characteristics constant, graduates from private institutions enjoy a slight 4 percent earnings advantage over public college graduates. Moreover, graduates from colleges with selectivity scores 100 points higher than comparison colleges averaged a 1 percent earnings premium.*"

100 points on the selectivity scale that source 58 uses would be most of the operational scale that the paper looked at, with the boundaries between 867 and 1011 being 144 points apart:

Source 58 - Figure 1:

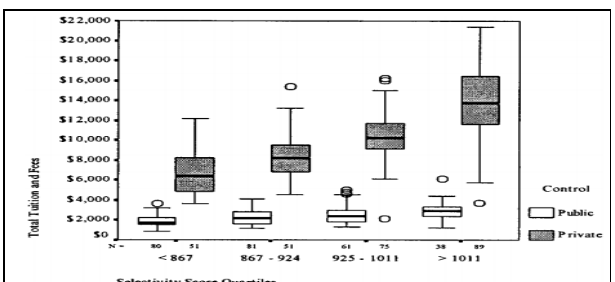


FIG. 1. Institutional selectivity vs. tuition and fees.

Source 59:

> "*After we adjust for students' unobserved characteristics, our findings lead us to question the view that school selectivity, as measured by the average SAT score of the freshmen who attend a. college, is an important determinant of students' subsequent incomes. Students who attended more selective colleges do not earn more than other students who were accepted and rejected by comparable schools but attended less selective colleges*"

This said, there is also significant pro-Black bias in college admissions because of affirmative action. With equal qualifications, Black applicants are roughly 21 times more likely to be admitted into an American college, while Hispanics are 3 times as likely, and Asians are 6% less likely:

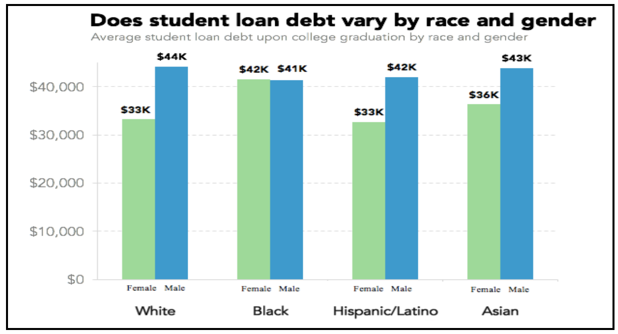
#:	School:	Black	Hispanic	Asian
737	Arizona State (Law)	1115.4	84.95	2.18
737	University of Nebraska (Law)	442.39	89.63	5.78
737	University of Arizona Law	250.03	18.15	2.54
738	University of Virginia (Law)	730.8	1.1	1.86
738	William and Mary (Law)	167.51	2.47	3.29
738	University of Maryland (Medical)	20.63	2.51	0.68
738	George Mason (Law)	1.13	1.09	1.74

Continued:

#:	School:	Black	Hispanic	Asian
[739](#)	William and Mary (Law)	267.0	0.66	0.66
[739](#)	University of Virginia (Undergrad)	106.0	2.81	0.94
[739](#)	North Carolina State (Undergrad)	13.0	1.93	0.64
[740](#)	Berkeley (Law)	121.6	18.2	1.6
[740](#)	UCLA (Undergrad)	5.15	1.92	0.85
[741](#)	University of Michigan	62.79	47.82	0.81
[742](#)	SUNY (Medical)	9.44	4.08	0.76
[742](#)	University of Washington (Medical)	4.01	4.86	0.9
[743](#)	Miami University (Undergrad)	7.99	2.16	2.14
[743](#)	Ohio State (Undergrad)	3.33	4.3	1.47
[744](#)	US Naval Academy	4.44	3.32	0.67
[744](#)	US Military Academy	1.94	1.2	0.68
All	All (Mean)	175.51	15.43	1.59
All	All (Median)	20.63	2.81	0.94

In selective colleges, it is estimated that the proportion of students who are White would increase from 66% to 75% if admissions were based solely on test scores [745]. Thinking about it another way, affirmative action gives Blacks a bonus worth the equivalent of 230 extra SAT points during admissions, Hispanics 185 points, legacies 160 points, and Asians -50 points [652].

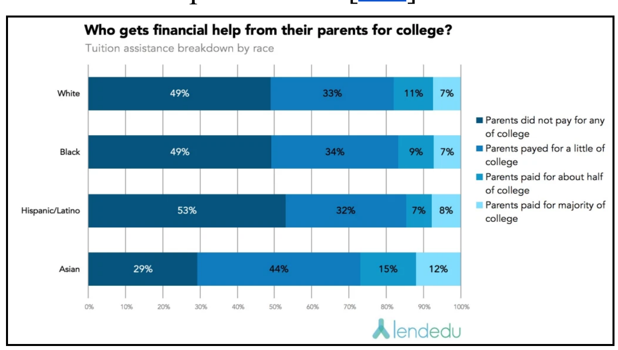
Does college debt disadvantage Blacks? The gap in debt is a function of Whites being more likely to pay it off; there is not really any gap in student loan debt **upon graduation** [746]:



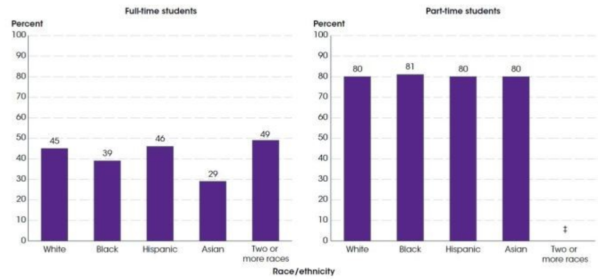
Once minorities get into college, they are given greater access to grants. Specifically, Minority students account for 38% of the student population and 40.4% of grant funding. White students account for 61.8% of all students and 59.3% of grant funding [749]:

Total Grants All Grants Race	Percentage Receiving Grants	Average Grant Amount Received	Total Grant Funding	Number of Grant Recipients	Percentage of Grant Recipients	Percentage of Total Grant Funding	Percentage of Student Population
Total	51.7%	$4,864	$52,646 million	10,822,900	100.0%	100.0%	100.0%
White	48.2%	$5,008	$31,230 million	6,235,700	57.6%	59.3%	61.8%
All Minority Students	57.4%	$4,672	$21,287 million	4,556,300	42.1%	40.4%	38.0%
Black or African-American	63.5%	$4,372	$8,113 million	1,855,800	17.1%	15.4%	14.0%
Hispanic or Latino	58.1%	$4,314	$7,425 million	1,720,900	15.9%	14.1%	14.1%
Asian	43.1%	$6,444	$3,430 million	532,200	4.9%	6.5%	5.9%
American Indian or Alaska Native	60.4%	$4,327	$461 million	106,600	1.0%	0.9%	0.8%
Native Hawaiian or Pacific Islander	49.3%	$4,097	$305 million	74,500	0.7%	0.6%	0.7%
More Than One Race	53.8%	$5,831	$1,553 million	266,300	2.5%	2.9%	2.4%

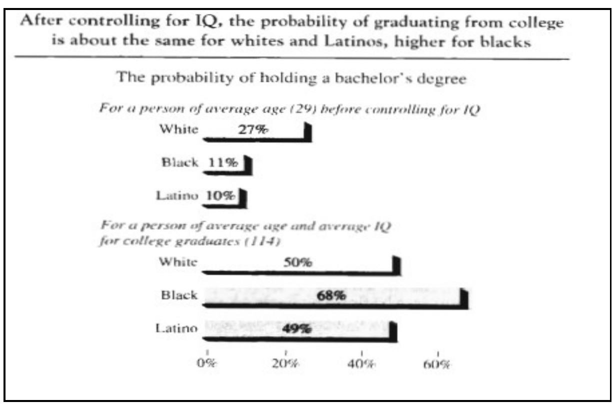
Black, Hispanic, and White students also have similar chances of their parents paying for a significant proportion of their college education while Asians are more likely than others to have parental aid [746]:



A related narrative is that Blacks can't focus as much on education because their poor financial situation means that they have to work to support themselves during college, but Whites are more likely to hold a job during high school and college [750]:



So given all of the financial privileges of Blacks, why are Whites more likely to graduate? Controlling for IQ, Whites and Hispanics are equally likely to graduate from college, and Blacks are more likely to graduate from college [666 - ch. 14 - p.320]:



In sum, the evidence suggests that a gap in educational quality cannot be responsible for any of the Black-White IQ gap. This is true because a gap in educational quality wouldn't matter if it did exist, and because Blacks get access to higher quality schools.

-Income:

Modeling SES as a background variable with SEM, the Black-White IQ gap is reduced from 1.164σ to 0.977σ [197]. However, the genetic correlation between SES and IQ shows this to be a spurious control:

Age	Correlation	% Genetic Mediation	Source #
7	.31	94%	624
12	.32	56%	624
16	.50	50%	417

In fact, guaranteed income experiments, adoption studies, heritability, etc shows income / shared environment to yield no IQ gains:

Source 698:

This guaranteed income experiment on children in North Carolina and Iowa produced no effect on GPA in Iowa and a 6.2% increase in GPA in North Carolina for young children. No effect was found in either state for high schoolers.

Source 696:

This analysis of 16 welfare experiments found that increased income improved teachers' ratings of student performance, but had no effect on test scores.

Source 699:

Differences in family income didn't predict sibling differences in most cognitive abilities with one exception: a $10,000 increase in income did predict a 0.22 SD increase in reading ability.

Heritability:

The general heritability of IQ risis with age to about .8 in adulthood while the influence of shared environment lowers to near zero, in addition to the twin studies this is confirmed by experiments of unrelated children adopted into the same homes.

Adoption & G:

Black-White differences are on g [more here] but IQ gains from adoption are not on g [306].

Source 700:

This guaranteed income experiment on poor Black children increased reading scores by .23 SD and had no effect on GPA for grades 4-6. It had no effect on reading scores and a negative effect on GPA (-.18SD) for grades 7 – 10.

TRAS:

The two largest, best studies in the transracial adoption literature find that Blacks have no lasting IQ gains from being adopted into White homes.

-The Flynn Effect:

There has been a consistent observation of raw IQ test scores rising over time. This was first dubbed the Flynn Effect in the book, *The Bell Curve* [666]. So, should we expect this to push the Black-White IQ gap towards shrinking over time? No, Whites have gained as much from the Flynn Effect as Blacks have [751].

This is already case closed, but moreover, the Black-White IQ gap is etiologically different from the Flynn effect. The Black-White gap is on g [see more here], while the Flynn Effect is not on g [274]. In addition, the Flynn Effect does not achieve measurement equivalence [see more here], while by contrast, the Black-White gap does [see more here].

-Spearman's Hypothesis:

The g-loadings of tests are highly correlated with heritabilities [355, 356, 357, 358, & 359]. If population group differences are greater on the more g-loaded and more heritable subtests, this implies that those differences have a partial genetic origin [663 & 7]. This is a well replicated finding [546 & 7 - pp. 369–379]. This is true even among three-year-olds administered eight subtests of the Stanford–Binet [323]. But this is just the relationship between Black-White gaps and g-loadings; are Black-White gaps larger on the more heritable tests too? Yes [356 & 777]. But this is just correlational, what is the actual Black-White gap in g? In modern day, using SEM/MGCFA, the Black-White difference in g is ~1.16 standard deviations [707 & 708]. Spearman's hypothesis has also been confirmed for differences between Whites and Native Americans [753], for the differences between Whites and Latin-American Hispanics [754], for the gaps between Korea and various other countries [1196], and for the differences between Jews and Whites [755]. Correlations are often stronger when data is more granular, and this is no exception; admixture analysis, with multiple different degrees of ancestry, confirms Spearman's hypothesis much more strongly than usual for both g-loadings and heritabilities [777]. For more on the validity of the method of correlated vectors, [see this].

X-Factors:

Because the environmental factors thus investigated have been investigated to death, many now resort to X-factors as a possible explanation. The basic idea is akin to Lewontin's seed metaphor: If we gave one pot of plant seeds good soil, good lighting, and plenty of water, and another pot poor soil, poor lighting, and meager wager, and we randomly distributed plant seeds between two pots, the differences in growth between the two pots would have a heritability of 0% despite the differences in growth within the two pots having a heritability of 100%.

The basic idea is the possibility of environmental factors that have the property that they are present in Blacks but not in Whites, or are present in Whites but not in Blacks, and that environmental X-factor variables are what contribute to the Black-White IQ gap.

What makes X-factor effects unlikely to exist in general is the well replicated finding of measurement equivalence [see more here]; as explained in source 197, the existence of an X-factor would likely show up as a differential property of Black intelligence. Given that the Black-White IQ gap is on g [see more here], If an X-factor existed, it would have to have all of the exact same psychometric properties as the general factor of intelligence, and it would need to interact with all other factors and items in the exact same way, which is extremely unlikely. Stereotype threat for example, would be a violation of measurement equivalence if it existed [756] (it doesn't; see more here). In the context of race, this would insinuate that Jensen's default hypothesis is correct [see also; 198 - pp. 217-218; 194 - p. 46; 200 - pp. 435; 199; 201 - p 43; 202; & 203 - pp. 3-4].

Also eyebrow raising is the well replicated observation that between-White heritability is the same as between-Black heritability and heritability between-Hispanics [300].

This stated, the following X-factors are the only ones that anybody can ever come up with, and the evidence is against them:

- [Stereotype Threat]
- [Colourism/Racism/Discrimination]
- [Test Bias]
- [Race-Unique Home Environment/Culture]

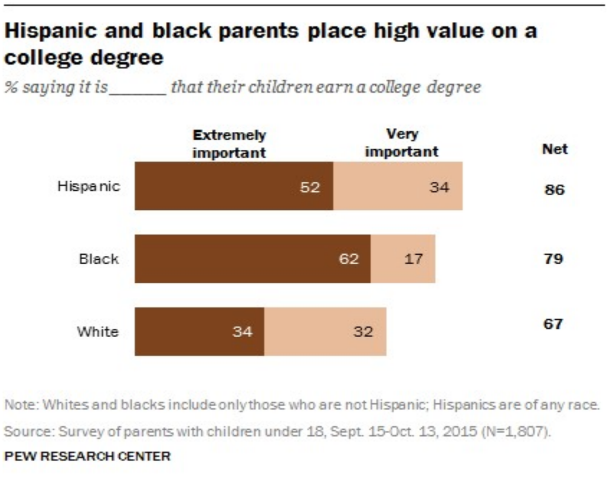
-Race-Unique Home Environment/Culture:

Presumably, if race-unique home environment mattered more than regular home environment, transracial adoption studies would have an effect on Black IQ despite normal adoption data having been shown no effects on g. While the transracial adoption literature isn't very high quality, the best interpretation of it does not seem to indicate that Blacks gain anything from adoption into White homes [see more here].

Also relevant are the various studies examining group-specific developmental theories of cognitive ability near unanimously finding no group-specific developmental variables for IQ [233 - pp. 170-171; 234; 235; 236; 7 - pp. 465-467; & 237].

When culture is invoked, it is oftentimes suggested that Blacks and Hispanics lag behind Whites and Asians because they have cultures that place less value on education. Given the previously established irrelevance of education [see more here] to the Black-White IQ gap, we may be inclined to dismiss it but perhaps such an attitude would generalize to other things that aren't immediately obvious. The problem with this hypothesis is that it isn't clear that such racial differences in culture actually exist. Black parents are more likely than White parents to say that it is important that their child gets a college degree [761]:

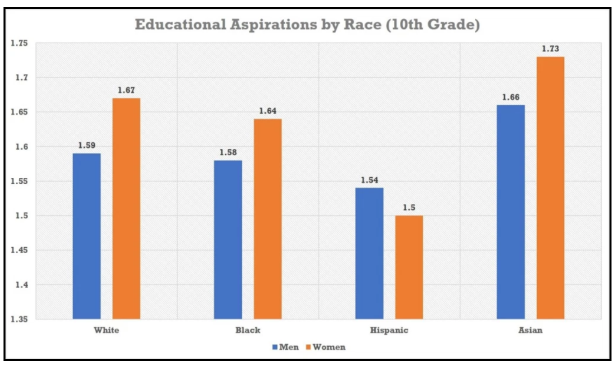
Source 761:



Consistent with this, Black and Hispanic students are also more likely than Whites and Asians to have parents who check to see that their homework is completed [762]:

Race/ethnicity	Average hours spent on homework per week by students who did homework outside of school	Percentage distribution of students who do homework outside of school by how frequently they do homework: Less than once per week	1 to 2 days per week	3 to 4 days per week	5 or more days per week	Percentage of students whose parents[1] check that homework is done
Total	**6.8**	**5.4**	**14.8**	**38.0**	**41.9**	**64.6**
White	6.8	4.2	12.9	38.6	44.3	57.2
Black	6.3	‡	20.1	41.0	29.7	83.1
Hispanic	6.4	5.9	17.7	36.6	39.9	75.6
Asian	10.3	#	13.8 !	18.5 !	67.7	59.0
Native Hawaiian/Pacific Islander	‡	‡	‡	‡	‡	‡
American Indian/Alaska Native	‡	‡	‡	‡	‡	‡
Two or more races	7.1	‡	10.5	32.9	50.5	65.9

There are some differences that favor Whites when you ask students to rate, on a 4 points scale, how far they intend to go in school. But these differences are less than 0.2 SD and so are practically negligible (SD = .49) [763]:



In another survey [764], racial differences on measures of family involvement in school, commitment to school, and family attitude towards education, were consistently found to either be practically insignificant (d<.20) or to favor minorities:

Source 764 - Table 4:

Table 4
Cohen's d-values for Whites versus Minorities Differences

	Motivation			Social Engagement				Self-Regulation		
Race / Ethnicity	***AD***	***CS***	***OPT***	***FA***	***FI***	***RSP***	***SSC***	***MF***	***TBA***	***OC***
Hispanic / Latino	.40	.11	.09	.16	.25	.20	.39	.28	.12	.34
American Indian / Alaska Native	.35	.19	.21	.23	.23	.25	.24	.32	.21	.37
Asian	-.24	-.08	.08	-.12	.16	-.11	-.17	-.25	-.29	-.37
Black / African American	.17	-.09	-.17	-.11	-.07	.16	.35	.47	.05	.44
Two or more races	.13	-.04	.10	-.06	.07	.22	.29	.29	.12	.27

Note. Reference group = White. The positive values indicate Whites score higher. *AD = Academic Discipline, CS = Commitment to School, OPT = Optimism, FA = Family Attitude toward Education, FI = Family Involvement, RSP = Relationships with School Personnel, SSC = School Safety Climate, MF = Managing Feelings, TBA = Thinking before Acting*, and *OC = Orderly Conduct.*

Thus, if we are to define a stereotype as an erroneous belief in a difference between two groups, then stereotypes about large racial differences in the value placed on education appear to be unjustified.

Trans-Race Adoption:

IQ gains from adoption are not on g [306], while Black-White IQ differences are driven by g [see more here]. However, perhaps transracial adoption may be a special case if it can capture race-specific family environment x-factors that aren't present in within-race comparisons.

-The Moore Study:

The only actual adoption study aside from the Minnesota study [765], the Moore study does support the environmentalist view. However, there are two fatal flaws. First, its sample size was tiny. Second, even the Blacks raised in Black homes scored higher on IQ tests than Whites typically do in the general population. Thus, the sample was not only small but also unrepresentative. Moore was studying a sample of Blacks in which there was no Black-White IQ gap to begin with. We don't know if gains are g, and we may also expect a follow up to look like the Minnesota Study.

-The Minnesota Study:

The Minnesota Transracial Adoption Study was set up to conclusively show that the Black-White IQ gap was not due to genes, with the authors studying White, Black, Mixed, and Asian/Indian children adopted into the families of White parents who had above average IQs and SES. It is better than the other transracial adoption studies because it has the largest sample size of 426, and because it is the only transracial adoption study to do a follow up later in life. Before the later follow up, there was an original writeup in 1976 when the children were 7 years old [766], at which point the authors concluded that their data supported an environmentalist position since the higher than average adoptive parent IQ & SES contributed to improvements across the board; Blacks were brought up to an IQ of 96.8, and the Mixed were brought up to an IQ of 109, which is above the White average. The Asian/Indian subjects were a small sample which is to be ignored.

Next, the same sample was retested at age 17. The new 1992 results [767] caused quite a controversy. Attrition substantially affected the White group, but not the other groups, here are the results after adjustment for attrition:

Children:	1975 IQ:	1986 IQ:
Black Adopted	95.4	89.4
Mixed Adopted	109.5	98.5
White Adopted	117.6	105.6
Biological Children of Adoptive Parents	116.4	109.4

At age 17 after correction for longitudinal attrition, Whites scored exactly the usual 15 points higher than Blacks, and the Mixed scored a point higher than the Hereditarian prediction of the Black-White average. Many may think the result of 89 for black IQ is still evidence that while smaller than initially thought, there were still some gains. However, as pointed out by Lynn [768], 89 is the average for Blacks from this area of the country. The fact that adoptees changed to resemble the general population with age, as well as some other racial IQ data on age effects [769], is in line with the [Wilson Effect]; given a genetic origin of the racial IQ differences, since the heritability of IQ increases with age, the racial gaps should come closer to resembling the general population with age.

Another great result was that some adoptive parents knew that the mixed children were mixed, while some adoptive parents thought that their mixed children were fully Black. Both groups of mixed children scored the same.

As an environmentalist defense in the 1992 writeup, the Minnesota Study authors point out that age at adoption is weakly related to the IQ gap, and that Blacks had later ages of adoption than Mulattoes who had later ages of adoption than Whites. However, this cannot account for more than 17% of variance [770].

It's also telling that the authors didn't think that the environmental differences between the groups were enough to matter until the second writeup when they stopped getting the results that they wanted.

One of the authors, Sandra Scarr, has also admitted to not being entirely forthright about the study in her tribute to Arthur Jensen [800]:

> *"My colleagues and I reported the data accurately and as fully as possible, and then tried to make the results palatable to environmentally committed colleagues. In retrospect, this was a mistake. The results of the transracial adoption study can be used to support either a genetic difference hypothesis or an environmental difference one (because the children have visible African ancestry). We should have been agnostic on the conclusions; Art would have been."*

Since the results of the Minnesota Study, Sandra Scarr has retired from her career in psychology and become a coffee farmer in Hawaii [771].

-The Eyferth Study:

While not technically an adoption study, the Eyferth Study is often brought up when talking about transracial adoption because it has the second largest sample size and because it's sort of like adoption since all children were raised by White mothers. The Eyferth Study in 1961 [772] collected the IQ scores of 181 children born of the bastard children of US soldiers who mated with German women following WWII. Some were half Black, some were full White. This is sort of like an adoption study, although perhaps better since the children belonged to the mothers from birth. Here are the results:

White Male	White Female	Mixed Male	Mixed Female
101	93	97	96

Is this against Hereditarian predictions? No, there was actually an IQ standard for getting into the military at the time, and because of the IQ gap, the bottom 30% of Blacks were rejected from the military while only the bottom 3% of Whites were rejected [773]. IQ is on a normal distribution, and doing the math for a truncated Gaussian on the page below, we would expect the IQ of the White fathers to average 102 and the IQ of the Black fathers in the Eyferth study to average 92.452:

<u>Finding the mean of a truncated Gaussian</u>

$$Gaussian: \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

to find the mean of a probability distribution, integrate over $xP(x)$

$$\frac{1}{f\sqrt{2\pi\sigma^2}}\int_a^b xe^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

where 'f' is a normalization factor (we are integrating only a piece of the distribution)

to make it easier on ourselves, we can shift our equation over so that $x-\mu => x$

$$mean = \frac{1}{f\sqrt{2\pi\sigma^2}}\int_{a-\mu}^{b-\mu} xe^{-\frac{x^2}{2\sigma^2}}dx + \mu$$

now do u-substitution

$$u = -\frac{x^2}{2\sigma^2}$$
$$du = -\frac{x}{\sigma^2}dx$$

substituting this in

$$mean = -\frac{\sigma^2}{f\sqrt{2\pi\sigma^2}}\int_{u_1}^{u_2} e^u du + \mu$$

integrating

$$= -\frac{\sigma^2 e^u}{f\sqrt{2\pi\sigma^2}}\Big|_{u_1}^{u_2} + \mu$$

substituting back to the original units

$$mean = -\frac{\sigma e^{-\frac{x^2}{2\sigma^2}}}{f\sqrt{2\pi}}\Big|_{a-\mu}^{b-\mu} + \mu$$

now we are in the position to solve.
Plugging in $b=\infty$, $\mu=85$, $\sigma=15$, $a=77.14$, $f=0.70$

$$mean = -\frac{15e^{-\frac{x^2}{2*15^2}}}{.7\sqrt{2\pi}}\Big|_{a-85}^{\infty-85} + 85$$
$$= \frac{15}{.7\sqrt{2\pi}}(e^{-\frac{(77.14-85)^2}{2*15^2}} - e^{-\frac{(\infty-85)^2}{2*15^2}}) + 85$$

in the end

$$= \frac{15}{.7\sqrt{2\pi}}e^{-\frac{7.86^2}{2*15^2}} + 85$$
$$= 92.452$$

With the White mothers expected to average 100, the Hereditarian prediction should be for the IQ scores of the children to average in between that of their parents. Here are the actual results in comparison to the Hereditarian predictions:

White Male Actual	White Female Actual	Mixed Male Actual	Mixed Female Actual
101	93	97	96
White Male Predicted	White Female Predicted	Mixed Male Predicted	Mixed Female Predicted
101	101	96.226	96.226

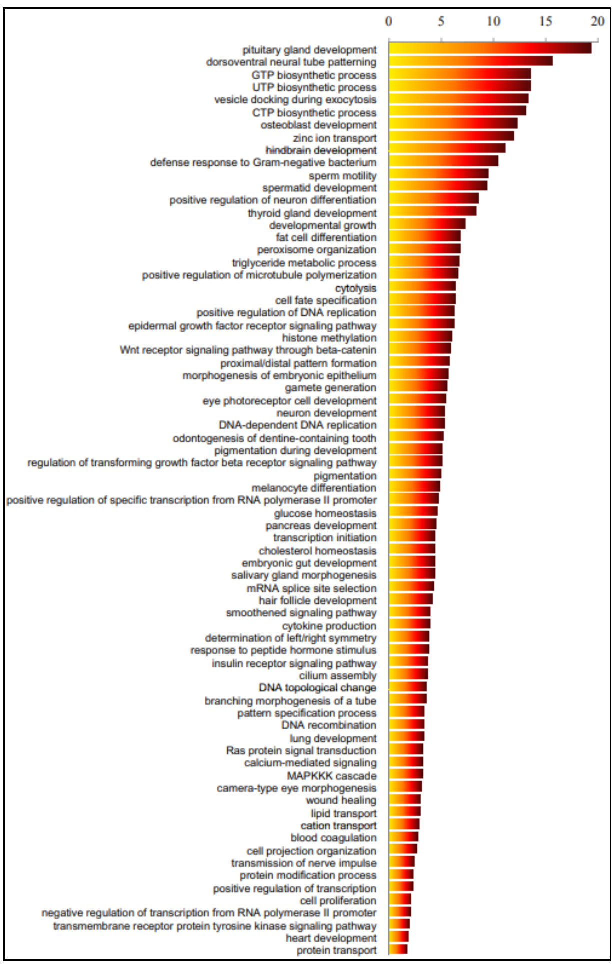
Results match hereditarian predictions near perfectly aside from the White female results, which we are unrepresentative.
For a review of the rest of the transracial adoption literature, see source 774.

[The Ingredients:]

Since the baking and preparation of our cakes doesn't seem to matter much, how about the ingredients? Going in, we would expect so for two reasons:

1. There is a well established Black-White difference in brain size, with an at least somewhat genetic origin, which accounts for 30% of the IQ gap [see more here].
2. Racial differences in terms of genes involved in brain function are larger than the racial differences in terms of genes involved in physical traits like skin colour or hair texture [610].

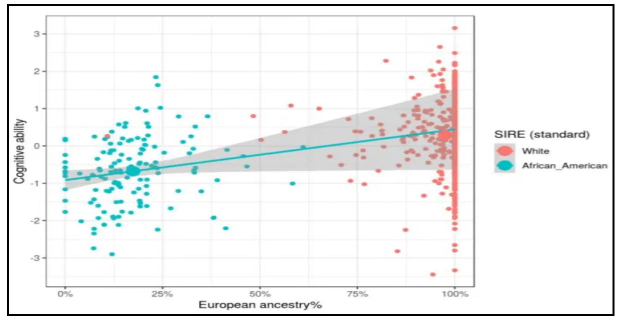
Source 610 - Figure 1:



λ values of GO categories in biological processes enriched for higher F_{ST} SNPs with P-value lower than 10^{-10}

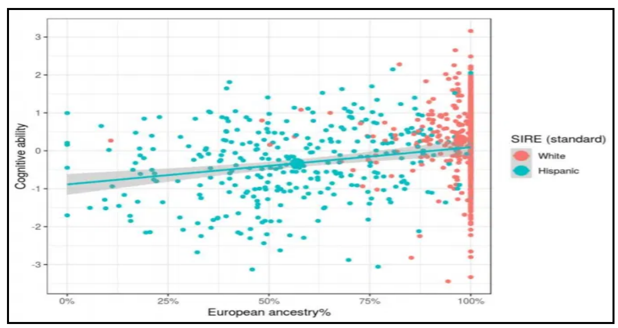
However, that just shows racial differences in terms of brain genes, not what effects those differences have on the Black-White IQ gap. Modern admixtures studies show there to be an association between molecularly assessed European admixture and IQ [752 & 777]:

Source 752 - Figure 3:

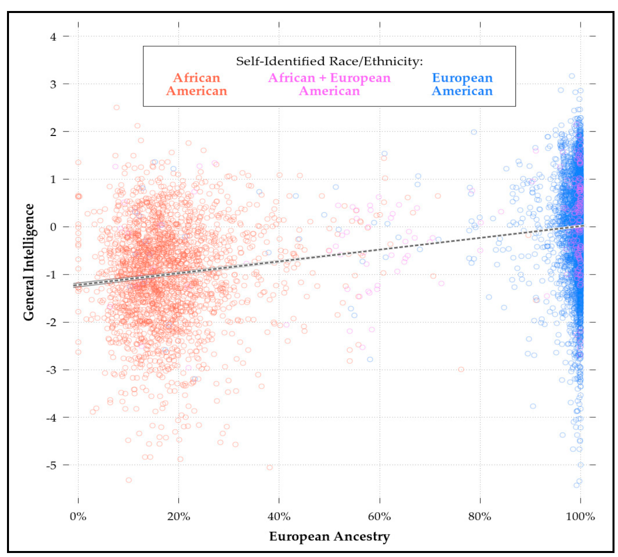


The same is found to be true of Hispanics:

Source 752 - Figure 4:



Source 777 - Figure 3:



Similarly to the analysis of known environmental variables [see more here], we can take the effect size for European Ancestry and the Black-White gap in European Ancestry, and solve for how many IQ points is accounted for by genetic ancestry. Here is the necessary information from source 777 (The following figures for ancestral makeup and deviations in ancestry are also largely consistent with source 799):

- The Black sample was, on average, 18.7% European in ancestry.
- The White sample was, on average, 98.6% European in ancestry.
- For the Black sample, 11.7% percentage points of European ancestry is 1 standard deviation of European ancestry.
- The effect size for European ancestry on IQ is r = 0.086.

This means that a 1 standard deviation increase in European ancestry in Blacks is associated with a 0.086 standard deviation increase in IQ for Blacks. The difference in ancestry between Blacks and Whites is 98.6% minus 18.7% equals 79.9, divided by 11.7, equals a ~6.83 standard deviation difference in European ancestry. 6.83 multiplied by the effect size 0.086 equals ~0.587 standard deviations of the IQ gap accounted for by the gap in European ancestry. The Black-White IQ gap in this sample was 14.72 points, so European ancestry accounts for ~60% of the Black-White IQ gap in this sample. The paper thus concludes that depending on the model, the between-group heritability of the Black-White IQ gap is **50%-70%** [777]. In addition, the sample was also only ~13.7 years old, so we should expect heritability to rise with age [see more here].

The same sort of thing has been found looking at population level data on IQ and the degree of European, African, and Native American admixtures in municipalities of South American countries [788, 789, & 790].

Classic racial phenotypes like skin colour, skull size, and nasal index (the ratio of nose width to nose length) have also been shown to be strong correlates of national IQ variation. This is true even when only comparing African nations:

Phenotype:	r:	Regions:	#:
Skin Colour	-.91	129 Nations	791
Skin Colour	-.86	143 Nations	792
Cranial Capacity	.77	143 Nations	792
Nasal Index	-.6	128 Nations	793
Nasal Index	-.58	48 African Nations	793
Nasal Index	-.16	80 Eurasian Nations	793

These sorts of phenotypic associations have been consistently found for decades as well. In America, there have been studies going back to the 1920s which looked at the correlation between IQ and racial phenotypes like skin colour or nose width among Blacks. Modest positive correlations between proxies for European ancestry and IQ are consistently produced [775 - pp. 546-563; 782]. Some more replications since the olden days include sources 778, 779, 228, 780, & 776. There were two notable exceptions. First, Strong (1913), as reported by Shuey [775], found that rates of mental retardation were similar in dark and light skinned Blacks in a sample of 122 Black Americans. This study is hard to interpret because it looks at the far left tail of the IQ distribution rather than the mean. Whites have a larger standard deviation in IQ than do Blacks and so lighter skinned Blacks may have a larger SD than do dark skinned Blacks. This in turn would lead them to be over-represented among those at both extremes of the IQ distribution relative to dark skinned Blacks. In any case, this is a single study and does little to change the weight of the totality of evidence in this literature. The second study worth mentioning [783] looked at the ancestry of 63 smart Black kids (IQs > 120), as reported by their parents, and found that their degree of White ancestry was lesser than that of a national comparison group. It was also found that the smartest subset of this group (IQs > 140) did not have more White ancestry than the rest of the group. However, as it turns out, the comparison group used by Witty and Jenkins [784] was, itself, an elite sample of Blacks that had higher than average White ancestry [785] which invalidates the whole study design. In any case, it's a single study with a sample size small enough that it, like Strong (1913), doesn't do much to change the total weight of the evidence.

One well known study in this literature [794] found that racial ancestry, as measured via blood analysis, did not correlate with IQ in a sample of 144 Blacks once SES status and skin colour were held constant, both of these however are, in a non-molecular analysis, genetically confounded variables that a sample of 144 would not be expected to survive. Additionally, blood group analysis is a very crude measure of racial ancestry.

-Colourism/Racism/Discrimination:

A direct response to the modern admixture work is the colourism hypothesis; that darker Blacks are slightly more discriminated against than lighter Blacks, and that this is responsible for the correlations between skin colour, European ancestry, and IQ.

This is falsified because molecularly measured ancestry is a better predictor of IQ than both self-identified race/ethnicity (SIRE) and skin colour [777]:

Source 777 - Table 3:

Table 3. Pairwise correlations among African-, African-European, and European-Americans.

	Cognitive Ability	SES	Euro. Ancestry	Afr. Ancestry	SIRE EA	SIRE AA	Color	EduPGS
Cognitive Ability	1							
SES	0.406 (7253)	1						
European Ancestry	0.411 (7321)	0.412 (7319)	1					
African Ancestry	−0.411 (7321)	−0.412 (7319)	−1.000 (7399)	1				
SIRE EA	0.408 (7321)	0.413 (7319)	0.964 (7399)	−0.964 (7399)	1			
SIRE AA	−0.387 (7321)	−0.395 (7319)	−0.928 (7399)	0.928 (7399)	-0.930 (7399)	1		
Color	−0.359 (5534)	−0.354 (5530)	−0.875 (5585)	0.875 (5585)	−0.838 (5585)	0.812 (5585)	1	
EduPGS	0.402 (7321)	0.445 (7319)	0.672 (7399)	−0.672 (7399)	0.645 (7399)	−0.630 (7399)	−0.614 (7399)	1

Note: All values significant at $p < 0.0001$. Pairwise N in parentheses. SES = socioeconomic status, SIRE = self-identified race/ethnicity, eduPGS = education polygenic score.

Accordingly, regression analysis shows ancestry to continue to predict IQ at p<0.001 when controlling for skin colour (model 2):

Source 777 - Table 5:

Table 5. Regression analysis for European ancestry as a predictor of g among monoracial African-Americans with controls for skin color (Model 2), and SES (Model 3) added.

	Model 1			Model 1b			Model 2			Model 3		
Predictor	*B*	SE	β	*B*	SE	β	*B*	SE	β	*B*	SE	β
Intercept	−1.16	0.04	−1.01	−0.82	0.06	−0.997	−1.05	0.09	−0.996	−0.83	0.09	−0.995
EUR	0.78 ***	0.19	0.09 ***				0.83 ***	0.24	0.10 ***	0.67 **	0.23	0.08 **
Skin Color				−0.13 ***	0.04	−0.09 ***	−0.07	0.04	−0.05	−0.06	0.04	−0.04
SES										0.36 ***	0.03	0.28 ***
Adjusted R^2	0.007			0.006			0.013			0.082		
N	2179			1526			1526			1500		

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Model 1b shows the results with color as an alternative predictor. EUR = European ancestry. SES = socioeconomic status.

The same is also shown to be true of other visual ancestry markers like eye colour and hair colour.

These results testing the colorism hypothesis are also replicated in source 752.

The experimentum crucis of an admixture study is the siblings fixed-effects design. The idea is that since full siblings have the same amount of African ancestry, the existence of a between-sibling correlation determines the existence of colorism. IQ correlates with skin colour across-families but not across-siblings; therefore, skin colour correlates with IQ because it's a proxy for ancestry [228].

Similarly, in the Minnesota Transracial Adoption Study [more here], there were two samples of mixed race children; in one, the parents believed the children to be fully Black while the others knew their kids' ancestries; the two groups ended up equal in IQ.

Additionally relevant is the discussion of any sort of racial difference in neuroanatomy. If European ancestry is related to IQ because of its effects on brain variables rather than because of its effects on physical appearance, then this also falsifies colorism. There is a well established finding of a Black-White gap in brain size, which is of at least partially genetic in origin, and which explains ~30% of the Black-White IQ gap [see more here]; we also know of racial differences in a few other neuroanatomical traits.

This is enough to lay the issue to rest, but there are also a few other predictions that a colourism model would make which have been falsified. Colorism may not actually be an X-factor; James Flynn has noted that the colourism hypothesis is intellectually lazy [757, p.60], writing that,

> *"But this is simply an escape from hard thinking and hard research. Racism is not some magic force that operates without a chain of causality. Racism harms people because of its effects and when we list those effects, lack of confidence, low self-image, emasculation of the male, the welfare mother home, poverty, it seems absurd to claim that any one of them does not vary significantly within both black and white America."*

So, if we are to accept the relevance of colourism, we should be able to see Black-White differences in self-esteem, positive affect, suicide rates, etc. However, the opposite is observed:

Source 758:

This meta-analysis of 354 studies on racial differences in self-esteem finds that Blacks are 0.19 standard deviations higher than Whites in self-esteem. This has been the case for the past 50 years.

Source 840:

In this U.S. nationally representative sample of 38,891, Blacks self reported being less stressed than Whites did.

Source 759:

In this nationally representative sample, Whites are .28σ higher in risk for a panic disorder, .28σ higher in risk for generalized anxiety disorder, .12σ higher in social phobia, and had the exact same rate of PTSD.

Source 760:

In this nationally representative sample of 15-40 year olds, Whites scored .27σ higher than Blacks in major depressive disorder.

Source 786:

In this sample of 11 private, non-profit healthcare organizations constituting the Mental Health Research Network, with a combined 7,523,956, replicates these results finding Whites to universally have more psychological disorders than minorities, aside from Blacks being more likely to have schizophrenia disorders and miscellaneous disorders:

Reproduced from source 786 - Table 2:

Disorder	Asian	Black	Hispanic	Mixed	Native Amer. & Alaska Native	Hawaiian/Pacific Islander
Anxiety disorder	0.43	0.65	0.83	0.68	1.09	0.47
Any psychiatric diagnosis	0.36	0.69	0.72	0.64	1.03	0.47
Bipolar disorder	0.24	0.65	0.44	0.65	1.34	0.33
Depressive disorder	0.32	0.68	0.70	0.66	0.99*	0.46
Schizophrenia spectrum disorder	0.77	1.98	0.72	0.88*	1.18*	0.67
Other psychosis	0.50	1.13	0.61	0.34	0.80	0.51

Odds ratios of mental disorders by US racial groups, compared to the White prevalence scaled as 1.00. * indicated statistical insignificance, all other values differed with $p<.001$.

So are Whites disadvantaged in regards to this? No, stress does not causally impact IQ or academic achievement [852, 853, 854, 855, 856, 857, 858, 859, & 860].

Closely related to the idea of stress, self-esteem, or positive affect having an effect on the Black-White IQ gap is the idea of stereotype threat. The idea of stereotype threat is that it occurs in a situation in which it is plausible that some members of a social group may exhibit behavior which is typical of a stereotype about their respective group. It is thought that belief in one's groups' stereotypes induces feelings of threat that cause the stereotypes to become a self-fulfilling prophecy, and that stereotype threat effects partially contribute to long standing racial and gender gaps in academic performance, intelligence, etc. It is thought that these effects can be tested with so-called "primes" in tests. For an example, let's say two groups are given a test, and for one group the start of their test says that racial groups consistently perform equally on the test, while the control group gets no such prime, or perhaps the prime says that some group performs worse. If the prime group and the control group have different performances, this is supposed to be evidence for stereotype threat.

Or at least that's the theory. Taken together, the body of evidence pertaining to stereotype threat does not support its existence whatsoever [see more here].

Given a colourism model, robust, replicable evidence of pro-Black discrimination [478, more here] would also have to be ignored, or conveniently be unrelated to the Black-White IQ gap. The mere stability of the gap [see more here] is also not predicted by the fall of racism, Jim Crow, etc.

-We Found (Some Of) The Genes:

With Genome-Wide Association Studies (GWAS), researchers straightforwardly record the correlation between having certain gene variants and having more of a certain trait. When recording which of the discovered variants a given individual has, researchers can count how many variants predict x rather than y, weigh by effect size, and the result is a polygenic score. Polygenic scores for educational attainment correlate with IQ and racially differ in distribution [777]:

Source 777 - Figure 4:

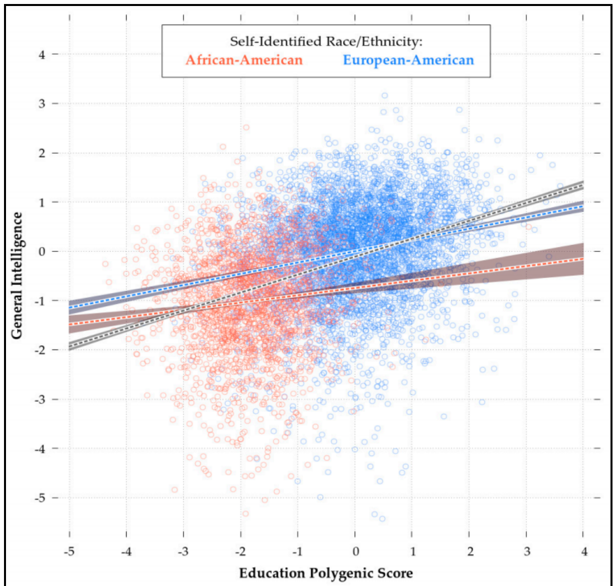


Figure 4. Regression plot for the predictive validity of MTAG 10k eduPGS with Respect to g in the African-American (Red; r = 0.112) and European-American (Blue; r = 0.227) Samples.

Polygenic scores were more predictive of general intelligence for Europeans (r = .227) than for Africans (r = .112), but controlling for the differential validity, the pattern remains. 20%-25% of the Black-White IQ gap can be naïvely explained by polygenic scores. This is important for two reasons:

1. It sets a minimum heritability.
2. The polygenic score evidence is relevant to another potential bias in admixture analysis; nonrandom mating:

There is a well established finding that people tend to select mates who are similar to themselves across a variety of traits, including psychological [see more here]. So, if the Whites who breed with Blacks are a non-random sample of Whites, who bred with Blacks because their polygenic IQ scores are lower than average Whites, then perhaps the ancestry correlation within mixed-race individuals is confounded by assortative mating. Of course, we would expect such a bias to be mirrored and canceled by an inverse assortative mating bias: the Blacks who breed with Whites should have polygenic scores that are higher than average Blacks; indeed, the evidence on assortative mating supports genetic similarity theory [see more here], meaning that assortative mating happens because we are after mates who are similar to ourselves on a genetic level. This is confirmed by molecular genetic evidence, but aside from this, we also know this because assortative mating effects are stronger on the psychological traits which are more heritable. We can also directly calculate the heritability of an individual's choice in friends (21%) and spouses (31%). If we are to expect effects on the admixture analysis based on this, we would expect opposing, cancelling forces. Theory aside, we know that nonrandom mating does not explain the admixture association because racially pure Blacks have similarly low polygenic scores [777].

Source 787 replicates the finding that the races differ in polygenic IQ scores, and responds to criticism by showing that controlling for general ancestry, only using gene variants common in all populations, and excluding recent mutations, all fails to eliminate the polygenic gap. It is also shown that there are large racial differences in polygenic scores when using polygenic scores constructed via within-family effect sizes, and that racial differences are larger when SNPs that have directionally different effects across races are removed. The paper thus provides significant evidence against the idea that various forms of population related bias in GWAS studies can account for the racial polygenic score gap and so strengthens the case for hereditarianism.

Finally, using variants derived from the supplementary data [749] of source 748, and population frequencies derived from the 1000 genomes project [747], there are over 200 variants that are at least 100% more common in Europeans than in Africans which increase intelligence with genome-wide statistical significance and are known to influence genes linked to the central nervous system:

rs159428 (T)	Europeans are 4192% more likely to have this allele than Africans	rs3858281 (A)	Europeans are 204% more likely to have this allele than Africans	rs9607805 (T)	Europeans are 136% more likely to have this allele than Africans
rs215603 (T)	Europeans are 113% more likely to have this allele than Africans	rs4141830 (T)	Europeans are 176% more likely to have this allele than Africans	rs9611186 (G)	Europeans are 119% more likely to have this allele than Africans
rs229168 (C)	Europeans are 134% more likely to have this allele than Africans	rs4242099 (A)	Europeans are 683% more likely to have this allele than Africans	rs9682919 (A)	Europeans are 233% more likely to have this allele than Africans
rs241558 (A)	Europeans are 750% more likely to have this allele than Africans	rs4314918 (A)	Europeans are 337% more likely to have this allele than Africans	rs9889827 (T)	Europeans are 201% more likely to have this allele than Africans
rs324885 (A)	Europeans are 340% more likely to have this allele than Africans	rs4336505 (A)	Europeans are 156% more likely to have this allele than Africans	rs9933476 (G)	Europeans are 260% more likely to have this allele than Africans
rs482818 (C)	Europeans are 191% more likely to have this allele than Africans	rs4370753 (T)	Europeans are 1141% more likely to have this allele than Africans	rs10057590 (A)	Europeans are 164% more likely to have this allele than Africans
rs589249 (A)	Europeans are 658% more likely to have this allele than Africans	rs4384309 (A)	Europeans are 131% more likely to have this allele than Africans	rs10078630 (A)	Europeans are 389% more likely to have this allele than Africans
rs602512 (A)	Europeans are 388% more likely to have this allele than Africans	rs4425734 (A)	Europeans are 280% more likely to have this allele than Africans	rs10124571 (T)	Europeans are 286% more likely to have this allele than Africans
rs604149 (A)	Europeans are 403% more likely to have this allele than Africans	rs4470366 (A)	Europeans are 202% more likely to have this allele than Africans	rs10205421 (G)	Europeans are 338% more likely to have this allele than Africans
rs628594 (A)	Europeans are 119% more likely to have this allele than Africans	rs4508164 (T)	Europeans are 526% more likely to have this allele than Africans	rs10233848 (A)	Europeans are 142% more likely to have this allele than Africans
rs685968 (A)	Europeans are 733% more likely to have this allele than Africans	rs4537697 (A)	Europeans are 222% more likely to have this allele than Africans	rs10401883 (T)	Europeans are 106% more likely to have this allele than Africans
rs708913 (A)	Europeans are 341% more likely to have this allele than Africans	rs4600898 (A)	Europeans are 184% more likely to have this allele than Africans	rs10402747 (T)	Europeans are 227% more likely to have this allele than Africans
rs708915 (T)	Europeans are 256% more likely to have this allele than Africans	rs4693234 (A)	Europeans are 2297% more likely to have this allele than Africans	rs10411958 (T)	Europeans are 144% more likely to have this allele than Africans
rs736471 (T)	Europeans are 136% more likely to have this allele than Africans	rs4699006 (C)	Europeans are 279% more likely to have this allele than Africans	rs10432638 (A)	Europeans are 1274% more likely to have this allele than Africans
rs737665 (A)	Europeans are 125% more likely to have this allele than Africans	rs4744250 (A)	Europeans are 445% more likely to have this allele than Africans	rs10476217 (G)	Europeans are 163% more likely to have this allele than Africans
rs761718 (A)	Europeans are 659% more likely to have this allele than Africans	rs4767921 (A)	Europeans are 160% more likely to have this allele than Africans	rs10740140 (A)	Europeans are 316% more likely to have this allele than Africans
rs801742 (A)	Europeans are 433% more likely to have this allele than Africans	rs4792250 (A)	Europeans are 2504% more likely to have this allele than Africans	rs10786832 (T)	Europeans are 196% more likely to have this allele than Africans
rs803379 (T)	Europeans are 866% more likely to have this allele than Africans	rs4841430 (G)	Europeans are 112% more likely to have this allele than Africans	rs10821080 (T)	Europeans are 439% more likely to have this allele than Africans
rs872123 (T)	Europeans are 188% more likely to have this allele than Africans	rs4860797 (A)	Europeans are 133% more likely to have this allele than Africans	rs10821977 (A)	Europeans are 559% more likely to have this allele than Africans
rs889169 (A)	Europeans are 285% more likely to have this allele than Africans	rs4881269 (A)	Europeans are 1011% more likely to have this allele than Africans	rs10863246 (A)	Europeans are 387% more likely to have this allele than Africans
rs918765 (C)	Europeans are 148% more likely to have this allele than Africans	rs4895650 (T)	Europeans are 2946% more likely to have this allele than Africans	rs10889949 (T)	Europeans are 115% more likely to have this allele than Africans
rs943262 (C)	Europeans are 108% more likely to have this allele than Africans	rs4916751 (T)	Europeans are 139% more likely to have this allele than Africans	rs10953137 (G)	Europeans are 130% more likely to have this allele than Africans
rs1016678 (A)	Europeans are 121% more likely to have this allele than Africans	rs4941735 (T)	Europeans are 180% more likely to have this allele than Africans	rs11125721 (A)	Europeans are 1437% more likely to have this allele than Africans
rs1050847 (T)	Europeans are 539% more likely to have this allele than Africans	rs4954174 (A)	Europeans are 109% more likely to have this allele than Africans	rs11192147 (A)	Europeans are 188% more likely to have this allele than Africans
rs1083845 (A)	Europeans are 574% more likely to have this allele than Africans	rs4985417 (A)	Europeans are 1305% more likely to have this allele than Africans	rs11210394 (T)	Europeans are 960% more likely to have this allele than Africans
rs1085639 (A)	Europeans are 177% more likely to have this allele than Africans	rs5001573 (C)	Europeans are 134% more likely to have this allele than Africans	rs11576565 (A)	Europeans are 1131% more likely to have this allele than Africans
rs1144709 (C)	Europeans are 168% more likely to have this allele than Africans	rs6119267 (C)	Europeans are 757% more likely to have this allele than Africans	rs11685902 (T)	Europeans are 318% more likely to have this allele than Africans
rs1167800 (G)	Europeans are 344% more likely to have this allele than Africans	rs6503513 (A)	Europeans are 178% more likely to have this allele than Africans	rs11768845 (T)	Europeans are 210% more likely to have this allele than Africans
rs1167827 (A)	Europeans are 1092% more likely to have this allele than Africans	rs6534338 (T)	Europeans are 197% more likely to have this allele than Africans	rs11775314 (T)	Europeans are 443% more likely to have this allele than Africans
rs1291817 (A)	Europeans are 179% more likely to have this allele than Africans	rs6535809 (A)	Europeans are 650% more likely to have this allele than Africans	rs11793831 (T)	Europeans are 350% more likely to have this allele than Africans
rs1291823 (A)	Europeans are 1050% more likely to have this allele than Africans	rs6546856 (T)	Europeans are 418% more likely to have this allele than Africans	rs11864066 (G)	Europeans are 421% more likely to have this allele than Africans
rs1317885 (T)	Europeans are 199% more likely to have this allele than Africans	rs6546857 (A)	Europeans are 104% more likely to have this allele than Africans	rs12136092 (A)	Europeans are 799% more likely to have this allele than Africans
rs1406238 (C)	Europeans are 231% more likely to have this allele than Africans	rs6585426 (T)	Europeans are 122% more likely to have this allele than Africans	rs12146618 (T)	Europeans are 359% more likely to have this allele than Africans
rs1408579 (T)	Europeans are 1487% more likely to have this allele than Africans	rs6708046 (T)	Europeans are 136% more likely to have this allele than Africans	rs12208753 (A)	Europeans are 158% more likely to have this allele than Africans
rs1487441 (A)	Europeans are 156% more likely to have this allele than Africans	rs6708515 (A)	Europeans are 299% more likely to have this allele than Africans	rs12404086 (T)	Europeans are 797% more likely to have this allele than Africans
rs1487445 (T)	Europeans are 183% more likely to have this allele than Africans	rs6721148 (T)	Europeans are 269% more likely to have this allele than Africans	rs12491181 (T)	Europeans are 830% more likely to have this allele than Africans
rs1519799 (A)	Europeans are 589% more likely to have this allele than Africans	rs6732160 (A)	Europeans are 110% more likely to have this allele than Africans	rs12548560 (A)	Europeans are 100% more likely to have this allele than Africans
rs1521162 (C)	Europeans are 925% more likely to have this allele than Africans	rs6752813 (T)	Europeans are 644% more likely to have this allele than Africans	rs12567360 (A)	Europeans are 183% more likely to have this allele than Africans
rs1599180 (A)	Europeans are 252% more likely to have this allele than Africans	rs6760964 (C)	Europeans are 110% more likely to have this allele than Africans	rs12701207 (T)	Europeans are 154% more likely to have this allele than Africans
rs1729799 (T)	Europeans are 340% more likely to have this allele than Africans	rs6779442 (G)	Europeans are 112% more likely to have this allele than Africans	rs12712784 (T)	Europeans are 718% more likely to have this allele than Africans
rs1893907 (C)	Europeans are 167% more likely to have this allele than Africans	rs6799337 (T)	Europeans are 277% more likely to have this allele than Africans	rs12773994 (T)	Europeans are 107% more likely to have this allele than Africans
rs1906252 (A)	Europeans are 158% more likely to have this allele than Africans	rs6823836 (T)	Europeans are 1034% more likely to have this allele than Africans	rs12799722 (A)	Europeans are 148% more likely to have this allele than Africans
rs1956334 (A)	Europeans are 169% more likely to have this allele than Africans	rs6833433 (A)	Europeans are 123% more likely to have this allele than Africans	rs12814021 (T)	Europeans are 187% more likely to have this allele than Africans
rs1963395 (A)	Europeans are 544% more likely to have this allele than Africans	rs6873367 (T)	Europeans are 120% more likely to have this allele than Africans	rs12919084 (T)	Europeans are 236% more likely to have this allele than Africans
rs2005827 (A)	Europeans are 1490% more likely to have this allele than Africans	rs6885410 (A)	Europeans are 262% more likely to have this allele than Africans	rs13004882 (T)	Europeans are 2165% more likely to have this allele than Africans
rs2049741 (T)	Europeans are 188% more likely to have this allele than Africans	rs6887429 (A)	Europeans are 132% more likely to have this allele than Africans	rs13019313 (A)	Europeans are 145% more likely to have this allele than Africans
rs2084131 (A)	Europeans are 468% more likely to have this allele than Africans	rs6917204 (T)	Europeans are 169% more likely to have this allele than Africans	rs13071190 (T)	Europeans are 289% more likely to have this allele than Africans
rs2099744 (A)	Europeans are 123% more likely to have this allele than Africans	rs7016874 (A)	Europeans are 1788% more likely to have this allele than Africans	rs13116036 (T)	Europeans are 2719% more likely to have this allele than Africans
rs2164552 (A)	Europeans are 258% more likely to have this allele than Africans	rs7131440 (T)	Europeans are 290% more likely to have this allele than Africans	rs13270346 (A)	Europeans are 2178% more likely to have this allele than Africans
rs2164553 (A)	Europeans are 258% more likely to have this allele than Africans	rs7131627 (A)	Europeans are 290% more likely to have this allele than Africans	rs13413443 (T)	Europeans are 1394% more likely to have this allele than Africans
rs2221534 (A)	Europeans are 584% more likely to have this allele than Africans	rs7156431 (C)	Europeans are 110% more likely to have this allele than Africans	rs13428598 (T)	Europeans are 417% more likely to have this allele than Africans
rs2250660 (C)	Europeans are 300% more likely to have this allele than Africans	rs7195605 (T)	Europeans are 216% more likely to have this allele than Africans	rs16950305 (G)	Europeans are 100% more likely to have this allele than Africans
rs2279574 (A)	Europeans are 102% more likely to have this allele than Africans	rs7197830 (A)	Europeans are 360% more likely to have this allele than Africans	rs17048855 (A)	Europeans are 595% more likely to have this allele than Africans
rs2364543 (T)	Europeans are 113% more likely to have this allele than Africans	rs7199601 (A)	Europeans are 275% more likely to have this allele than Africans	rs17106817 (T)	Europeans are 145% more likely to have this allele than Africans
rs2430926 (T)	Europeans are 393% more likely to have this allele than Africans	rs7403957 (A)	Europeans are 485% more likely to have this allele than Africans	rs17266097 (T)	Europeans are 1218% more likely to have this allele than Africans
rs2543489 (A)	Europeans are 147% more likely to have this allele than Africans	rs7534577 (A)	Europeans are 535% more likely to have this allele than Africans	rs17608413 (T)	Europeans are 109% more likely to have this allele than Africans
rs2615333 (T)	Europeans are 131% more likely to have this allele than Africans	rs7575938 (A)	Europeans are 128% more likely to have this allele than Africans	rs34389480 (A)	Europeans are 2605% more likely to have this allele than Africans
rs2640564 (T)	Europeans are 232% more likely to have this allele than Africans	rs7624274 (A)	Europeans are 231% more likely to have this allele than Africans	rs56115971 (A)	Europeans are 193% more likely to have this allele than Africans
rs2678897 (G)	Europeans are 311% more likely to have this allele than Africans	rs7779072 (T)	Europeans are 229% more likely to have this allele than Africans	rs56131188 (T)	Europeans are 307% more likely to have this allele than Africans
rs2710691 (T)	Europeans are 382% more likely to have this allele than Africans	rs7896910 (A)	Europeans are 2007% more likely to have this allele than Africans	rs56146203 (A)	Europeans are 195% more likely to have this allele than Africans
rs2819336 (T)	Europeans are 248% more likely to have this allele than Africans	rs7963801 (T)	Europeans are 2981% more likely to have this allele than Africans	rs59692040 (T)	Europeans are 130% more likely to have this allele than Africans
rs2885198 (A)	Europeans are 468% more likely to have this allele than Africans	rs8056960 (C)	Europeans are 336% more likely to have this allele than Africans	rs61372148 (T)	Europeans are 260% more likely to have this allele than Africans
rs2899319 (A)	Europeans are 214% more likely to have this allele than Africans	rs9309922 (T)	Europeans are 108% more likely to have this allele than Africans	rs61527214 (A)	Europeans are 164% more likely to have this allele than Africans
rs2989868 (A)	Europeans are 335% more likely to have this allele than Africans	rs9328534 (T)	Europeans are 176% more likely to have this allele than Africans	rs61816194 (T)	Europeans are 467% more likely to have this allele than Africans
rs3002284 (T)	Europeans are 210% more likely to have this allele than Africans	rs9387711 (T)	Europeans are 139% more likely to have this allele than Africans	rs61854130 (A)	Europeans are 319% more likely to have this allele than Africans
rs3748400 (T)	Europeans are 569% more likely to have this allele than Africans	rs9388490 (T)	Europeans are 121% more likely to have this allele than Africans	rs62383977 (T)	Europeans are 388% more likely to have this allele than Africans
rs3768480 (C)	Europeans are 906% more likely to have this allele than Africans	rs9397806 (A)	Europeans are 113% more likely to have this allele than Africans	rs66493370 (A)	Europeans are 358% more likely to have this allele than Africans
rs3809912 (T)	Europeans are 563% more likely to have this allele than Africans	rs9513780 (T)	Europeans are 261% more likely to have this allele than Africans	rs72845663 (T)	Europeans are 176% more likely to have this allele than Africans
rs3813405 (A)	Europeans are 210% more likely to have this allele than Africans	rs9556702 (A)	Europeans are 1639% more likely to have this allele than Africans		

-Admixture Analysis Is The Bee's Knees:

Prominent environmentalists have explicitly endorsed this kind of admixture analysis before the results were in and they had to think up excuses. For example, Templeton [795] writes:

> *"**There is a way of testing if differences in phenotypic means between two populations have a genetic basis.** The test was developed by Mendel and requires that the populations be crossed and that the hybrids and their descendants be raised in a "common garden" (i.e., a common environment). Despite the extreme interest in the genetic basis of between-population differences in intelligence,*

Continued:

only a handful of studies have even attempted to use this standard research design of genetics. These few studies (Green, 1972; Loehlin, Vandenberg, & Osborne, 1973; Scarr, Pakstis, Katz, & Barker, 1977) have several common features. First, they take advantage of the strong tendency of humans to interbreed when brought into physical proximity. For example, in the Americas, geographically differentiated human populations of European and sub-Saharan African origin were brought together and began to hybridize. However, most matings still occurred within populations. Given this assortative mating, the genetic impact of hybridization is extremely sensitive to the cultural environment. In North America, the hybrids were culturally classified as blacks, and hence most subsequent matings involving the hybrids were into the population of African origin. Therefore, a broad range of variation in degree of European and African ancestry can be found among North American individuals who are all culturally classified as being members of the same "race", in this case blacks (a "common garden" cultural classification). In Latin America, different cultures have different ways of classifying hybrids, but in general a number of alternative categories are available and social class is a more powerful determinant of mating than is physical appearance (e.g., skin colour). As a consequence, individuals in Latin America can be culturally classified into a single social entity that genetically represents a broad range of variation in amount of European and African ancestry. Thus, these studies use a "common garden" design in a cultural sense that nevertheless includes hybrid individuals and their descendants. Second, these studies quantify the degree of European and African ancestry in a population of individuals that is culturally classified as being a single "race." Because the original geographically disparate populations do show genetic differences due to isolation by distance, the degree of European and African ancestry of a specific individual can be estimated using blood group

Continued:

and molecular genetic markers. ***Finally, the shared premise of these studies is that if a trait that differentiates European and sub-Saharan Africans has a genetic basis, it should show variation in the hybrid population that correlates with the degree of African ancestry.*** *This is indeed the case for many morphological traits, such as skin colour (Scarr et al., 1977). However, there is no significant correlation with the degree of African ancestry for any cognitive test result, either within the cultural environment of being "black" (Loehlin et al., 1973; Scarr et al.,1977) or in the cultural environment of being "white" (Green, 1972).* ***Hence, even though these populations differ in their average test scores, there is no evidence for any genetic differentiation among these populations at genetic loci that influence these IQ test scores."***

As another example, in Nisbett's book [796], he specifically advocates using admixture studies in his discussion:

"Racial Ancestry and IQ
All of the research reported above is most consistent with the proposition that the genetic contribution to the black/white dif-ference is nil, but the evidence is not terribly probative one way or the other because it is indirect. ***The only direct evidence on the question of genetics concerns the racial ancestry of a given individual.*** *The genes in the U.S. "black" population are about zo percent European (Parra et al., 1998; Parra, Kittles, and Shriver, 1004). Some blacks have completely African ancestry, many have at least some European ancestry, and some—about to percent—have mostly European ancestry. Does it make a difference how African versus European a black person is?* ***A hereditarian model demands that blacks with more European genes have higher IQs. Herrnstein and Murray (1994) and Rushton and Jensen (2005), as it happens, scarcely deal with this direct evidence.****…*
...So what do we have in the way of ***studies that***

Continued:

examine the effects of racial ancestry—by far the most direct way to assess the contribution of genes versus the environment to the black/white IQ gap? *We have one flawed adoption study with results consistent with the hypothesis that the gap is substantially genetic in origin, and we have two less-flawed adoption studies, one of which indicates slightly superior African genes and one of which suggests no genetic difference. We have downs of studies looking at racial ancestry as indicated by skin colour and "negroidness" of features that provide scant support for the genetic theory. In addition, three different studies of Europeanness of blood groups, using two different designs, indicate no support for the genetic theory. One study of illegitimate children in Germany demonstrates no superiority for children of white fathers as compared to children of black fathers. One study shows that exceptionally bright "black children have no more European ancestry than the best-available estimate for the population as a whole. And one study indicates that A is more advantageous for a mixed-race child to be raised by a family having a white mother than by a family having a black mother. All of these racial ancestry studies are subject to alternative interpretations Most of these alternatives boil down to the possibility that there was self-selection for IQ in black-white unions. If whites who mated with blacks had much lower IQs than whites in general, their European genes would convey little IQ advantage. Similarly, if blacks who mated with whites had much higher IQs than blacks in general, their African genes would not have been a drawback. Yet the extent to which white genes contributing to mixed-race unions would have to be inferior to white genes in general, or black genes would have to be superior to black genes in general, would have to be very extreme*

Continued:

to result in no IQ difference at all between children of purely African heritage and those of partially European origin. Moreover, self-selection by IQ was probably not very great during the slave era, when most black-white unions probably took place. It is unlikely, for example, that the white males who mated with

Continued:

*black females had on average a lower IQ than other white males. Indeed, if such unions mostly involved white male slave-owners and black female slaves, which seems likely to be the case (Parra et al., 1998), and if economic status was slightly positively related to IQ (as it is now), thew whites probably had IQs slightly above average. The black female partners were nor likely chosen on the bask of IQ, as opposed to comeliness. Similarly, it scarcely seems likely that either black or white soldiers in World War II were selecting their German mates on the basis of IQ. Several studies, moreover, are immune to the self-selection hypothesis. In particular, the study involving black and white children raised in an institutional setting, and the study involving black children adopted into either black or white middle-class homes, could not be explained by self-selection for IQ in mating. In short, though one would never know it by reading Herrnstein and Murray's book (1994) or Rushton and Jensen's article (zoos), the great mass of evidence on racial ancestry—****the only direct evidence we have—points toward no contribution at all of genetics to the black/white gap.****"*

As we can see, admixture analysis is widely considered to be the bee's knees.

Race & Neuroanatomy:

Brain size is one of the most well established neurological influences on the general factor of intelligence [see more here]. As will be argued here, there is also a well established racial gap in brain size, and there are multiple lines of evidence that the brain size differences are genetic in origin; The gaps exist in the womb, they have persisted across time, they are ubiquitous across the world, they are consistent with racial differences in a myriad of other traits that coevolve with brain size, there is some evidence that they evolved in response to climate, and intermediate ancestry results in intermediate brain size. This Black-White gap in brain size accounts for 30% of the Black-White IQ gap [812].

The Gaps:

Many are wary of this topic following Stephen Jay Gould's [257] highly influential critique of the subject. In it, Gould argues that researchers involved in this line of work allow their biases to inflate gaps. As a case study, Gould accuses this of a long since deceased researcher, Samuel George Morton, and he accuses Morton of excluding contradictory data from his tables. However, reanalysis of Morton's skulls reveals that errors disfavor Whites, and that the supposedly excluded data was in the very book that Gould cited [813].

It is thus revealed that there has long since been good evidence that there are racial differences in brain volume.

-1. Endocranial Volume:

Aggregated data on a sample of ~20,000, using the same method as Morton where skulls are filled with a substance to measure internal volume, replicates the size differences [814].

-2. MRI:

The first study comparing the brain size of different racial groups via MRI was done in 1994 [815]. The previous findings were confirmed: Blacks have smaller brains than Whites. The same finding was reproduced by source 816, though the study was statistically underpowered, as is [typical] of Neuroscience. For more detailed analysis of racial differences in specific brain regions, see source 817. Notably, racial ancestry can be predicted from brain shape [618]. Racial differences in Neuroanatomy go beyond the straightforwardly physical as well, racial ancestry constitutes a bias in functional MRI (fMRI) [818].

-3. Head Size:

On the opposite end of the spectrum of measurement approaches from MRI, we have raw head sizes. The advantage of this approach is that it can be done inexpensively on large, representative samples of living people. The disadvantage is obvious: the operationalization of brain size; raw head size is less related to intelligence than other measures [361] because while head size is influenced by brain size, there are other influences which reduce the usefulness of head size. So, it's merely a matter of gathering a large amount of evidence, and samples are impressive as expected [819, 820, & 821].

-4. Autopsies:

The final way to measure brain size is to simply rip a brain out of a skull during an autopsy and measure its volume. There is plenty of evidence here, some of it going quite far back [822 - p.137 & 361]. One highly influential critique of the autopsy literature 823] cited and popularized by Gould argued that the literature was invalid because it failed to control for a wide variety of variables such as, but not limited to, age of death, nutritional intake early in life, occupational status, cause of death, time of death, temperature the brain was kept in after death, and the exact place the brain was cut from the spinal cord. The socioeconomic variables are obviously genetically confounded and thus fallacious to control for, but most are valid. This being said, there's no reason to think that the random error would systematically differ by race in the variables such as where the brain stem is cut. Thus such problems should be dealt with simply by aggregating a large amount of data, as has been done [822 - p.137 & 361].

The Cause Of The Size Gap:

-1. Gaps During Youth (Newborns):

Most environmentalists have given up denying the existence of racial gaps in head size, but they have only retreated a few yards. This paper [824], released by a couple quite prominent environmentalists, claims no brain size gap at birth, and doubts genetic mediation between brain size and IQ. Not that there was ever any serious doubt, but multiple papers have evidenced a genetic correlation, some released several years before this paper [363, 364, & 683]. On the claim that there is no racial gap in IQ at birth, they cite source 825; there are a couple of issues:

1. They say it is at birth, but their study is about autopsies, i.e. it is conditioned upon infant death, which may be a disruption.
2. They also want to condition on term length. This is spurious because of the racial differences in gestation; [see coevolution].

The sample size, 782, is also greatly overshadowed by the rest of the evidence.

Source 819:

Analyzing the Collaborative Perinatal Project, which has longitudinal head size data on 53,000 children, 17,000 of them European and 19,000 of them African, the expected brain size differences are replicated.

Source 826:

Though talking of "fetal outcomes", this study is about newborns. This cohort study of 21,500 splits results into high infant head circumference versus average infant head circumference and compares demographics of the two groups. While this statistical approach is poor, and simple d-values would have been preferable, results are still clear. Infants of high head circumference were more likely to be White than infants of average head circumference (82% vs. 74%).

Source 827:

With a sample of 27,229 newborns, Whites and Hispanics had head circumferences .4 cm larger than those of Blacks. Additionally, both gender and racial differences increased with gestational age.

Source 828:

The usual gaps in head circumference are found in a sample of 1,539 infants, though there is no Black group to compare to:

Source 828 - in Table 1:

Head circumference (cm)	
White	34.9 ± 1.5
Asian Indian	34.2 ± 1.4*
Chinese	34.3 ± 1.4*
Other Asian	33.7 ± 1.3*
Hispanic	34.5 ± 1.6*
Other	34.6 ± 1.4*
All	34.5 ± 1.5

* = Group difference from whites is significant (P<0.05)

Prenatal Differences:

Multiple studies also provide evidence that the racial differences in brain size exist in fetuses prior to birth [829, 830, & 831].

-2. Persistence:

Considering the autopsy data [822, p.137], the Black-White doesn't seem to have gotten any smaller over the course of the 20th century:

Male Racial Brain Size Differences 1860-1980

Year	White Male	Asian Male	Black Male	White/Black Gap	White/Asian Gap
1860-1879	1400	1337	1300	100	63
1880-1899	1362	1357	1191	171	5
1900-1919	1381	1367	1274	107	14
1920-1939	1393	1379	1276	117	14
1960-1980	1408	1397	1286	122	11

-3. Ubiquity:

The finding of racial differences in brain size is not one peculiar to any one place in the world; the ubiquity requires a difficult explanation from any cultural theories [814].

-4. Coevolution:

Source 832 took 37 anatomical features identified as co-evolving with the brain in 3 human evolution textbooks, and used the list to compare with 5 forensic anthropology textbooks to look at the racial distributions of these traits. The distributions lined up with the traits as expected in ~80% of cases.

Across 234 mammalian species, brain size correlates with longevity, gestation time, birth weight, litter size, age of first mating, body weight, and body length [833]. These traits differ by race as predicted from the brain size data [822, ch.10].

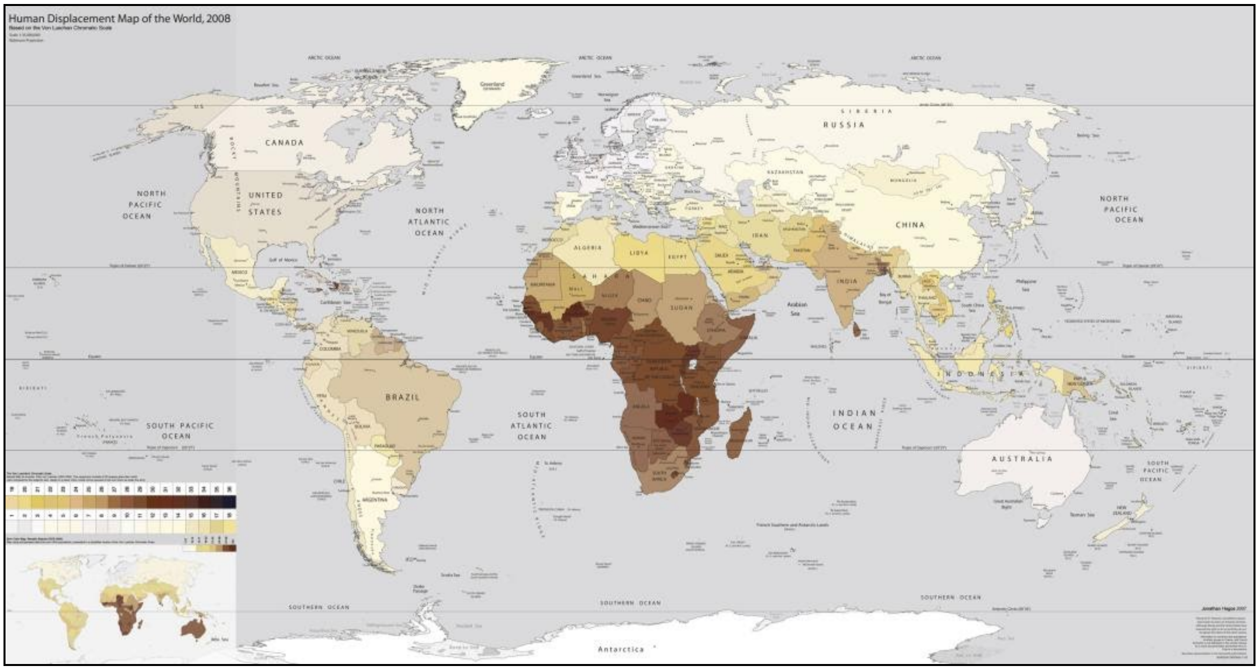
-5. Climate:

There is evidence that the brain size differences evolved in response to climate. There are various hypotheses that could be applied to this; for example, longer, colder winters may require farmers to save up more food during summer to ward off starvation during the winter, when the land temporarily halts productivity. There is a ~.75 correlation between a population's latitude and its brain size [834 & 835]. Analysis of 175 skulls dated 10,000 - 1,900,000 years old, brain size correlations -.41 with winter temperature and .61 with latitude [836].

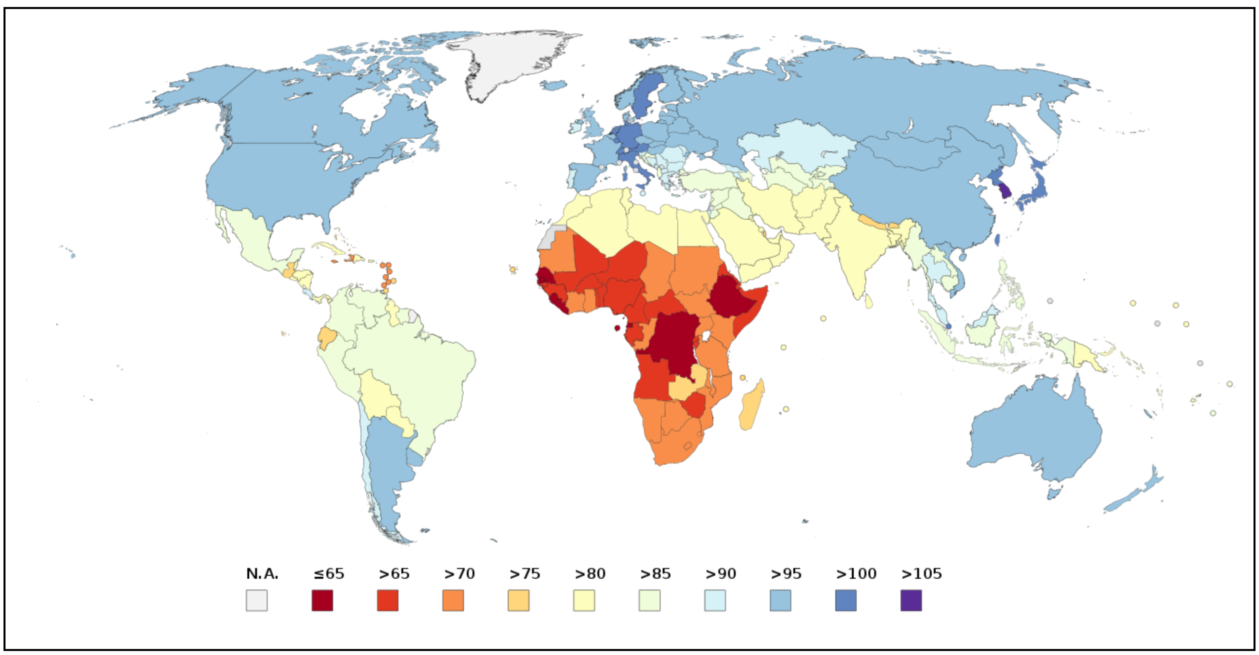
-6. Ancestry & Brain Size:

There has long since been evidence that intermediate racial ancestry results in intermediate brain size [837 & 838].

Skin Tone Map Of The World:



IQ Map Of The World:



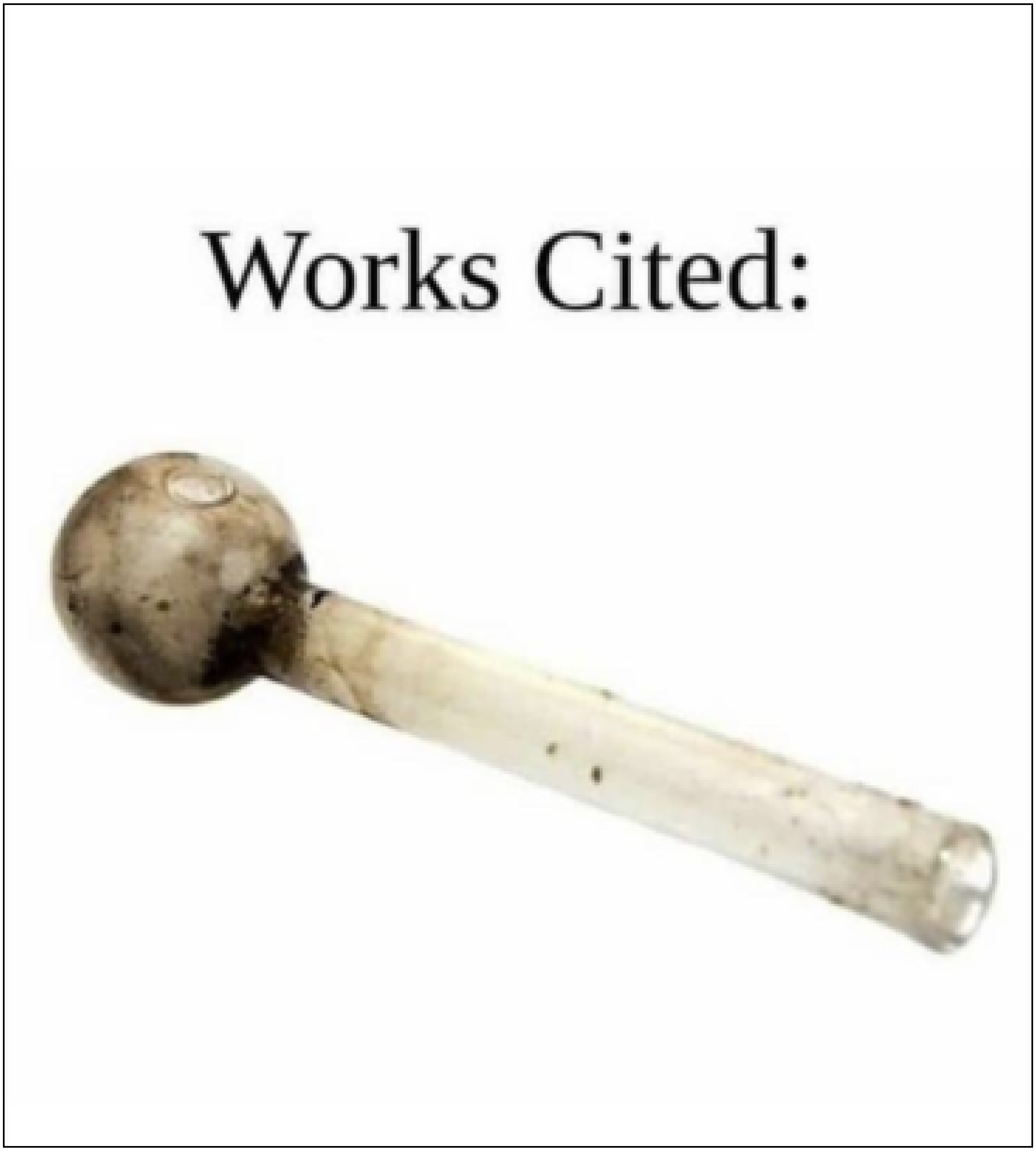
Source List

The list of links to all sources used throughout the document, sorted by the assigned source number which is held constant. If a link is broken, the user has several options. First, you can go to either https://archive.is, https://archive.today, or to https://archive.org and paste in the link to the wayback machines. These sites usually have a working snapshot of whatever link you need. Second, MLA citations of all sources are given so you can manually google for source names or search in journals or libraries or whatever. Third, as many links as possible are doi links put into Sci-hub. You can usually paste a source's doi into https://scholar.google.com and it will give you links to the source, or at least the source's citation. Sci-hub links are a tool to bypass paywalls and read articles for free. Sometimes Sci-hub domains go down, but you can usually find another Sci-hub site which is still up, https://sci-hub.tw might be unavailable while https://sci-hub.se or https://sci-hub.ee or https://scihubtw.tw is available. All you have to do is take the provided doi link and paste it into a working Sci-hub site to get full access to a paper. If you can't find the doi link, sometimes https://search.crossref.org can help to find a doi. In addition, I download pdfs of all of the sources I reference which are freely available to readers on mega.nz google drive. The google doc automatically updates in real time, but the folders on mega and google drive have to be manually updated whenever I feel like I haven't updated them recently enough.

Mega.nz archive: https://mega.nz/folder/PKRHUAiL#KEW3CC_Pa7yCZ4E99Tj-0Q

Google Drive archive:

https://drive.google.com/drive/folders/1N-6RAfTKbAwsspYQ8ENmDHZQg-c6ihR7?usp=sharing


Works Cited:

[1] Springer statement on SCIgen-generated papers in conference proceedings 2014 Retrieved from https://www.springer.com/about+springer/media/statements?SGWID=0-1760813-6-1456249-0

[2] Labbé, Cyril, and Dominique Labbé. "Duplicate and fake publications in the scientific literature: how many SCIgen papers in computer science?." Scientometrics 94.1 (2013): 379-396. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11192-012-0781-y

[3] Bohannon, John. "Who's afraid of peer review?." (2013): 60-65. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.342.6154.60

[4] Armstrong, J. Scott. "Peer review for journals: Evidence on quality control, fairness, and innovation." Science and engineering ethics 3.1 (1997): 63-84. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11948-997-0017-3

[5] Brembs, Björn. "Prestigious science journals struggle to reach even average reliability." Frontiers in human neuroscience 12 (2018): 37. Retrieved from https://sci-hub.se/https://doi.org/10.3389/fnhum.2018.00037

[6] Brembs, Björn, Katherine Button, and Marcus Munafò. "Deep impact: unintended consequences of journal rank." Frontiers in human Neuroscience 7 (2013): 291. Retrieved from https://sci-hub.se/https://doi.org/10.3389/fnhum.2013.00291

[7] Jensen, Arthur Robert. The g factor: The science of mental ability. Vol. 648. Westport, CT: Praeger, 1998. Retrieved from https://emilkirkegaard.dk/en/wp-content/uploads/The-g-factor-the-science-of-mental-ability-Arthur-R.-Jensen.pdf

[8] Brown, Eric N., and S. Ramaswamy. "Quality of protein crystal structures." Acta Crystallographica Section D: Biological Crystallography 63.9 (2007): 941-950. Retrieved from https://sci-hub.se/https://doi.org/10.1107/S0907444907033847

[9] Munafo, M. R., G. Stothart, and J. Flint. "Bias in genetic association studies and impact factor." Molecular psychiatry 14.2 (2009): 119-120. Retrieved from https://sci-hub.se/https://doi.org/10.1038/mp.2008.77

[10] Ziemann, Mark, Yotam Eren, and Assam El-Osta. "Gene name errors are widespread in the scientific literature." Genome biology 17.1 (2016): 1-3. Retrieved from https://sci-hub.se/https://doi.org/10.1186/s13059-016-1044-7

[11] Macleod, Malcolm R., et al. "Risk of bias in reports of in vivo research: a focus for improvement." PLoS Biol 13.10 (2015): e1002273. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pbio.1002273

[12] Szucs, Denes, and John PA Ioannidis. "Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature." PLoS biology 15.3 (2017): e2000797. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pbio.2000797

[13] Fang, Ferric C., and Arturo Casadevall. "Retracted science and the retraction index." (2011): 3855-3859. Retrieved from https://sci-hub.se/https://doi.org/10.1128/IAI.05661-11

[14] Nuijten, Michèle B., et al. "Effect sizes, power, and biases in intelligence research: A meta-meta-analysis." (2018). Retrieved from https://psyarxiv.com/ytsvw

[15] Ernst, E., K. L. Resch, and E. M. Uher. "Reviewer bias." Annals of Internal Medicine 116.11 (1992): 958-958. Retrieved from https://sci-hub.se/https://doi.org/10.7326/0003-4819-116-11-958_2

[16] Epstein, William M. "Confirmational response bias among social work journals." Science, Technology, & Human Values 15.1 (1990): 9-38. Retrieved from https://sci-hub.se/https://doi.org/10.1177/016224399001500102

[17] Emerson, Gwendolyn B., et al. "Testing for the presence of positive-outcome bias in peer review: a randomized controlled trial." Archives of internal medicine 170.21 (2010): 1934-1939. Retrieved from https://sci-hub.se/https://doi.org/10.1001/archinternmed.2010.406

[18] Stribling, Jeremy, Daniel Aguayo, and Maxwell Krohn. "Rooter: A methodology for the typical unification of access points and redundancy." Journal of Irreproducible Results 49.3 (2005): 5. Retrieved from http://pdos.csail.mit.edu/scigen/rooter.pdf

[19] Schroter, Sara, et al. "What errors do peer reviewers detect, and does training improve their ability to detect them?." Journal of the Royal Society of Medicine 101.10 (2008): 507-514. Retrieved from https://sci-hub.se/https://doi.org/10.1258/jrsm.2008.080062

[20] Rothwell, Peter M., and Christopher N. Martyn. "Reproducibility of peer review in clinical neuroscience: Is agreement between reviewers any greater than would be expected by chance alone?." Brain 123.9 (2000): 1964-1969. Retrieved from https://sci-hub.se/https://doi.org/10.1093/brain/123.9.1964

[21] Ross, Joseph S., et al. "Effect of blinded peer review on abstract acceptance." Jama 295.14 (2006): 1675-1680. Retrieved from https://sci-hub.se/https://doi.org/10.1001/jama.295.14.1675

[22] Tomkins, Andrew, Min Zhang, and William D. Heavlin. "Reviewer bias in single-versus double-blind peer review." Proceedings of the National Academy of Sciences 114.48 (2017): 12708-12713. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.1707323114

[23] Vasilevsky, Nicole A., et al. "On the reproducibility of science: unique identification of research resources in the biomedical literature." PeerJ 1 (2013): e148. Retrieved from https://core.ac.uk/download/pdf/26001642.pdf

[24] Downey, Autumn S., and Steve Olson, eds. Sharing clinical research data: workshop summary. National Academies Press, 2013. Retrieved from https://www.nap.edu/read/18267/chapter/1

[25] Rothwell, Peter M., and Christopher N. Martyn. "Reproducibility of peer review in clinical neuroscience: Is agreement between reviewers any greater than would be expected by chance alone?." Brain 123.9 (2000): 1964-1969. Retrieved from https://sci-hub.se/https://doi.org/10.1093/brain/123.9.1964

[26] Bornmann, Lutz, Rüdiger Mutz, and Hans-Dieter Daniel. "A reliability-generalization study of journal peer reviews: A multilevel meta-analysis of inter-rater reliability and its determinants." PloS one 5.12 (2010): e14331. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pone.0014331

[27] Cicchetti, Domenic V. "The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation." Behavioral and brain sciences 14.1 (1991): 119-135. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0140525X00065675

[28] Jefferson, Tom, et al. "Effects of editorial peer review: a systematic review." Jama 287.21 (2002): 2784-2786. Retrieved from https://sci-hub.se/https://doi.org/10.1001/jama.287.21.2784

[29] Elvik, Rune. "Are road safety evaluation studies published in peer reviewed journals more valid than similar studies not published in peer reviewed journals?." Accident Analysis & Prevention 30.1 (1998): 101-118. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0001-4575(97)00068-7

[30] MacNealy, Mary Sue, Bruce W. Speck, and Noel Clements. "Publishing in Technical Communication Journals from the Successful Author's Point of View." Technical communication (1994): 240-259. Retrieved from https://www.jstor.org/stable/43090322?seq=1

[31] Bradley, James V. "Pernicious publication practices." Bulletin of the Psychonomic Society 18.1 (1981): 31-34. Retrieved from https://sci-hub.se/https://doi.org/10.3758/BF03333562

[32] Fletcher, Robert H., and Suzanne W. Fletcher. "Evidence for the effectiveness of peer review." Science and Engineering Ethics 3.1 (1997): 35-50. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11948-997-0015-5

[33] Evans, James T., Howard I. Nadjari, and Sherry A. Burchell. "Quotational and reference accuracy in surgical journals: A continuing peer review problem." JAMA 263.10 (1990): 1353-1354. Retrieved from https://sci-hub.se/https://doi.org/10.1001/jama.1990.03440100059009

[34] Armstrong, J. Scott. "Barriers to scientific contributions: The author's formula." Behavioral and Brain Sciences 5.2 (1982): 197-199. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0140525X00011201

[35] Begg, Colin B., and Jesse A. Berlin. "Publication bias: a problem in interpreting medical data." Journal of the Royal Statistical Society: Series A (Statistics in Society) 151.3 (1988): 419-445. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2982993

[36] Gallup. "What America Needs to Know about Higher Education Redesign." (2014). Retrieved from https://www.gallup.com/file/services/176759/2013%20Gallup-Lumina%20Foundation%20Report.pdf

[37] Eichorn, Philip, and Alfred Yankauer. "Do authors check their references? A survey of accuracy of references in three public health journals." American Journal of Public Health 77.8 (1987): 1011-1012. Retrieved from https://sci-hub.se/https://doi.org/10.2105/AJPH.77.8.1011

[38] Greenwald, Anthony G. "Consequences of prejudice against the null hypothesis." Psychological bulletin 82.1 (1975): 1. Retrieved from https://sci-hub.se/https://doi.org/10.1037/h0076157

[39] Hubbard, Raymond, and J. Scott Armstrong. "Replications and extensions in marketing: Rarely published but quite contrary." International Journal of Research in Marketing 11.3 (1994): 233-248. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0167-8116(94)90003-5

[40] Salsburg, David S. "The religion of statistics as practiced in medical journals." The American Statistician 39.3 (1985): 220-223. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2683942

[41] Sterling, Theodore D., Wilf L. Rosenbaum, and James J. Weinkam. "Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa." The American Statistician 49.1 (1995): 108-112. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2684823

[42] Cohen, Jacob. "The earth is round (p<.05)." American psychologist 49.12 (1994): 997. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0003-066X.49.12.997

[43] McCloskey, Deirdre N., and Stephen T. Ziliak. "The standard error of regressions." Journal of economic literature 34.1 (1996): 97-114. Retrieved from http://ww.w.deirdremccloskey.com/docs/pdf/Article_189.pdf

[44] Atkinson, Donald R., Michael J. Furlong, and Bruce E. Wampold. "Statistical significance, reviewer evaluations, and the scientific process: Is there a (statistically) significant relationship?." Journal of Counseling Psychology 29.2 (1982): 189. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-0167.29.2.189

[45] Sterling, Theodore D. "Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa." Journal of the American statistical association 54.285 (1959): 30-34. Retrieved from https://sci-hub.se/https://doi.org/10.1080/01621459.1959.10501497

[46] Armstrong, Jon Scott. Long-range forecasting. New York ETC.: Wiley, 1985. Retrieved from https://b-ok.cc/book/810586/e75efc

[47] Lau, Richard R. "An analysis of the accuracy of “trial heat” polls during the 1992 presidential election." Public Opinion Quarterly 58.1 (1994): 2-20. Retrieved from https://sci-hub.se/https://doi.org/10.1086/269405

[48] Armstrong, J. Scott. "Unintelligible management research and academic prestige." Interfaces 10.2 (1980): 80-86. Retrieved from https://sci-hub.se/https://doi.org/10.1287/inte.10.2.80

[49] Jauch, Lawrence R., and Jerry L. Wall. "What they do when they get your manuscript: A survey of Academy of Management reviewer practices." Academy of Management Journal 32.1 (1989): 157-173. Retrieved from https://sci-hub.se/https://doi.org/10.2307/256424

[50] Kronick, David A. "Scientific journals: a review article." (1982): 265-269. Retrieved from https://sci-hub.se/https://doi.org/10.1086/601244

[51] Lock, Stephen, and Jane Smith. "What do peer reviewers do?." Jama 263.10 (1990): 1341-1343. Retrieved from https://sci-hub.se/https://doi.org/10.1001/jama.1990.03440100045006

[52] Yankauer, Alfred. "Who are the peer reviewers and how much do they review?." Jama 263.10 (1990): 1338-1340. Retrieved from https://sci-hub.se/https://doi.org/10.1001/jama.1990.03440100042005

[53] Wilson, Jean D. "Peer review and publication. Presidential address before the 70th annual meeting of the American Society for Clinical Investigation, San Francisco, California, 30 April 1978." The Journal of Clinical Investigation 61.6 (1978): 1697-1701. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC372697/pdf/jcinvest00666-0281.pdf

[54] Lock, Stephen, and Jane Smith. "PEER-REVIEW AT WORK." *Scholarly Publishing* 17.4 (1986): 304-316.

[55] Hargens, Lowell L. "Variation in journal peer review systems: Possible causes and consequences." Jama 263.10 (1990): 1348-1352. Retrieved from https://sci-hub.se/https://doi.org/10.1001/jama.1990.03440100052008

[56] Patterson, Samuel C., and Shannon K. Smithey. "Monitoring Scholarly Journal Publication in Political Science: The Role of the APSR." PS: Political Science & Politics 23.4 (1990): 647-656. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S104909650003403X

[57] Mueller, Ralph O. "The impact of college selectivity on income for men and women." Research in Higher Education 29.2 (1988): 175-191. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF00992285

[58] Thomas, Scott L. "Deferred costs and economic returns to college major, quality, and performance." Research in Higher Education 41.3 (2000): 281-313. Retrieved from https://sci-hub.se/https://doi.org/10.1023/A:1007003510102

[59] Dale, Stacy Berg, and Alan B. Krueger. "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables." The Quarterly Journal of Economics 117.4 (2002): 1491-1527. Retrieved from https://sci-hub.se/https://doi.org/10.1162/003355302320935089

[60] Smith, Tom W., Davern, Michael, Freese, Jeremy, and Morgan, Stephen L., General Social Surveys, 1972-2018 [machine-readable data file] /Principal Investigator, Smith, Tom W.; Co-Principal Investigators, Michael Davern, Jeremy Freese and Stephen L. Morgan; Sponsored by National Science Foundation. --NORC ed.-- Chicago: NORC, 2019.
1 data file (64,814 logical records) + 1 codebook (3,758 pp.). -- (National Data Program for the Social Sciences, no. 25). Retrieved from http://gss.norc.org/

[61] Kirsch, Irwin S. Adult literacy in America: A first look at the results of the National Adult Literacy Survey. US Government Printing Office, Superintendent of Documents, Washington, DC 20402 (Stock No. 065-000-00588-3)., 1993. Retrieved from https://nces.ed.gov/pubs93/93275.pdf

[62] U.S. Department of Education. National Center for Education Statistics. Literacy in the Labor Force: Results from the NationalAdult Literacy Survey, NCES 1999–470, by Andrew Sum. Project Officer: Andrew Kolstad. Washington DC. 1999. Retrieved from https://nces.ed.gov/pubs99/1999470.pdf

[63] National Science Foundation, National Center for Science and Engineering Statistics. 2017. Doctorate Recipients from U.S. Universities: 2015. Special Report NSF 17-306. Arlington, VA. Available at https://www.nsf.gov/statistics/2017/nsf17306/static/report/nsf17306.pdf

[64] Schmidt, Frank L., and John Hunter. "General mental ability in the world of work: occupational attainment and job performance." Journal of personality and social psychology 86.1 (2004): 162. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-3514.86.1.162

[65] Lindsey, Robert V., et al. "Improving students' long-term knowledge retention through personalized review." Psychological science 25.3 (2014): 639-647. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0956797613504302

[66] Bacon, Donald R., and Kim A. Stewart. "How fast do students forget what they learn in consumer behavior? A longitudinal study." Journal of Marketing Education 28.3 (2006): 181-192. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0273475306291463

[67] Faught, Brent E., et al. *How much do students remember over time? Longitudinal knowledge retention in traditional versus accelerated learning environments*. Toronto: Higher Education Quality Council of Ontario, 2016. Retrieved from https://collections.ola.org/mon/31001/338257.pdf

[68] Marcel, F. D. "Knowledge loss of medical students on first year basic science courses at the University of Saskatchewan." BMC medical education 6.1 (2006): 1-6. Retrieved from https://sci-hub.se/https://doi.org/10.1186/1472-6920-6-5

[69] HELLO HELLO HELLO https://www.bitchute.com/channel/1MUVaTZWgbLj/

[70] Kohen, Andrew I., and Paul H. Kipps. "Factors determining student retention of economic knowledge after completing the principles-of-microeconomics course." The Journal of Economic Education 10.2 (1979): 38-48. Retrieved from https://sci-hub.se/https://doi.org/10.2307/1182376

[71] Custers, Eugène JFM. "Long-term retention of basic science knowledge: a review study." Advances in Health Sciences Education 15.1 (2010): 109-128. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10459-008-9101-y

[72] Grove, William M., et al. "Clinical versus mechanical prediction: a meta-analysis." Psychological assessment 12.1 (2000): 19. Retrieved from https://sci-hub.se/https://doi.org/10.1037/1040-3590.12.1.19

[73] of Menie, Michael A. Woodley, and Heitor BF Fernandes. "The secular decline in general intelligence from decreasing developmental stability: Theoretical and empirical considerations." Personality and Individual Differences 92 (2016): 194-199. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2015.12.035

[74] Anderson, Melissa S., et al. "The perverse effects of competition on scientists' work and relationships." Science and engineering ethics 13.4 (2007): 437-461. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11948-007-9042-5

[75] Abbott, A., Cyranoski, D., Jones, N., Maher, B., Schiermeier, Q., & Van Noorden, R. (2010). Metrics: Do metrics matter? Nature, 465(7300), 860–862. Retrieved from https://sci-hub.se/https://doi.org/10.1038/465860a

[76] PLoS Medicine Editors. "The impact factor game." PLoS Med 3.6 (2006): e291. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pmed.0030291

[77] Fanelli, Daniele. "Do pressures to publish increase scientists' bias? An empirical support from US States Data." PloS one 5.4 (2010): e10271. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pone.0010271

[78] Garfield, Eugene. "Citation indexes for science." Science 122.3159 (1953): 108-111. Retrieved from http://garfield.library.upenn.edu/essays/v6p468y1983.pdf

[79] Lexchin, Joel, et al. "Pharmaceutical industry sponsorship and research outcome and quality: systematic review." Bmj 326.7400 (2003): 1167-1170. Retrieved from https://sci-hub.se/https://doi.org/10.1136/bmj.326.7400.1167

[80] Bekelman, Justin E., Yan Li, and Cary P. Gross. "Scope and impact of financial conflicts of interest in biomedical research: a systematic review." Jama 289.4 (2003): 454-465. Retrieved from https://sci-hub.se/https://doi.org/10.1001/jama.289.4.454

[81] van Wesel, Maarten. "Evaluation by citation: Trends in publication behavior, evaluation criteria, and the strive for high impact publications." Science and Engineering Ethics 22.1 (2016): 199-225. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11948-015-9638-0

[82] Bjork, Bo-Christer, Annikki Roos, and Mari Lauri. "Scientific journal publishing: yearly volume and open access availability." Information Research: An International Electronic Journal 14.1 (2009). Retrieved from https://pdfs.semanticscholar.org/912d/22ea1243d94719b9dcf451c6c619316da1d1.pdf

[83] Jinha, Arif E. "Article 50 million: an estimate of the number of scholarly articles in existence." Learned Publishing 23.3 (2010): 258-263. Retrieved from https://sci-hub.se/https://doi.org/10.1087/20100308

[84] Gordon, Michael D. "Citation ranking versus subjective evaluation in the determination of journal hierachies in the social sciences." Journal of the American Society for Information Science 33.1 (1982): 55-57. Retrieved from https://sci-hub.se/https://doi.org/10.1002/asi.4630330109

[85] Saha, Somnath, Sanjay Saint, and Dimitri A. Christakis. "Impact factor: a valid measure of journal quality?." Journal of the Medical Library Association 91.1 (2003): 42. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC141186/pdf/i0025-7338-091-01-0042.pdf

[86] Yue, Weiping, Concepción S. Wilson, and Francois Boller. "Peer assessment of journal quality in clinical neurology." Journal of the Medical Library Association 95.1 (2007): 70. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1773051/pdf/i0025-7338-095-01-0070.pdf

[87] Sønderstrup-Andersen, Eva, and Hans Sønderstrup-Andersen. "An investigation into diabetes researcher's perceptions of the Journal Impact Factor—reconsidering evaluating research." Scientometrics 76.2 (2008): 391-406. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11192-007-1924-4

[88] Jacobs, Nele, et al. "Heritability estimates of intelligence in twins: Effect of chorion type." Behavior genetics 31.2 (2001): 209-217. Retrieved from https://sci-hub.se/https://doi.org/10.1023/A:1010257512183

[89] Fang, Ferric C., R. Grant Steen, and Arturo Casadevall. "Misconduct accounts for the majority of retracted scientific publications." Proceedings of the National Academy of Sciences 109.42 (2012): 17028-17033. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.1212247109

[90] Steen, R. Grant. "Retractions in the scientific literature: is the incidence of research fraud increasing?." Journal of medical ethics 37.4 (2011): 249-253. Retrieved from https://sci-hub.se/https://doi.org/10.1136/jme.2010.040923

[91] Davis, Mark S., Michelle Riske-Morris, and Sebastian R. Diaz. "Causal factors implicated in research misconduct: Evidence from ORI case files." Science and engineering ethics 13.4 (2007): 395-414. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11948-007-9045-2

[92] Anderson, Melissa S., Brian C. Martinson, and Raymond De Vries. "Normative dissonance in science: Results from a national survey of US scientists." Journal of Empirical Research on Human Research Ethics 2.4 (2007): 3-14. Retrieved from https://sci-hub.se/https://doi.org/10.1525/jer.2007.2.4.3

[93] Louis, Karen Seashore, Melissa S. Anderson, and Lenn Rosenberg. "Academic misconduct and values: The department's influence." The Review of Higher Education 18.4 (1995): 393-422. Retrieved from https://sci-hub.se/https://doi.org/10.1353/rhe.1995.0007

[94] Munafo, M. R., I. J. Matheson, and J. Flint. "Association of the DRD2 gene Taq1A polymorphism and alcoholism: a meta-analysis of case–control studies and evidence of publication bias." Molecular psychiatry 12.5 (2007): 454-461. Retrieved from https://sci-hub.se/https://doi.org/10.1038/sj.mp.4001938

[95] Song, Fujian, et al. "Publication and related biases: a review." Health technology assessment 4.10 (2000): 1-115. Retrieved from https://pdfs.semanticscholar.org/f28b/a558556cb522e7d2d7ee23b0b9ef5686f2f6.pdf

[96] Dwan, Kerry, et al. "Systematic review of the empirical evidence of study publication bias and outcome reporting bias." PloS one 3.8 (2008): e3081. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pone.0003081

[97] Hopewell, Sally, et al. "Time to publication for results of clinical trials." Cochrane database of systematic reviews 2 (2005). Retrieved from https://sci-hub.se/https://doi.org/10.1002/14651858.MR000011.pub2

[98] Scherer, Roberta W., et al. "Full publication of results initially presented in abstracts." Cochrane Database of Systematic Reviews 11 (2007). Retrieved from https://sci-hub.se/https://doi.org/10.1002/14651858.MR000005.pub3

[99] Leimu, Roosa, and Julia Koricheva. "What determines the citation frequency of ecological papers?." Trends in Ecology & Evolution 20.1 (2005): 28-32. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.tree.2004.10.010

[100] Eysenck, Hans JÃrgen. The structure and measurement of intelligence. Transaction Publishers, 1979. Retrieved from https://hanseysenck.com/wp-content/uploads/2019/12/1979_eysenck_-_the_structure_and_measurement_of_intelligence.pdf

[101] Etter, Jean-François, and John Stapleton. "Citations to trials of nicotine replacement therapy were biased toward positive results and high-impact-factor journals." Journal of Clinical Epidemiology 62.8 (2009): 831-837. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jclinepi.2008.09.015

[102] Kjaergard, Lise L., and Christian Gluud. "Citation bias of hepato-biliary randomized clinical trials." Journal of clinical epidemiology 55.4 (2002): 407-410. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0895-4356(01)00513-3

[103] Murtaugh, Paul A. "Journal quality, effect size, and publication bias in meta-analysis." Ecology 83.4 (2002): 1162-1166. Retrieved from https://sci-hub.se/https://doi.org/10.1890/0012-9658(2002)083[1162:JQESAP]2.0.CO;2

[104] Gigerenzer G, Swijtink Z, Porter T, Daston L, Beatty J, et al. (1990) The empire of chance: How probability changed science and everyday life. Cambridge: Cambridge University Press. Sorry, no pdf: https://www.amazon.com/Empire-Chance-Probability-Changed-Everyday-ebook/dp/B00E3UR71S

[105] Sharpe, Donald. "Beyond significance testing: Reforming data analysis methods in behavioral research." (2004): 317. Retrieved from https://sci-hub.se/https://doi.org/10.1037/h0087004

[106] Kyzas, Panayiotis A., Despina Denaxa-Kyza, and John PA Ioannidis. "Almost all articles on cancer prognostic markers report statistically significant results." European Journal of Cancer 43.17 (2007): 2559-2579. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.ejca.2007.08.030

[107] Csada, Ryan D., Paul C. James, and Richard HM Espie. "The" file drawer problem" of non-significant results: does it apply to biological research?." Oikos (1996): 591-593. Retrieved from https://sci-hub.se/https://doi.org/10.2307/3546355

[108] Jennions, Michael D., and Anders P. Moeller. "Publication bias in ecology and evolution: an empirical assessment using the 'trim and fill' method." Biological Reviews 77.2 (2002): 211-222. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S1464793101005875

[109] Sterling, Theodore D., Wilf L. Rosenbaum, and James J. Weinkam. "Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa." The American Statistician 49.1 (1995): 108-112. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2684823

[110] Mookerjee, Rajen. "A meta-analysis of the export growth hypothesis." Economics letters 91.3 (2006): 395-401. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.econlet.2005.12.020

[111] Mackintosh, Nicholas, and Nicholas John Mackintosh. IQ and human intelligence. Oxford University Press, 2011. Retrieved from https://ia803103.us.archive.org/9/items/N.J.MackintoshIQAndHumanIntelligence1998OxfordUniversityPress/N.%20J.%20Mackintosh%20-%20IQ%20and%20Human%20Intelligence%20%281998%2C%20Oxford%20University%20Press%29.pdf

[112] Gerber, Alan S., and Neil Malhotra. "Publication bias in empirical sociological research: Do arbitrary significance levels distort published results?." Sociological Methods & Research 37.1 (2008): 3-30. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0049124108318973

[113] Marsh, David M., and Teresa J. Hanlon. "Seeing what we want to see: Confirmation bias in animal behavior research." Ethology 113.11 (2007): 1089-1098. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1439-0310.2007.01406.x

[114] Rosenthal, Robert. "Experimenter effects in behavioral research." (1976). Retrieved from https://www.gwern.net/docs/statistics/bias/1976-rosenthal-experimenterexpectancyeffects.pdf

[115] Nickerson, Raymond S. "Confirmation bias: A ubiquitous phenomenon in many guises." Review of general psychology 2.2 (1998): 175-220. Retrieved from https://sci-hub.se/https://doi.org/10.1037/1089-2680.2.2.175

[116] Mahoney, Michael J. "Psychology of the scientist: An evaluative review." Social studies of Science 9.3 (1979): 349-375. Retrieved from https://sci-hub.se/https://doi.org/10.1177/030631277900900304

[117] Felson, Jacob. "What can we learn from twin studies? A comprehensive evaluation of the equal environments assumption." Social science research 43 (2014): 184-199. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.ssresearch.2013.10.004

[118] Kerr, Norbert L. "HARKing: Hypothesizing after the results are known." Personality and social psychology review 2.3 (1998): 196-217. Retrieved from https://sci-hub.se/https://doi.org/10.1207/s15327957pspr0203_4

[119] Chan, An-Wen, et al. "Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles." Jama 291.20 (2004): 2457-2465. Retrieved from https://sci-hub.se/https://doi.org/10.1001/jama.291.20.2457

[120] De Vries, Raymond, Melissa S. Anderson, and Brian C. Martinson. "Normal misbehavior: Scientists talk about the ethics of research." Journal of Empirical Research on Human Research Ethics 1.1 (2006): 43-50. Retrieved from https://sci-hub.se/https://doi.org/10.1525/jer.2006.1.1.43

[121] Fanelli, Daniele. "How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data." PloS one 4.5 (2009): e5738. Retrieved from https://projects.iq.harvard.edu/files/ras-rcr/files/plosone-fanelli-how_many_fabricate_research.pdf

[122] Klein, Daniel B., and Charlotta Stern. "How politically diverse are the social sciences and humanities? Survey evidence from six fields." Academic Questions, Transaction Publishers, Forthcoming (2004). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.120.9897&rep=rep1&type=pdf

[123] Jewish Virtual Library, Talk Magazine, Stolen Children: Interview with Gitta Sereny. Retrieved from https://archive.is/GdU8i

[124] Baker, Monya. "Reproducibility crisis." Nature 533.26 (2016): 353-66. Retrieved from https://www.nature.com/news/polopoly_fs/1.19970!/menu/main/topColumns/topLeftColumn/pdf/533452a.pdf

[125] Open Science Collaboration. "Estimating the reproducibility of psychological science." Science 349.6251 (2015). Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.aac4716

[126] Camerer, Colin F., et al. "Evaluating replicability of laboratory experiments in economics." Science 351.6280 (2016): 1433-1436. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.aaf0918

[127] Begley, C. Glenn, and Lee M. Ellis. "Raise standards for preclinical cancer research." Nature 483.7391 (2012): 531-533. Retrieved from https://www.nature.com/articles/483531a.pdf

[128] Boekel, Wouter, et al. "A purely confirmatory replication study of structural brain-behavior correlations." Cortex 66 (2015): 115-133. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.cortex.2014.11.019

[129] Camerer, Colin F., et al. "Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015." Nature Human Behaviour 2.9 (2018): 637-644. Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41562-018-0399-z

[130] Forsell, Eskil, et al. "Predicting replication outcomes in the Many Labs 2 study." Journal of Economic Psychology 75 (2019): 102117. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.joep.2018.10.009

[131] Vines, Timothy H., et al. "The availability of research data declines rapidly with article age." Current biology 24.1 (2014): 94-97. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.cub.2013.11.014

[132] Kirkegaard 2017 https://emilkirkegaard.dk/en/?p=6864

[133] Steen, R. Grant. "Retractions in the scientific literature: do authors deliberately commit research fraud?." Journal of medical ethics 37.2 (2011): 113-117. Retrieved from https://sci-hub.se/https://doi.org/10.1136/jme.2010.038125

[134] Duarte, José L., et al. "Political diversity will improve social psychological science 1." Behavioral and Brain Sciences 38 (2015). Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0140525X14000430

[135] Langbert, Mitchell. "Homogenous: The political affiliations of elite liberal arts college faculty." Academic Questions 31.2 (2018): 186-197. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s12129-018-9700-x

[136] Fanelli, Daniele. "“Positive” results increase down the hierarchy of the sciences." PloS one 5.4 (2010): e10068. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pone.0010068

[137] Duncan, Greg J., and Katherine Magnuson. "Investing in preschool programs." Journal of economic perspectives 27.2 (2013): 109-32. Retrieved from https://sci-hub.se/https://doi.org/10.1257/jep.27.2.109

[138] https://archive.is/tpVsY

[139] Prinz, Florian, Thomas Schlange, and Khusru Asadullah. "Believe it or not: how much can we rely on published data on potential drug targets?." Nature reviews Drug discovery 10.9 (2011): 712-712. Retrieved from https://sci-hub.se/https://doi.org/10.1038/nrd3439-c1

[140] Carroll, John B. Human cognitive abilities: A survey of factor-analytic studies. Cambridge University Press, 1993. Retrieved from https://b-ok.cc/book/850847/764ee8

[141] the math proof for why r^2 = r * r: https://economictheoryblog.com/2014/11/05/proof/

[142] te Nijenhuis, Jan, Birthe Jongeneel-Grimen, and Emil OW Kirkegaard. "Are Headstart gains on the g factor? A meta-analysis." Intelligence 46 (2014): 209-215. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2014.07.001

[143] Winegard, Bo M., et al. "Low-status groups as a domain of liberal bias." Manuscript submitted for publication (2019). Retrieved from https://www.researchgate.net/publication/326144740_Low-status_groups_as_a_domain_of_liberal_bias

[144] Cova, Florian, et al. "Estimating the reproducibility of experimental philosophy." Review of Philosophy and Psychology (2018): 1-36. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s13164-018-0400-9

[145] Tate, Brian W., and Michael A. Mcdaniel. "Race differences in personality: An evaluation of moderators and publication bias." (2008). Retrieved from https://www.researchgate.net/publication/237215311_Race_Differences_in_Personality_An_Evaluation_of_Moderators_and_Publication_Bias

[146] https://archive.is/MutW2

[147] https://archive.is/dgMUs

[148] Church, A. Timothy. "Personality measurement in cross-cultural perspective." Journal of Personality 69.6 (2001): 979-1006. Retrieved from https://sci-hub.se/https://doi.org/10.1111/1467-6494.696172

[149] Dana, Richard H. "Culture and methodology in personality assessment." Handbook of multicultural mental health. Academic Press, 2000. 97-120. Retrieved from https://sci-hub.se/https://doi.org/10.1016/B978-012199370-2/50007-9

[150] Sesardic, Neven. Making sense of heritability. Cambridge University Press, 2005. Retrieved from https://lesacreduprintemps19.files.wordpress.com/2012/11/making-sense-of-heritability-neven-sesardic.pdf

[151] Rindermann, Heiner, David Becker, and Thomas R. Coyle. "Survey of expert opinion on intelligence: Intelligence research, experts' background, controversial issues, and the media." Intelligence 78 (2020): 101406. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2019.101406

[152] Dumas-Mallet, Estelle, et al. "Low statistical power in biomedical science: a review of three human research domains." Royal Society open science 4.2 (2017): 160254. Retrieved from https://sci-hub.se/https://doi.org/10.1098/rsos.160254

[153] Lortie-Forgues, Hugues, and Matthew Inglis. "Rigorous large-scale educational RCTs are often uninformative: Should we be concerned?." Educational Researcher 48.3 (2019): 158-166. Retrieved from https://sci-hub.se/https://doi.org/10.3102/0013189X19832850

[154] Szucs, Denes, and John PA Ioannidis. "Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature." PLoS biology 15.3 (2017): e2000797. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pbio.2000797

[155] Smaldino, Paul E., and Richard McElreath. "The natural selection of bad science." Royal Society open science 3.9 (2016): 160384. Retrieved from https://sci-hub.se/https://doi.org/10.1098/rsos.160384

[156] Button, Katherine S., et al. "Power failure: why small sample size undermines the reliability of neuroscience." Nature reviews neuroscience 14.5 (2013): 365-376. Retrieved from https://sci-hub.se/https://doi.org/10.1038/nrn3475

[157] Kirkegaard emails retrieved from https://archive.is/Tbclc

[158] Barnett, Jennifer H., Linda Scoriels, and Marcus R. Munafò. "Meta-analysis of the cognitive effects of the catechol-O-methyltransferase gene Val158/108Met polymorphism." Biological psychiatry 64.2 (2008): 137-144. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.biopsych.2008.01.005

[159] Yang, Lin, et al. "Psychiatric illness and intellectual disability in the Prader–Willi syndrome with different molecular defects-a meta analysis." PLoS one 8.8 (2013): e72640. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pone.0072640

[160] Zhang, Jian-Ping, et al. "Meta-analysis of genetic variation in DTNBP1 and general cognitive ability." Biological psychiatry 68.12 (2010): 1126-1133. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.biopsych.2010.09.016

[161] Luciano, Michelle, et al. "Testing replication of a 5-SNP set for general cognitive ability in six population samples." European Journal of Human Genetics 16.11 (2008): 1388-1395. Retrieved from https://sci-hub.se/https://doi.org/10.1038/ejhg.2008.100

[162] Beaujean, A. Alexander. "Heritability of cognitive abilities as measured by mental chronometric tasks: A meta-analysis." Intelligence 33.2 (2005): 187-201. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2004.08.001

[163] Oh yeah yeah https://www.youtube.com/watch?v=7fFSupGfZME

[164] https://minecraft.gamepedia.com/Cake

[165] Scarr-Salapatek, Sandra. "Race, social class, and IQ." Science 174.4016 (1971): 1285-1295. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.174.4016.1285

[166] Turkheimer, Eric, et al. "The Scarr–Rowe interaction between measured socioeconomic status and the heritability of cognitive ability." Experience and development: A festschrift in honor of Sandra Wood Scarr 2 (2009): 81-97. Retrieved from https://b-ok.cc/book/3676817/57d26b

[167] Figlio, David N., et al. "Socioeconomic status and genetic influences on cognitive development." Proceedings of the National Academy of Sciences 114.51 (2017): 13441-13446. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.1708491114

[168] google scholar https://scholar.google.com/scholar?hl=en&as_sdt=0%2C44&q=Socioeconomic+Status+Modifies+Heritability+of+IQ+in+Young+Children&btnG=

[169] Reeve, Charlie L., and Jennifer E. Charles. "Survey of opinions on the primacy of g and social consequences of ability testing: A comparison of expert and non-expert views." Intelligence 36.6 (2008): 681-688. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2008.03.007

[170] https://www.dailywire.com/news/new-york-times-hires-openly-anti-white-racist-it-joseph-curl

[171] https://archive.is/Qe27p

[172] Haier, Richard J. The neuroscience of intelligence. Cambridge University Press, 2016. Retrieved from https://ia800105.us.archive.org/17/items/TheNeuroscienceOfIntelligence/The-Neuroscience-of-Intelligence.pdf

[173] Vakhtin, Andrei A., et al. "Functional brain networks contributing to the Parieto-Frontal Integration Theory of Intelligence." Neuroimage 103 (2014): 349-354. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.neuroimage.2014.09.055

[174] Chabris, Christopher F. "Cognitive and neurobiological mechanisms of the Law of General Intelligence." (2007). Retrieved from http://www.chabris.com/Chabris2007a.pdf

[175] Thompson, Bruce. Exploratory and confirmatory factor analysis. American Psychological Association, 2004. Retrieved from https://b-ok.cc/book/901756/c5c4d0

[176] Brown, Timothy A. *Confirmatory factor analysis for applied research*. Guilford publications, 2015. Retrieved from https://b-ok.cc/book/2481257/a7ecc5

[177] Irvine, Sidney Herbert. *Human abilities in cultural context*. Cambridge University Press, 1988. Retrieved from https://b-ok.cc/book/875136/2d51c6

[178] Jensen, Arthur R. "Effects of inbreeding on mental-ability factors." Personality and Individual Differences 4.1 (1983): 71-87. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0191-8869(83)90054-5

[179] te Nijenhuis, Jan, Kenya Kura, and Yoon-Mi Hur. "The correlation between g loadings and heritability in Japan: A meta-analysis." Intelligence 46 (2014): 275-282. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2014.07.008

[180] Taylor, Ronald L., and Edward W. Ziegler. "Comparison of the first principal factor on the WISC-R across ethnic groups." Educational and Psychological Measurement 47.3 (1987): 691-694. Retrieved from https://sci-hub.se/https://doi.org/10.1177/001316448704700318

[181] Warne, Russell T., and Cassidy Burningham. "Spearman's g found in 31 non-Western nations: Strong evidence that g is a universal phenomenon." *Psychological bulletin* 145.3 (2019): 237. Retrieved from https://psycnet.apa.org/fulltext/2019-01683-001.pdf

[182] Galsworthy, Michael J., Rosalind Arden, and Christopher F. Chabris. "Animal models of general cognitive ability for genetic research into cognitive functioning." Behavior genetics of cognition across the lifespan. Springer, New York, NY, 2014. 257-278. Retrieved from https://sci-hub.se/https://doi.org/10.1007/978-1-4614-7447-0_9

[183] of Menie, Michael A. Woodley, Heitor BF Fernandes, and William D. Hopkins. "The more g-loaded, the more heritable, evolvable, and phenotypically variable: Homology with humans in chimpanzee cognitive abilities." *Intelligence* 50 (2015): 159-163. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.04.002

[184] Jensen, Arthur R. "Bias in mental testing." (1980). Retrieved from https://arthurjensen.net/wp-content/uploads/2020/04/Bias-in-Mental-Testing-Arthur-R.-Jensen.pdf

[185] Brown, Robert T., Cecil R. Reynolds, and Jean S. Whitaker. "Bias in mental testing since Bias in Mental Testing." School Psychology Quarterly 14.3 (1999): 208. Retrieved from https://sci-hub.se/https://doi.org/10.1037/h0089007

[186] Valencia, Richard R., and Lisa A. Suzuki. *Intelligence testing and minority students: Foundations, performance factors, and assessment issues*. Vol. 3. Sage Publications, 2000. Retrieved from https://b-ok.cc/book/2610928/7072e1

[187] Gustafsson, Jan-Eric. "The relevance of factor analysis for the study of group differences." Multivariate Behavioral Research 27.2 (1992): 239-247. Retrieved from http://arthurjensen.net/wp-content/uploads/2014/06/1992-gustafsson.pdf

[188] Carretta, Thomas R., and Malcolm James Ree. "Near identity of cognitive structure in sex and ethnic groups." Personality and Individual Differences 19.2 (1995): 149-155. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0191-8869(95)00031-Z

[189] Ree, Malcolm James, and Thomas R. Carretta. "Group differences in aptitude factor structure on the ASVAB." Educational and Psychological Measurement 55.2 (1995): 268-277. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0013164495055002011

[190] Kaufman, Alan S., James C. Kaufman, and James E. McLean. "Factor structure of the Kaufman Adolescent and Adult Intelligence Test (KAIT) for whites, African Americans, and hispanics." Educational and Psychological Measurement 55.3 (1995): 365-376. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0013164495055003001

[191] Pandolfi, Vincent. "Assessment of factor models underlying the WISC-III in White, Black, and Hispanic subgroups of the standardization sample." (1998): 4464-4464. https://www.gwern.net/docs/iq/1997-pandolfi.pdf

[192] Keith, Timothy Z., et al. "Construct bias in the Differential Ability Scales? Confirmatory and hierarchical factor structure across three ethnic groups." Journal of Psychoeducational Assessment 17.3 (1999): 249-268. Retrieved from https://sci-hub.se/https://doi.org/10.1177/073428299901700305

[193] Reed, Cametra L. *An investigation of measurement invariance in the WISC III: examining a sample of referred African American and Caucasian students*. Diss. University of South Florida, 2000. Retrieved from https://www.dropbox.com/s/xtf331p1239vnnz/2000-reed.pdf?dl=0

[194] Dolan, Conor V. "Investigating Spearman's hypothesis by means of multi-group confirmatory factor analysis." Multivariate Behavioral Research 35.1 (2000): 21-50. Retrieved from https://sci-hub.se/https://doi.org/10.1207/S15327906MBR3501_2

[195] Dolan, Conor V., and Ellen L. Hamaker. "Investigating Black-White differences in psychometric IQ: Multi-group confirmatory factor analyses of the WISC-R and K-ABC and a critique of the method of correlated vectors." Advances in psychology research 6 (2001): 31-59. Retrieved from https://www.researchgate.net/profile/Conor_Dolan/publication/46716651_Investigating_Black-White_differences_in_psychometric_IQ_Multi-group_confirmatory_factor_analyses_of_the_WISC-R_and_K-ABC_and_a_critique_of_the_method_of_correlated_vectors/links/02e7e51e592b11f729000000/Investigating-Black-White-differences-in-psychometric-IQ-Multi-group-confirmatory-factor-analyses-of-the-WISC-R-and-K-ABC-and-a-critique-of-the-method-of-correlated-vectors.pdf

[196] Kush, Joseph C., et al. "Construct validity of the WISC-III for White and Black students from the WISC-III standardization sample and for Black students referred for psychological evaluation." School Psychology Review 30.1 (2001): 70-88. Retrieved from http://edpsychassociates.com/Papers/WISC3BlackFactor.pdf

[197] Lubke, Gitta H., et al. "On the relationship between sources of within-and between-group differences and measurement invariance in the common factor model." Intelligence 31.6 (2003): 543-566. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0160-2896(03)00051-5

[198] Dolan, Conor V., and Peter CM Molenaar. "Testing specific hypotheses concerning latent group differences in multi-group covariance structure analysis with structured means." Multivariate Behavioral Research 29.3 (1994): 203-222. Retrieved from https://sci-hub.se/https://doi.org/10.1207/s15327906mbr2903_1

[199] Dalliard, Marc. "The elusive x-factor: A critique of JM Kaplan's model of race and IQ." Open Differ. Psychol 27 (2014). Retrieved from https://openpsych.net/files/papers/Dalliard_2014a.pdf

[200] Hunt, Earl. Human intelligence. Cambridge University Press, 2010. Retrieved from https://ia800105.us.archive.org/6/items/EarlHuntHumanIntelligence2010/Earl%20Hunt%20-%20Human%20Intelligence%20%282010%2C%20Cambridge%20University%20Press%29.pdf

[201] Warne, Russell T., Mayson C. Astle, and Jessica C. Hill. "What do undergraduates learn about human intelligence? An analysis of introductory psychology textbooks." Archives of Scientific Psychology 6.1 (2018): 32. Retrieved from https://www.gwern.net/docs/iq/2018-warne.pdf#page=12

[202] Adolf, Janne, et al. "Measurement invariance within and between individuals: A distinct problem in testing the equivalence of intra-and inter-individual model structures." Frontiers in psychology 5 (2014): 883. Retrieved from https://sci-hub.se/https://doi.org/10.3389/fpsyg.2014.00883

[203] Jak, Suzanne, and Terrence D. Jorgensen. "Relating measurement invariance, cross-level invariance, and multilevel reliability." Frontiers in psychology 8 (2017): 1640. Retrieved from https://sci-hub.se/https://doi.org/10.3389/fpsyg.2017.01640

[204] Floyd, Rachele L., Kathleen Gathercoal, and Gale Roid. "No Evidence for Ethnic and Racial Bias in the Tryout Edition of the Merrill-Palmer Scale—Revised." Psychological reports 94.1 (2004): 217-220. Retrieved from https://sci-hub.se/https://doi.org/10.2466/pr0.94.1.217-220

[205] Edwards, Oliver W., and Thomas D. Oakland. "Factorial invariance of Woodcock-Johnson III scores for African Americans and Caucasian Americans." Journal of Psychoeducational Assessment 24.4 (2006): 358-366. Retrieved rom https://sci-hub.se/https://doi.org/10.1177/0734282906289595

[206] Koretz, Daniel, and Young-Suk Kim. "Changes in the Black-White Test score Gap in the Elementary School Grades. CSE Report 715." National Center for Research on Evaluation, Standards, and Student Testing (CRESST) (2007). Retrieved from https://files.eric.ed.gov/fulltext/ED503292.pdf

[207] Fryer Jr, Roland G., and Steven D. Levitt. "The black-white test score gap through third grade." American law and economics review 8.2 (2006): 249-281. Retrieved from https://sci-hub.se/https://doi.org/10.1093/aler/ahl003

[208] Fryer Jr, Roland G., and Steven D. Levitt. "Understanding the black-white test score gap in the first two years of school." Review of economics and statistics 86.2 (2004): 447-464. Retrieved from https://sci-hub.se/https://doi.org/10.1162/003465304323031049

[209] Roznowski, Mary, and Janet Reith. "Examining the measurement quality of tests containing differentially functioning items: do biased items result in poor measurement?." Educational and psychological Measurement 59.2 (1999): 248-269. Retrieved from https://sci-hub.se/https://doi.org/10.1177/00131649921969839

[210] U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), Long-Term Trend Assessments, 2004, 2008 and 2012. Retrieved from http://nces.ed.gov/nationsreportcard/tdw/analysis/scaling_avoidviolat_results.aspx

[211] Kane, Harrison. "Race differences on the UNIT: Evidence from multi-sample confirmatory analysis." Mankind Quarterly 48.3 (2008): 283. Retrieved from https://www.dropbox.com/s/smhl4u8fzx4cppn/kane.pdf?dl=0

[212] Huaqing Qi, Cathy, and Scott C. Marley. "Differential item functioning analysis of the Preschool Language Scale—4 between English-speaking Hispanic and European American children from low-income families." Topics in Early Childhood Special Education 29.3 (2009): 171-180. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0271121409332674

[213] Carretta, Thomas R., and Malcolm James Ree. "Factor structure of the air force officer qualifying test: Analysis and comparison." Military Psychology 8.1 (1996): 29-42. Retrieved from https://sci-hub.se/https://doi.org/10.1207/s15327876mp0801_3

[214] Kane, Harrison D., and Thomas D. Oakland. "Group Differences in Cognitive Ability: A CHC Theory Framework." Mankind Quarterly 50.4 (2010): 318. Retrieved from https://www.dropbox.com/s/k6bi2i4lpg784im/group%20diffs%20mq%20kane%20oakland.pdf?dl=0

[215] Trundt, Katherine Marie. "Construct bias in the differential ability scales, (DAS-II): a comparison among African American, Asian, Hispanic, and White ethnic groups." (2013). Retrieved from https://mh19871004.files.wordpress.com/2014/05/construct-bias-in-the-differential-ability-scales-second-edition-das-ii-a-comparison-among-african-american-asian-hispanic-and-white-ethnic-groups-trundt-2013.pdf

[216] Beaujean, A. Alexander, and Sean M. McGlaughlin. "Invariance in the Reynolds Intellectual Assessment Scales for Black and White referred students." Psychological assessment 26.4 (2014): 1394. Retrieved from https://sci-hub.se/https://doi.org/10.1037/pas0000029

[217] Frisby, Craig L., and A. Alexander Beaujean. "Testing Spearman's hypotheses using a bi-factor model with WAIS-IV/WMS-IV standardization data." Intelligence 51 (2015): 79-97. Retrieved from http://emilkirkegaard.dk/en/wp-content/uploads/Testing-Spearmans-hypotheses-using-a-bi-factor-model-with-WAIS-IVWMS-IV-standardization-data.pdf

[218] Blankson, A. Nayena, and John J. McArdle. "Measurement invariance of cognitive abilities across ethnicity, gender, and time among older Americans." Journals of Gerontology Series B: Psychological Sciences and Social Sciences 70.3 (2015): 386-397. Retrieved from https://sci-hub.se/https://doi.org/10.1093/geronb/gbt106

[219] Scheiber, Caroline. Do the Kaufman Tests of Cognitive Ability and Academic Achievement Display Ethnic Bias for Students in Grades 1 Through 12?. Alliant International University, 2016. Retrieved from https://www.dropbox.com/s/836rkigugjsdbv8/Do_the_Kaufman_Tests_of_Cognit.pdf?dl=0

[220] Barnes, Lisa L., et al. "Examination of the factor structure of a global cognitive function battery across race and time." Journal of the International Neuropsychological Society 22.1 (2016): 66-75. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S1355617715001113

[221] Scheiber, Caroline. "Do the Kaufman tests of cognitive ability and academic achievement display construct bias across a representative sample of Black, Hispanic, and Caucasian school-age children in grades 1 through 12?." Psychological assessment 28.8 (2016): 942. Retrieved from https://www.dropbox.com/s/6lbd7dzvna8be7f/2016-scheiber.pdf?dl=0

[222] Jensen, Arthur R., and Frank CJ McGurk. "Black-white bias in 'cultural' and 'noncultural' test items." Personality and individual differences 8.3 (1987): 295-301. Retrieved from http://arthurjensen.net/wp-content/uploads/2014/06/Black-White-Bias-in-Cultural-and-Noncultural-Test-Items-1987-by-Arthur-Robert-Jensen-Frank-C.-J.-McGurk.pdf

[223] Millsap, Roger E. "Invariance in measurement and prediction revisited." Psychometrika 72.4 (2007): 461-473. Retrieved from https://sci-hub.se/https://doi.org/10.1007/S11336-007-9039-7

[224] Borsboom, Denny. "The attack of the psychometricians." Psychometrika 71.3 (2006): 425. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11336-006-1447-6

[225] Scheiber, Caroline. "Is the Cattell–Horn–Carroll-Based Factor Structure of the Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V) Construct Invariant for a Representative Sample of African–American, Hispanic, and Caucasian Male and Female Students Ages 6 to 16 Years?." Journal of Pediatric Neuropsychology 2.3-4 (2016): 79-88. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s40817-016-0019-7

[226] Trundt, K. M., Keith, T. Z., Caemmerer, J. M., & Smith, L. V. (2018). Testing for Construct Bias in the Differential Ability Scales: A Comparison Among African American, Asian, Hispanic, and Caucasian Children. Journal of Psychoeducational Assessment, 36(7), 670-683. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0734282917698303

[227] Hu, Meng. “Modeling the Spearman's Hypothesis Using MGCFA - the Woodcock-johnson Data.” PsyArXiv, 2 Apr. 2017. Web. Retrieved from https://psyarxiv.com/3z9gd

[228] Hu, M., Lasker, J., Kirkegaard, E. O., & Fuerst, J. G. (2019). Filling in the Gaps: The Association between Intelligence and Both Color and Parent-Reported Ancestry in the National Longitudinal Survey of Youth 1997. Psych, 1(1), 240-261. Retrieved from https://sci-hub.se/https://doi.org/10.3390/psych1010017

[229] Al-Bursan, Ismael S., et al. "Sex Differences in 32,347 Jordanian 4th Graders on the National Exam of Mathematics." Journal of Individual Differences (2018). Retrieved from https://www.researchgate.net/publication/329197116_Sex_Differences_in_32347_Jordanian_4th_Graders_on_the_National_Exam_of_Mathematics

[230] Jensen, Arthur R., and Frank CJ McGurk. "Black-white bias in ‘cultural’and ‘noncultural’test items." Personality and individual differences 8.3 (1987): 295-301. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0191-8869(87)90029-8

[231] Kirkegaard, Emil OW, Julius D. Bjerrekær, and Noah Carl. "Cognitive ability and political preferences in Denmark." Open Quantitative Sociology & Political ScienceE 1.1 (2017). Retrieved from https://www.researchgate.net/profile/Emil_O_W_Kirkegaard/publication/313890514_Cognitive_Ability_and_Political_Preferences_in_Denmark/links/58adc44b92851cf7ae85ac79/Cognitive-Ability-and-Political-Preferences-in-Denmark.pdf

[232] Kirkegaard, Emil OW, and Julius D. Bjerrekær. "The OKCupid dataset: A very large public dataset of dating site users." Open Differential Psychology 46 (2016): 1-10. Retrieved from the link below (Warning: direct download link)
https://www.researchgate.net/profile/Emil_O_W_Kirkegaard/project/The-OKCupid-dataset-A-very-large-public-dataset-of-dating-site-users/attachment/573b1d7e08aea7adff2f3c8d/AS:362725751443460@1463491966863/download/paper.pdf?context=ProjectUpdatesLog

[233] Loehlin, John C. "Race differences in intelligence." (1975). Retrieved from https://b-ok.cc/book/2474317/30342f

[234] Rowe, David C., Alexander T. Vazsonyi, and Daniel J. Flannery. "No more than skin deep: Ethnic and racial similarity in developmental process." Psychological Review 101.3 (1994): 396. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0033-295X.101.3.396

[235] Rowe, David C., Alexander T. Vazsonyi, and Daniel J. Flannery. "Ethnic and racial similarity in developmental process: A study of academic achievement." Psychological Science 6.1 (1995): 33-38. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-9280.1995.tb00301.x

[236] Rowe, David C., and Hobart H. Cleveland. "Academic achievement in Blacks and Whites: Are the developmental processes similar?." Intelligence 23.3 (1996): 205-228. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0160-2896(96)90004-5

[237] Rowe, David C., and Joseph Ed Rodgers. "Under the skin: On the impartial treatment of genetic and environmental hypotheses of racial differences." American Psychologist 60.1 (2005): 60. Retrieved from https://mh19871004.files.wordpress.com/2014/05/under-the-skin-on-the-impartial-treatment-of-genetic-and-environmental-hypotheses-of-racial-differences-rowe-2005.pdf

[238] Thorndike, Robert L. "Stability of factor loadings." Personality and Individual Differences 8.4 (1987): 585-586. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0191-8869(87)90224-8

[239] Johnson, Wendy, et al. "Just one g: Consistent results from three test batteries." Intelligence 32.1 (2004): 95-107. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0160-2896(03)00062-X

[240] Johnson, W., Nijenhuis, J. T., & Bouchard Jr, T. J. (2008). Still just 1g: Consistent results from five test batteries. Intelligence, 36(1), 81-95. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2007.06.001

[241] Quiroga, M. Ángeles, et al. "Can we reliably measure the general factor of intelligence (g) through commercial video games? Yes, we can!." Intelligence 53 (2015): 1-7. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.08.004

[242] Keith, T. Z., Kranzler, J. H., & Flanagan, D. P. (2001). What does the Cognitive Assessment System (CAS) measure? Joint confirmatory factor analysis of the CAS and the Woodcock-Johnson Tests of Cognitive Ability. School Psychology Review, 30(1), 89-119. Retrieved from http://www.iapsych.com/wj3ewok/LinkedDocuments/Keith2001.pdf

[243] Floyd, R. G., Bergeron, R., Hamilton, G., & Parra, G. R. (2010). How do executive functions fit with the Cattell–Horn–Carroll model? Some evidence from a joint factor analysis of the Delis–Kaplan executive function system and the Woodcock–Johnson III tests of cognitive abilities. Psychology in the Schools, 47(7), 721-738. Retrieved from https://sci-hub.se/https://doi.org/10.1002/pits.20500

[244] Floyd, R. G., Reynolds, M. R., Farmer, R. L., Kranzler, J. H., & Volpe, R. (2013). Are the General Factors From Different Child And Adolescent Intelligence Tests the Same? Results From a Five-Sample, Six-Test Analysis. School Psychology Review, 42(4). Retrieved from https://emilkirkegaard.dk/en/wp-content/uploads/Are-the-General-Factors-From-Different-Child-And-Adolescent-Intelligence-Tests-the-Same.pdf

[245] Roth, Bettina, et al. "Intelligence and school grades: A meta-analysis." Intelligence 53 (2015): 118-137. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.09.002

[246] Brodnick, R. J., and Malcolm James Ree. "A structural model of academic performance, socioeconomic status, and Spearman's g." Educational and Psychological Measurement 55.4 (1995): 583-594. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0013164495055004006

[247] Frey, Meredith C., and Douglas K. Detterman. "Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability." Psychological science 15.6 (2004): 373-378. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.0956-7976.2004.00687.x

[248] Beaujean, A. Alexander, et al. "Validation of the Frey and Detterman (2004) IQ prediction equations using the Reynolds Intellectual Assessment Scales." Personality and Individual Differences 41.2 (2006): 353-357. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2006.01.014

[249] Deary, Ian J., et al. "Intelligence and educational achievement." Intelligence 35.1 (2007): 13-21. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2006.02.001

[250] Tucker-Drob, Elliot M., and Timothy C. Bates. "Large cross-national differences in gene× socioeconomic status interaction on intelligence." Psychological science 27.2 (2016): 138-149. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0956797615612727

[251] Robertson, Kimberley Ferriman, et al. "Beyond the threshold hypothesis: Even among the gifted and top math/science graduate students, cognitive abilities, vocational interests, and lifestyle preferences matter for career choice, performance, and persistence." Current Directions in Psychological Science 19.6 (2010): 346-351. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0963721410391442

[252] Park, Gregory, David Lubinski, and Camilla P. Benbow. "Ability differences among people who have commensurate degrees matter for scientific creativity." Psychological Science 19.10 (2008): 957-961. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-9280.2008.02182.x

[253] Strenze, Tarmo. "Intelligence and socioeconomic success: A meta-analytic review of longitudinal research." Intelligence 35.5 (2007): 401-426. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2006.09.004

[254] Jensen, Arthur R., and Li-Jen Weng. "What is a good g?." (1994): 231-258. Retrieved from http://arthurjensen.net/wp-content/uploads/2014/06/What-Is-a-Good-g-1994-by-Arthur-Robert-Jensen-Li-Jen-Weng.pdf

[255] Ree, Malcolm James, and James A. Earles. "The stability of g across different methods of estimation." Intelligence 15.3 (1991): 271-278. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(91)90036-D

[256] Revelle, William, and Joshua Wilt. "The general factor of personality: A general critique." Journal of research in personality 47.5 (2013): 493-504. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jrp.2013.04.012

[257] Gould, Stephen Jay, and Steven James Gold. The mismeasure of man. WW Norton & company, 1996. Retrieved from https://b-ok.cc/book/651282/e99763

[258] Thurstone, Louis Leon. "Multiple-factor analysis; a development and expansion of The Vectors of Mind." (1947). https://www.amazon.com/Multiple-Factor-Analysis-Development-Expansion-Vectors/dp/B0007F2ZWO

[259] McGrew, Kevin S. "CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research." (2009): 1-10. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2008.08.004

[260] Spearman, Charles. "The abilities of man: Their nature and measurement." (1927). Retrieved from https://ia800308.us.archive.org/27/items/abilitiesofman031969mbp/abilitiesofman031969mbp.pdf

[261] Schubert, Anna-Lena. "A meta-analysis of the worst performance rule." Intelligence 73 (2019): 88-100. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2019.02.003

[262] Dutton, Edward, Dimitri van der Linden, and Richard Lynn. "The negative Flynn Effect: A systematic literature review." Intelligence 59 (2016): 163-169. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2016.10.002

[263] Woodley, Michael A., and Gerhard Meisenberg. "In the Netherlands the anti-Flynn effect is a Jensen effect." Personality and Individual Differences 54.8 (2013): 871-876. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2012.12.022

[264] Wicherts, Jelte M., et al. "Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect." Intelligence 32.5 (2004): 509-537. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2004.07.002

[265] Beaujean, A. Alexander. Using item response theory to assess the Lynn-Flynn effect. Diss. University of Missouri-Columbia, 2006. Retrieved from http://www.azmonyar.com/DownloadPDF/89592758.pdf

[266] Beaujean, A. Alexander, and Steven J. Osterlind. "Using item response theory to assess the Flynn effect in the National Longitudinal Study of Youth 79 Children and Young Adults data." Intelligence 36.5 (2008): 455-463. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2007.10.004

[267] Must, Olev, et al. "Comparability of IQ scores over time." Intelligence 37.1 (2009): 25-33. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2008.05.002

[268] Shiu, William, et al. "An item-level examination of the Flynn effect on the National Intelligence Test in Estonia." Intelligence 41.6 (2013): 770-779. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2013.05.007

[269] Must, Olev, and Aasa Must. "Changes in test-taking patterns over time." Intelligence 41.6 (2013): 780-790. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2013.04.005

[270] Woodley, Michael Anthony, et al. "Controlling for increased guessing enhances the independence of the Flynn effect from g: The return of the Brand effect." Intelligence 43 (2014): 27-34. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2013.12.004

[271] Must, Olev, and Aasa Must. "Speed and the Flynn Effect." Intelligence 68 (2018): 37-47. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2018.03.001

[272] Beaujean, A. Alexander, and Yanyan Sheng. "Examining the Flynn effect in the general social survey vocabulary test using item response theory." Personality and Individual Differences 48.3 (2010): 294-298. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2009.10.019

[273] Wai, Jonathan, and Martha Putallaz. "The Flynn effect puzzle: A 30-year examination from the right tail of the ability distribution provides some missing pieces." Intelligence 39.6 (2011): 443-455. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2011.07.006

[274] te Nijenhuis, Jan, and Henk Van Der Flier. "Is the Flynn effect on g?: A meta-analysis." Intelligence 41.6 (2013): 802-807. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2013.03.001

[275] te Nijenhuis, Jan, Annelies EM van Vianen, and Henk van der Flier. "Score gains on g-loaded tests: No g." Intelligence 35.3 (2007): 283-300. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2006.07.006

[276] Sala, Giovanni, and Fernand Gobet. "Cognitive training does not enhance general cognition." Trends in cognitive sciences 23.1 (2019): 9-20. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.tics.2018.10.004

[277] Fox, Mark C., and Ainsley L. Mitchum. "A knowledge-based theory of rising scores on “culture-free” tests." Journal of Experimental Psychology: General 142.3 (2013): 979. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0030155

[278] Gignac, Gilles E. "Raven's is not a pure measure of general intelligence: Implications for g factor theory and the brief measurement of g." Intelligence 52 (2015): 71-79. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.07.006

[279] Pietschnig, Jakob, Ulrich S. Tran, and Martin Voracek. "Item-response theory modeling of IQ gains (the Flynn effect) on crystallized intelligence: Rodgers' hypothesis yes, Brand's hypothesis perhaps." Intelligence 41.6 (2013): 791-801. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2013.06.005

[280] Fox, Mark C., and Ainsley L. Mitchum. "Confirming the cognition of rising scores: Fox and Mitchum (2013) predicts violations of measurement invariance in series completion between age-matched cohorts." PloS one 9.5 (2014): e95780. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pone.0095780

[281] Beaujean, Alexander, and Yanyan Sheng. "Assessing the Flynn effect in the Wechsler scales." Journal of Individual Differences (2014). Retrieved from https://www.gwern.net/docs/iq/2014-beaujean.pdf

[282] Major, Jason T., Wendy Johnson, and Thomas J. Bouchard Jr. "The dependability of the general factor of intelligence: Why small, single-factor models do not adequately represent g." Intelligence 39.5 (2011): 418-433. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2011.07.002

[283] Kolen, M. J., & Brennan, R. L. (2004). Test equating, scaling, and linking: Methods and practices (2nd ed.). New York, NY: Springer. Retrieved from https://b-ok.cc/book/2135959/110c86

[284] Benson, Nicholas, A. Alexander Beaujean, and Gordon E. Taub. "Using score equating and measurement invariance to examine the Flynn effect in the Wechsler Adult Intelligence Scale." Multivariate behavioral research 50.4 (2015): 398-415. Retrieved from https://sci-hub.se/https://doi.org/10.1080/00273171.2015.1022642

[285] Pietschnig, Jakob, and Georg Gittler. "A reversal of the Flynn effect for spatial perception in German-speaking countries: Evidence from a cross-temporal IRT-based meta-analysis (1977–2014)." Intelligence 53 (2015): 145-153. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.10.004

[286] Ståhl, Tomas, and Jan-Willem Van Prooijen. "Epistemic rationality: Skepticism toward unfounded beliefs requires sufficient cognitive ability and motivation to be rational." Personality and Individual Differences 122 (2018): 155-163. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2017.10.026

[287] Soto, Christopher J. "How replicable are links between personality traits and consequential life outcomes? The Life Outcomes of Personality Replication Project." Psychological Science 30.5 (2019): 711-727. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0956797619831612

[288] Lyu, Xiao-Kang, et al. "Beyond psychology: prevalence of p value and confidence interval misinterpretation across different fields." Journal of Pacific Rim Psychology 14 (2020). Retrieved from https://sci-hub.se/https://doi.org/10.1017/prp.2019.28

[289] Zuckerman, Miron, et al. "Contemporary issues in the analysis of data: A survey of 551 psychologists." Psychological Science 4.1 (1993): 49-53. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-9280.1993.tb00556.x

[290] Hoekstra, Rink, et al. "Robust misinterpretation of confidence intervals." Psychonomic bulletin & review 21.5 (2014): 1157-1164. Retrieved from https://sci-hub.se/https://doi.org/10.3758/s13423-013-0572-3

[291] Haller, Heiko, and Stefan Krauss. "Misinterpretations of significance: A problem students share with their teachers." Methods of Psychological Research 7.1 (2002): 1-20. Retrieved from https://www.researchgate.net/profile/Heiko_Haller/publication/27262211_Misinterpretations_of_Significance_A_Problem_Students_Share_with_Their_Teachers/links/0deec53198736a7dae000000.pdf

[292] McShane, Blakeley B., and David Gal. "Blinding us to the obvious? The effect of statistical training on the evaluation of evidence." Management Science 62.6 (2016): 1707-1718. Retrieved from https://sci-hub.se/https://doi.org/10.1287/mnsc.2015.2212

[293] McShane, Blakeley B., and David Gal. "Statistical significance and the dichotomization of evidence." Journal of the American Statistical Association 112.519 (2017): 885-895. Retrieved from https://sci-hub.se/https://doi.org/10.1080/01621459.2017.1289846

[294] Lecoutre, Marie-Paule, Jacques Poitevineau, and Bruno Lecoutre. "Even statisticians are not immune to misinterpretations of Null Hypothesis Significance Tests." International Journal of Psychology 38.1 (2003): 37-45. Retrieved from https://sci-hub.se/https://doi.org/10.1080/00207590244000250

[295] Altman, Douglas G. "Statistical reviewing for medical journals." Statistics in medicine 17.23 (1998): 2661-2674. Retrieved from https://sci-hub.se/https://doi.org/10.1002/(SICI)1097-0258(19981215)17:23%3C2661::AID-SIM33%3E3.0.CO;2-B

[296] Loehlin, John C., and Robert Cosby Nichols. "HEREDITY, ENVIRONMENT, AND PERSONALITY: A STUDY OF 850 SETS OF TWINS." (1976). Retrieved from https://b-ok.cc/book/3510869/ffb3a1

[297] Scarr, Sandra, and Louise Carter-Saltzman. "Twin method: Defense of a critical assumption." Behavior genetics 9.6 (1979): 527-542. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF01067349

[298] Conley, Dalton, et al. "Heritability and the equal environments assumption: Evidence from multiple samples of misclassified twins." Behavior genetics 43.5 (2013): 415-426. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-013-9602-1

[299] Van Beijsterveldt, C. E. M., et al. "Chorionicity and heritability estimates from twin studies: The prenatal environment of twins and their resemblance across a large number of traits." Behavior Genetics 46.3 (2016): 304-314. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-015-9745-3

[300] Pesta, Bryan J., et al. "Racial and ethnic group differences in the heritability of intelligence: A systematic review and meta-analysis." Intelligence 78 (2020): 101408. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2019.101408

[301] Rose, R. J., I. A. Uchida, and J. C. Christian. "Placentation effects on cognitive resemblance of adult monozygotes." Progress in clinical and biological research 69 (1981): 35. https://www.researchgate.net/publication/16000273_Placentation_effects_on_cognitive_resemblance_of_adult_monozygotes

[302] Sundet, Jon Martin, et al. "On the question of secular trends in the heritability of intelligence test scores: A study of Norwegian twins." Intelligence 12.1 (1988): 47-59. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(88)90022-0

[303] Benyamin, Beben, et al. "Large, consistent estimates of the heritability of cognitive ability in two entire populations of 11-year-old twins from Scottish mental surveys of 1932 and 1947." Behavior genetics 35.5 (2005): 525-534. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-005-3556-x

[304] McGue, Matt, et al. "The environments of adopted and non-adopted youth: Evidence on range restriction from the Sibling Interaction and Behavior Study (SIBS)." Behavior genetics 37.3 (2007): 449-462. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-007-9142-7

[305] Protzko, John. "The environment in raising early intelligence: A meta-analysis of the fadeout effect." Intelligence 53 (2015): 202-210. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.10.006

[306] te Nijenhuis, Jan, Birthe Jongeneel-Grimen, and Elijah L. Armstrong. "Are adoption gains on the g factor? A meta-analysis." Personality and Individual Differences 73 (2015): 56-60. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2014.09.022

[307] Domingue, Benjamin W., et al. "The social genome of friends and schoolmates in the National Longitudinal Study of Adolescent to Adult Health." Proceedings of the National Academy of Sciences 115.4 (2018): 702-707. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.1711803115

[308] Bouchard Jr, Thomas J., and Matt McGue. "Genetic and environmental influences on human psychological differences." Journal of neurobiology 54.1 (2003): 4-45. Retrieved from https://sci-hub.se/https://doi.org/10.1002/neu.10160

[309] Rushton, J. Philippe, and Trudy Ann Bons. "Mate choice and friendship in twins: evidence for genetic similarity." Psychological Science 16.7 (2005): 555-559. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.0956-7976.2005.01574.x

[310] Russell, Robin JH, and Pamela A. Wells. "Personality similarity and quality of marriage." Personality and Individual Differences 12.5 (1991): 407-412. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0191-8869(91)90057-I

[311] Tesser, Abraham. "The importance of heritability in psychological research: The case of attitudes." Psychological review 100.1 (1993): 129. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0033-295X.100.1.129

[312] Caspi, Avshalom, and Ellen S. Herbener. "Continuity and change: Assortative marriage and the consistency of personality in adulthood." Journal of personality and social psychology 58.2 (1990): 250. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-3514.58.2.250

[313] Luo, Shanhong, and Eva C. Klohnen. "Assortative mating and marital quality in newlyweds: a couple-centered approach." Journal of personality and social psychology 88.2 (2005): 304. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-3514.88.2.304

[314] Mascie-Taylor, CG Nicholas, and Steven G. Vandenberg. "Assortative mating for IQ and personality due to propinquity and personal preference." Behavior Genetics 18.3 (1988): 339-345. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF01260934

[315] Escorial, Sergio, and Carmen Martín-Buro. "The role of personality and intelligence in assortative mating." The Spanish journal of psychology 15.2 (2012): 680-687. Retrieved from https://sci-hub.se/https://doi.org/10.5209/rev_SJOP.2012.v15.n2.38879

[316] Robinson, Matthew R., et al. "Genetic evidence of assortative mating in humans." Nature Human Behaviour 1.1 (2017): 1-13. Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41562-016-0016

[317] Helgason, Agnar, et al. "An association between the kinship and fertility of human couples." Science 319.5864 (2008): 813-816. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.1150232

[318] Bouchard, Thomas J. "The Wilson effect: the increase in heritability of IQ with age." Twin Research and Human Genetics 16.5 (2013): 923-930. Retrieved from https://sci-hub.se/https://doi.org/10.1017/thg.2013.54

[319] Haworth, Catherine MA, et al. "The heritability of general cognitive ability increases linearly from childhood to young adulthood." Molecular psychiatry 15.11 (2010): 1112-1120. Retrieved from https://sci-hub.se/https://doi.org/10.1038/mp.2009.55

[320] McGue, Matt, et al. "Behavioral genetics of cognitive ability: A life-span perspective." (1993). Retrieved from https://sci-hub.se/https://doi.org/10.1037/10131-003

[321] Plomin, Robert, et al. "Nature, nurture, and cognitive development from 1 to 16 years: A parent-offspring adoption study." Psychological Science 8.6 (1997): 442-447. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-9280.1997.tb00458.x

[322] Trzaskowski, M., et al. "DNA evidence for strong genetic stability and increasing heritability of intelligence from age 7 to 12." Molecular psychiatry 19.3 (2014): 380-384. Retrieved from https://sci-hub.se/https://doi.org/10.1038/mp.2012.191

[323] Peoples, Charisse E., Joseph F. Fagan III, and Dennis Drotar. "The influence of race on 3-year-old children's performance on the Stanford-Binet." Intelligence 21.1 (1995): 69-82. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(95)90039-X

[324] Krapohl, Eva, et al. "The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence." Proceedings of the national academy of sciences 111.42 (2014): 15273-15278. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.1408777111

[325] Branigan, Amelia R., Kenneth J. McCallum, and Jeremy Freese. "Variation in the heritability of educational attainment: An international meta-analysis." Social forces 92.1 (2013): 109-140. Retrieved from https://sci-hub.se/https://doi.org/10.1093/sf/sot076

[326] Hyytinen, Ari, et al. "Heritability of lifetime income." Helsinki Center of Economic Research Discussion Paper 364 (2013). Retrieved from https://sci-hub.se/https://doi.org/10.2139/ssrn.2253264

[327] Barnett, Adrian., and Doubleday, Zoe. “Meta-Research: The growth of acronyms in the scientific literature” (2020). Retrieved from https://elifesciences.org/articles/60080

[328] Murray, Charles. Income inequality and IQ. AEI Press, c/o Publisher Resources Inc., 1224 Heil Quaker Boulevard, PO Box 7001, La Vergne, TN 37086-7001, 1998. Retrieved from https://www.aei.org/wp-content/uploads/2011/10/20040302_book443.pdf

[329] Benyamin, Beben, et al. "Childhood intelligence is heritable, highly polygenic and associated with FNBP1L." Molecular psychiatry 19.2 (2014): 253-258. Retrieved from https://sci-hub.se/https://doi.org/10.1038/mp.2012.184

[330] Marioni, Riccardo E., et al. "Molecular genetic contributions to socioeconomic status and intelligence." Intelligence 44 (2014): 26-32. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2014.02.006

[331] Davies, Gail, et al. "Genome-wide association studies establish that human intelligence is highly heritable and polygenic." Molecular psychiatry 16.10 (2011): 996-1005. Retrieved from https://sci-hub.se/https://doi.org/10.1038/mp.2011.85

[332] Tucker-Drob, Elliot M., and Daniel A. Briley. "Continuity of genetic and environmental influences on cognition across the life span: A meta-analysis of longitudinal twin and adoption studies." Psychological bulletin 140.4 (2014): 949. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0035893

[333] Te Nijenhuis, Jan, Olga F. Voskuijl, and Natasja B. Schijve. "Practice and coaching on IQ tests: Quite a lot of g." International Journal of Selection and Assessment 9.4 (2001): 302-308. Retrieved from https://sci-hub.se/https://doi.org/10.1111/1468-2389.00182

[334] https://www.unz.com/author/anatoly-karlin/

[335] He, Meian, et al. "Meta-analysis of genome-wide association studies of adult height in East Asians identifies 17 novel loci." Human molecular genetics 24.6 (2015): 1791-1800. Retrieved from

https://sci-hub.se/https://doi.org/10.1093/hmg/ddu583

[336] Wood, Andrew R., et al. "Defining the role of common variation in the genomic and biological architecture of adult human height." Nature genetics 46.11 (2014): 1173-1186. Retrieved from https://sci-hub.se/https://doi.org/10.1038/ng.3097

[337] Rietveld, Cornelius A., et al. "GWAS of 126,559 individuals identifies genetic variants associated with educational attainment." science 340.6139 (2013): 1467-1471. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.1235488

[338] Okbay, Aysu, et al. "Genome-wide association study identifies 74 loci associated with educational attainment." Nature 533.7604 (2016): 539-542. Retrieved from https://sci-hub.se/https://doi.org/10.1038/nature17671

[339] Yang, Jian, et al. "Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index." Nature genetics 47.10 (2015): 1114. Retrieved from https://sci-hub.se/https://doi.org/10.1038/ng.3390

[340] Hill, W. David, et al. "Genomic analysis of family data reveals additional genetic effects on intelligence and personality." Molecular psychiatry 23.12 (2018): 2347-2362. Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41380-017-0005-1

[341] Branwen, Gwern. "Embryo selection for intelligence." (2016). Retrieved from https://www.gwern.net/Embryo-selection
Archive: https://archive.is/RrfSs

[342] Speed, Doug, et al. "Reevaluation of SNP heritability in complex human traits." Nature genetics 49.7 (2017): 986-992. Retrieved from https://sci-hub.se/https://doi.org/10.1038/ng.3865

[343] Turkheimer, Eric, et al. "Socioeconomic status modifies heritability of IQ in young children." Psychological science 14.6 (2003): 623-628. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x

[344] Gottesman, I. I., et al. "Genetics, environment, and behavior: Implications for educational policy." (1972). Retrieved from http://arthurjensen.net/wp-content/uploads/2014/06/1972-ehrman-geneticsenvironmentandbehavior.pdf

[345] Plomin, Robert, and Yulia Kovas. "Generalist genes and learning disabilities." Psychological bulletin 131.4 (2005): 592. Retrieved from https://www.sci-hub.se/https://doi.org/10.1037/0033-2909.131.4.592

[346] Davis, Oliver SP, Claire MA Haworth, and Robert Plomin. "Learning abilities and disabilities: Generalist genes in early adolescence." Cognitive neuropsychiatry 14.4-5 (2009): 312-331. Retrieved from https://www.sci-hub.se/https://doi.org/10.1080/13546800902797106

[347] Trzaskowski, Maciej, et al. "DNA evidence for strong genome-wide pleiotropy of cognitive and learning abilities." Behavior genetics 43.4 (2013): 267-273. Retrieved from https://www.sci-hub.se/https://doi.org/10.1007/s10519-013-9594-x

[348] Petrill, Stephen A. "Molarity versus modularity of cognitive functioning? A behavioral genetic perspective." Current Directions in Psychological Science 6.4 (1997): 96-99. Retrieved from https://www.sci-hub.se/https://doi.org/10.1111/1467-8721.ep11512833

[349] Turkheimer, Eric. "Three laws of behavior genetics and what they mean." Current directions in psychological science 9.5 (2000): 160-164. Retrieved from https://www.sci-hub.se/https://doi.org/10.1111/1467-8721.00084

[350] Harlaar, Nicole, et al. "Associations between reading achievement and independent reading in early elementary school: A genetically informative cross-lagged study." Child Development 82.6 (2011): 2123-2137. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-8624.2011.01658.x

[351] Hasselbalch, Ann Louise, et al. "Studies of twins indicate that genetics influence dietary intake." The Journal of nutrition 138.12 (2008): 2406-2412. Retrieved from https://sci-hub.se/https://doi.org/10.3945/jn.108.087668

[352] Maia, José AR, Martine Thomis, and Gaston Beunen. "Genetic factors in physical activity levels: a twin study." American journal of preventive medicine 23.2 (2002): 87-91. Retrieved from https://sci-hub.se/https://doi.org/10.1016/s0749-3797(02)00478-6

[353] Beunen, Gaston, and Martine Thomis. "Genetic determinants of sports participation and daily physical activity." International Journal of Obesity 23.3 (1999): S55-S63. Retrieved from https://sci-hub.se/https://doi.org/10.1038/sj.ijo.0800885

[354] Eriksson, Marit, Finn Rasmussen, and Per Tynelius. "Genetic factors in physical activity and the equal environment assumption–the Swedish young male twins study." Behavior genetics 36.2 (2006): 238-247. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-005-9018-7

[355] Jensen, Arthur R. "4. The g Beyond Factor Analysis." (1987). Retrieved from https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1006&context=buroscogpsych

[356] Rushton, J. Philippe, and Arthur R. Jensen. "The rise and fall of the Flynn effect as a reason to expect a narrowing of the Black–White IQ gap." (2010): 213-219. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2009.12.002

[357] Rijsdijk, Frühling V., P. A. Vernon, and Dorret I. Boomsma. "Application of hierarchical genetic models to Raven and WAIS subtests: a Dutch twin study." Behavior Genetics 32.3 (2002): 199-210. Retrieved from https://sci-hub.se/https://doi.org/10.1023/A:1016021128949

[358] Pedersen, Nancy Lee, et al. "A quantitative genetic analysis of cognitive abilities during the second half of the life span." Psychological Science 3.6 (1992): 346-353. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-9280.1992.tb00045.x

[359] Kan, Kees-Jan, et al. "On the nature and nurture of intelligence and specific cognitive abilities: The more heritable, the more culture dependent." Psychological Science 24.12 (2013): 2420-2428. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0956797613493292

[360] Elliott, Maxwell L., et al. "What Is the Test-Retest Reliability of Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis." Psychological Science (2020): 0956797620916786. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0956797620916786

[361] Rushton, J. Philippe, and C. Davison Ankney. "Whole brain size and general mental ability: a review." International Journal of Neuroscience 119.5 (2009): 692-732. Retrieved from https://sci-hub.se/https://doi.org/10.1080/00207450802325843

[362] Pietschnig, Jakob, et al. "Meta-analysis of associations between human brain volume and intelligence differences: How strong are they and what do they mean?." Neuroscience & Biobehavioral Reviews 57 (2015): 411-432. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.neubiorev.2015.09.017

[363] Posthuma, Daniëlle, et al. "The association between brain volume and intelligence is of genetic origin." Nature neuroscience 5.2 (2002): 83-84. Retrieved from https://sci-hub.se/https://doi.org/10.1038/nn0202-83

[364] Hagenaars, Saskia P., et al. "Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N= 112 151) and 24 GWAS consortia." Molecular psychiatry 21.11 (2016): 1624-1632. Retrieved from https://sci-hub.se/https://doi.org/10.1038/mp.2015.225

[365] Strenze, Tarmo. "Intelligence and success." Handbook of intelligence. Springer, New York, NY, 2015. 405-413. Retrieved from https://sci-hub.se/https://doi.org/10.1007/978-1-4939-1562-0_25

[366] Roth, Gerhard, and Ursula Dicke. "Evolution of the brain and intelligence." Trends in cognitive sciences 9.5 (2005): 250-257. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.tics.2005.03.005

[367] Geary, David C. "Efficiency of Mitochondrial Functioning as the Fundamental Biological Mechanism of General Intelligence (g)." (2018). Retrieved from https://www.researchgate.net/profile/David_Geary/publication/327635060_Efficiency_of_Mitochondrial_Functioning_as_the_Fundamental_Biological_Mechanism_of_General_Intelligence_g/links/5de7d714299bf10bc340245f/Efficiency-of-Mitochondrial-Functioning-as-the-Fundamental-Biological-Mechanism-of-General-Intelligence-g.pdf
The mitochondria is the powerhouse of the cell. The mitochondria is the powerhouse of the cell. The mitochondria is the powerhouse of the cell. The mitochondria is the powerhouse of the cell. The mitochondria is the powerhouse of the cell. The mitochondria is the powerhouse of the cell. The mitochondria is the powerhouse of the cell. ☺

[368] Jung, Rex E., and Richard J. Haier. "The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence." Behavioral and Brain Sciences 30.2 (2007): 135. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0140525X07001185

[369] Ritchie, Stuart J., et al. "Beyond a bigger brain: Multivariable structural brain imaging and intelligence." Intelligence 51 (2015): 47-56. https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.05.001

[370] Luders, Eileen, et al. "Neuroanatomical correlates of intelligence." Intelligence 37.2 (2009): 156-163. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2008.07.002

[371] Gregory, Michael D., et al. "Regional variations in brain gyrification are associated with general cognitive ability in humans." Current Biology 26.10 (2016): 1301-1305. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.cub.2016.03.021

[372] Penke, L., et al. "Brain white matter tract integrity as a neural foundation for general intelligence." Molecular psychiatry 17.10 (2012): 1026-1030. Retrieved from https://sci-hub.se/https://doi.org/10.1038/mp.2012.66

[373] Brans, Rachel GH, et al. "Brain plasticity and intellectual ability are influenced by shared genes." Journal of Neuroscience 30.16 (2010): 5519-5524. Retrieved from https://sci-hub.se/https://doi.org/10.1523/JNEUROSCI.5841-09.2010

[374] Haier, Richard J., et al. "Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography." Intelligence 12.2 (1988): 199-217. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(88)90016-5

[375] Neubauer, Aljoscha C., and Andreas Fink. "Intelligence and neural efficiency." Neuroscience & Biobehavioral Reviews 33.7 (2009): 1004-1023. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.neubiorev.2009.04.001

[376] Stanovich, Keith E., Richard F. West, and Maggie E. Toplak. The rationality quotient: Toward a test of rational thinking. MIT press, 2016. Retrieved from https://b-ok.cc/book/2862073/35db7b

[377] McDaniel, Michael A., et al. "Use of situational judgment tests to predict job performance: a clarification of the literature." Journal of Applied Psychology 86.4 (2001): 730. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0021-9010.86.4.730

[378] Rhodes, Nancy, and Wendy Wood. "Self-esteem and intelligence affect influenceability: The mediating role of message reception." Psychological bulletin 111.1 (1992): 156. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0033-2909.111.1.156

[379] Andersson, Ola, et al. "Risk aversion relates to cognitive ability: Preferences or noise?." Journal of the European Economic Association 14.5 (2016): 1129-1154. Retrieved from https://sci-hub.se/https://doi.org/10.1111/jeea.12179

[380] Levine, Ross, and Yona Rubinstein. "Smart and illicit: who becomes an entrepreneur and do they earn more?." The Quarterly Journal of Economics 132.2 (2017): 963-1018. Retrieved from https://sci-hub.se/https://doi.org/10.3386/w19276

[381] Judge, Timothy A., Amy E. Colbert, and Remus Ilies. "Intelligence and leadership: a quantitative review and test of theoretical propositions." Journal of Applied Psychology 89.3 (2004): 542. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0021-9010.89.3.542

[382] Calvin, Catherine M., et al. "Intelligence in youth and all-cause-mortality: systematic review with meta-analysis." International journal of epidemiology 40.3 (2011): 626-644. Retrieved from https://sci-hub.se/https://doi.org/10.1093/ije/dyq190

[383] Arden, Rosalind, et al. "The association between intelligence and lifespan is mostly genetic." International Journal of Epidemiology 45.1 (2016): 178-185. Retrieved from https://sci-hub.se/https://doi.org/10.1093/ije/dyv112

[384] Nyborg, Helmuth, ed. The scientific study of general intelligence: Tribute to Arthur Jensen. Elsevier, 2003. Retrieved from http://emilkirkegaard.dk/en/wp-content/uploads/Helmuth_Nyborg_The_Scientific_Study_of_General_IBookos.org_.pdf

[385] Schwartz, Joseph A., et al. "Intelligence and criminal behavior in a total birth cohort: An examination of functional form, dimensions of intelligence, and the nature of offending." Intelligence 51 (2015): 109-118. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.06.001

[386] Frisell, Thomas, Yudi Pawitan, and Niklas Långström. "Is the association between general cognitive ability and violent crime caused by family-level confounders?." PloS one 7.7 (2012): e41783. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pone.0041783

[387] Levine, Stephen Z. "Elaboration on the association between IQ and parental SES with subsequent crime." Personality and Individual Differences 50.8 (2011): 1233-1237. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2011.02.016

[388] Yun, Ilhong, and Julak Lee. "IQ and delinquency: The differential detection hypothesis revisited." Youth violence and juvenile justice 11.3 (2013): 196-211. Retrieved from https://sci-hub.se/https://doi.org/10.1177/1541204012463410

[389] Moffitt, Terrie E., and Phil A. Silva. "IQ and delinquency: A direct test of the differential detection hypothesis." Journal of abnormal psychology 97.3 (1988): 330. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0021-843X.97.3.330

[390] Fergusson, David M., L. John Horwood, and Elizabeth M. Ridder. "Show me the child at seven II: Childhood intelligence and later outcomes in adolescence and young adulthood." Journal of Child Psychology and Psychiatry 46.8 (2005): 850-858. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1469-7610.2005.01472.x

[391] Poropat, Arthur E. "A meta-analysis of the five-factor model of personality and academic performance." Psychological bulletin 135.2 (2009): 322. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0014996

[392] Mathias, Samuel R., et al. "Minimal Relationship between Local Gyrification and General Cognitive Ability in Humans." Cerebral Cortex 30.6 (2020): 3439-3450. Retrieved from https://sci-hub.se/https://doi.org/10.1093/cercor/bhz319

[393] Hunter, John E., and Ronda F. Hunter. "Validity and utility of alternative predictors of job performance." Psychological bulletin 96.1 (1984): 72. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0033-2909.96.1.72

[394] Roth, Philip L., Philip Bobko, and Lynn A. McFarland. "A meta-analysis of work sample test validity: Updating and integrating some classic literature." Personnel Psychology 58.4 (2005): 1009-1037. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1744-6570.2005.00714.x

[395] Colquitt, Jason A., Jeffrey A. LePine, and Raymond A. Noe. "Toward an integrative theory of training motivation: a meta-analytic path analysis of 20 years of research." Journal of applied psychology 85.5 (2000): 678. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0021-9010.85.5.678

[396] Kuncel, Nathan R., Sarah A. Hezlett, and Deniz S. Ones. "Academic performance, career potential, creativity, and job performance: Can one construct predict them all?." Journal of personality and social psychology 86.1 (2004): 148. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-3514.86.1.148

[397] Schmitt, Neal, et al. "Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics." personnel psychology 37.3 (1984): 407-422. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1744-6570.1984.tb00519.x

[398] Berry, Christopher M., Paul R. Sackett, and Richard N. Landers. "Revisiting interview–cognitive ability relationships: Attending to specific range restriction mechanisms in meta-analysis." Personnel Psychology 60.4 (2007): 837-874. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1744-6570.2007.00093.x

[399] Ron Nelson, J., Gregory J. Benner, and Jorge Gonzalez. "Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review." Learning Disabilities Research & Practice 18.4 (2003): 255-267. Retrieved from https://sci-hub.se/https://doi.org/10.1111/1540-5826.00080

[400] Lynn, Richard. Dysgenics: Genetic deterioration in modern populations. Vol. 1. No. 4. Westport, CT: Praeger, 1996. Retrieved from https://www.gwern.net/docs/genetics/selection/1996-lynn-dysgenics.pdf

[401] Lopez, Carolina, Daniel Stahl, and Kate Tchanturia. "Estimated intelligence quotient in anorexia nervosa: a systematic review and meta-analysis of the literature." Annals of General Psychiatry 9.1 (2010): 40. Retrieved from https://sci-hub.se/https://doi.org/10.1186/1744-859X-9-40

[402] Mann, Richard D. "A review of the relationships between personality and performance in small groups." Psychological bulletin 56.4 (1959): 241. Retrieved from https://sci-hub.se/https://doi.org/10.1037/h0044587

[403] Kim, Kyung Hee. "Can only intelligent people be creative? A meta-analysis." Journal of Secondary Gifted Education 16.2-3 (2005): 57-66. Retrieved from https://sci-hub.se/https://doi.org/10.4219/jsge-2005-473

[404] DeNeve, Kristina M., and Harris Cooper. "The happy personality: A meta-analysis of 137 personality traits and subjective well-being." Psychological bulletin 124.2 (1998): 197. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0033-2909.124.2.197

[405] Steel, Piers. "The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure." Psychological bulletin 133.1 (2007): 65. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0033-2909.133.1.65

[406] Griffeth, Rodger W., Peter W. Hom, and Stefan Gaertner. "A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium." Journal of management 26.3 (2000): 463-488. Retrieved from https://sci-hub.se/https://doi.org/10.1177/014920630002600305

[407] Feingold, Alan. "Good-looking people are not what we think." Psychological bulletin 111.2 (1992): 304. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0033-2909.111.2.304

[408] Gendreau, Paul, Tracy Little, and Claire Goggin. "A meta-analysis of the predictors of adult offender recidivism: What works!." Criminology 34.4 (1996): 575-608. https://sci-hub.se/https://doi.org/10.1111/j.1745-9125.1996.tb01220.x

[409] Arthur, Winfred, Gerald V. Barret, and Ralph A. Alexander. "Prediction of vehicular accident involvement: A meta-analysis." Human performance 4.2 (1991): 89-105. Retrieved from https://sci-hub.se/https://doi.org/10.1207/s15327043hup0402_1

[410] Woodberry, Kristen A., Anthony J. Giuliano, and Larry J. Seidman. "Premorbid IQ in schizophrenia: a meta-analytic review." American Journal of Psychiatry 165.5 (2008): 579-587. Retrieved from https://sci-hub.se/https://doi.org/10.1176/appi.ajp.2008.07081242

[411] Bourhis, John, and Mike Allen. "Meta-analysis of the relationship between communication apprehension and cognitive performance." Communication Education 41.1 (1992): 68-76. Retrieved from https://sci-hub.se/https://doi.org/10.1080/03634529209378871

[412] Dalliard. IQ and Permanent Income: Sizing Up the “IQ Paradox”. Human Varieties. (2016) Retrieved from https://humanvarieties.org/2016/01/31/iq-and-permanent-income-sizing-up-the-iq-paradox/

[413] Ree, Malcolm James, James A. Earles, and Mark S. Teachout. "Predicting job performance: Not much more than g.." Journal of Applied Psychology 79.4 (1994): 518. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0021-9010.79.4.518

[414] Bertua, Cristina, Neil Anderson, and Jesús F. Salgado. "The predictive validity of cognitive ability tests: A UK meta-analysis." Journal of Occupational and Organizational psychology 78.3 (2005): 387-409. Retrieved from https://sci-hub.se/https://doi.org/10.1348/096317905X26994

[415] Postlethwaite, Bennett Eugene. "Fluid ability, crystallized ability, and performance across multiple domains: a meta-analysis." (2011). Retrieved from https://ir.uiowa.edu/cgi/viewcontent.cgi?article=2639&context=etd

[416] Te Nijenhuis, Jan, Olga F. Voskuijl, and Natasja B. Schijve. "Practice and coaching on IQ tests: Quite a lot of g." International Journal of Selection and Assessment 9.4 (2001): 302-308. Retrieved from https://sci-hub.se/https://doi.org/10.1111/1468-2389.00182

[417] Krapohl, Eva, and Robert Plomin. "Genetic link between family socioeconomic status and children’s educational achievement estimated from genome-wide SNPs." Molecular psychiatry 21.3 (2016): 437-443. Retrieved from https://sci-hub.se/https://doi.org/10.1038/mp.2015.2

[418] Chalfin, Aaron, et al. "Police Force Size and Civilian Race." NBER Working Paper w28202 (2020). Retrieved from https://www.nber.org/papers/w28202?utm_campaign=ntwh&utm_medium=email&utm_source=ntwg24

[419] Colom, Roberto, and Carmen E. Flores-Mendoza. "Intelligence predicts scholastic achievement irrespective of SES factors: Evidence from Brazil." Intelligence 35.3 (2007): 243-251. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2006.07.008

[420] Gonzalez, Maria, and Lisa A. Jones. "Our Struggle Is My Struggle: Solidarity Feminism as an Intersectional Reply to Neoliberal and Choice Feminism." Retrieved from https://pdfs.semanticscholar.org/2212/3ee1e882051a50be8c467abe641932c156a0.pdf

[421] O'Connell, Michael. "The power of cognitive ability in explaining educational test performance, relative to other ostensible contenders." Intelligence 66 (2018): 122-127. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2017.11.011

[422] Thienpont, Kristiaan, and Gino Verleye. "Cognitive ability and occupational status in a British cohort." Journal of Biosocial Science 36.3 (2004): 333. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0021932003006229

[423] Kemp, Leslie CD. "Environmental and other characteristics determining attainment in primary schools." British Journal of Educational Psychology 25.2 (1955): 67-77. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.2044-8279.1955.tb01339.x

[424] Saunders, Peter. "Social mobility in Britain: an empirical evaluation of two competing explanations." Sociology 31.2 (1997): 261-288. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0038038597031002005

[425] Hegelund, Emilie Rune, et al. "The influence of familial factors on the association between IQ and educational and occupational achievement: A sibling approach." Personality and Individual Differences 149 (2019): 100-107. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2019.05.045

[426] Schmidt, Frank L., and John E. Hunter. "The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings." Psychological bulletin 124.2 (1998): 262. https://sci-hub.se/https://doi.org/10.1037/0033-2909.124.2.262

[427] https://areomagazine.com/2018/10/02/academic-grievance-studies-and-the-corruption-of-scholarship/

[428] Fryer Jr, Roland G., Steven D. Levitt, and John A. List. "Exploring the impact of financial incentives on stereotype threat: Evidence from a pilot study." American Economic Review 98.2 (2008): 370-75. Retrieved from https://sci-hub.se/https://doi.org/10.1257/aer.98.2.370

[429] Stricker, Lawrence J., and William C. Ward. "Stereotype Threat, Inquiring About Test Takers' Ethnicity and Gender, and Standardized Test Performance 1." Journal of Applied Social Psychology 34.4 (2004): 665-693. Retrieved from https://sci-hub.se/10.1111/j.1559-1816.2004.tb02564.x/full

[430] Wei, Thomas E. Stereotype threat, gender, and math performance: Evidence from the national assessment of educational progress. Working Paper, Harvard University, 2008. Retrieved from https://pdfs.semanticscholar.org/61dd/a726d044ce880fd7eb98f2e9c3380780e3fe.pdf

[431] Shewach, Oren R., Paul R. Sackett, and Sander Quint. "Stereotype threat effects in settings with features likely versus unlikely in operational test settings: A meta-analysis." Journal of Applied Psychology 104.12 (2019): 1514. Retrieved from https://www.gwern.net/docs/psychology/2019-shewach.pdf

[432] Flore, Paulette. "Stereotype threat and differential item functioning: A critical assessment." (2018). Retrieved from https://research.tilburguniversity.edu/files/23445144/Flore_Stereotype_7_3_2018.pdf

[433] Flore, Paulette C., and Jelte M. Wicherts. "Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis." Journal of school psychology 53.1 (2015): 25-44. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jsp.2014.10.002

[434] Flore, Paulette C., Joris Mulder, and Jelte M. Wicherts. "The influence of gender stereotype threat on mathematics test scores of Dutch high school students: A registered report." Comprehensive Results in Social Psychology 3.2 (2018): 140-174. Retrieved from https://sci-hub.se/https://doi.org/10.1080/23743603.2018.1559647

[435] Ganley, Colleen M., et al. "An examination of stereotype threat effects on girls' mathematics performance." Developmental psychology 49.10 (2013): 1886. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0031412

[436] Stricker, Lawrence J. "The challenge of stereotype threat for the testing community." Presidential address to the division of evaluation, measurement, and statistics. 2007 american educational research association annual meeting. 2008. Retrieved from https://www.ets.org/Media/Research/pdf/RM-08-12.pdf

[437] Stricker, Lawrence J., and Isaac I. Bejar. "Test Difficulty and Stereotype Threat on the GRE® General Test." ETS Research Report Series 1999.2 (1999): i-47. Retrieved from https://sci-hub.se/https://doi.org/10.1002/j.2333-8504.1999.tb01817.x

[438] Appel, Markus, Silvana Weber, and Nicole Kronberger. "The influence of stereotype threat on immigrants: Review and meta-analysis." Frontiers in Psychology 6 (2015): 900. Retrieved from https://sci-hub.se/https://doi.org/10.3389/fpsyg.2015.00900

[439] ISIR program 2009. Retrieved from https://www.isironline.org/wp-content/uploads/2014/05/program2009.pdf

[440] Wicherts' 2014 CV https://web.archive.org/web/20140905190054/http://wicherts.socsci.uva.nl/CVJMW.pdf

[441] Kirkegaard 2018 "Jelte Wicherts' lost stereotype threat study for African Americans" https://emilkirkegaard.dk/en/?p=7534

[442] Wicherts' 2011 CV https://web.archive.org/web/20110813150923/http://wicherts.socsci.uva.nl/CVJMW.pdf

[443] Forscher, Patrick S., et al. "A multi-site examination of stereotype threat in black college students across varying operationalizations." (2019). Retrieved from https://psyarxiv.com/6hju9/

[444] Media Research Center https://emilkirkegaard.dk/en/wp-content/uploads/The-LIberal-Media-Exposed.pdf

[445] https://web.archive.org/web/20200310155733/https://www.nbcnews.com/think/opinion/are-white-people-jokes-racist-let-fellow-white-person-explain-ncna899981

[446] https://archive.is/222Zd

[447] https://archive.is/LZKzz

[448] https://archive.is/K0tIg

[449] https://archive.is/YRNVD

[450] https://archive.is/pPoR5

[451] https://archive.is/at54V

[452] https://archive.is/AWHAK

[453] https://archive.is/yjhMn

[454] https://archive.is/7X5KD

[455] Uhlmann, Eric L., et al. "The motivated use of moral principles." Judgment and Decision making 4.6 (2009): 479-491. Retrieved from http://journal.sjdm.org/9616/jdm9616.pdf

[456] https://archive.is/aULIS

[457] https://archive.is/yGx6l

[458] https://archive.is/bUQQz

[459] https://archive.is/PVHNV

[460] Winegard, Bo & Clark, Cory & Bunnel, Ethan. (2019). The Ideology of Censorship. Retrieved from https://www.researchgate.net/publication/333677484_The_Ideology_of_Censorship

[461] Cooley, E., Brown-Iannuzzi, J. L., Lei, R. F., & Cipolli, W. III. (2019). Complex intersections of race and class: Among social liberals, learning about White privilege reduces sympathy, increases blame, and decreases external attributions for White people struggling with poverty. Journal of Experimental Psychology: General, 148(12), 2218–2228. https://sci-hub.se/https://doi.org/10.1037/xge0000605

[462] Tetlock, Philip E., et al. "The psychology of the unthinkable: taboo trade-offs, forbidden base rates, and heretical counterfactuals." Journal of personality and social psychology 78.5 (2000): 853. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-3514.78.5.853

[463] Heiphetz, Larisa A. "Dehumanization and perceptions of immoral intergroup behavior." (2020). Retrieved from https://academiccommons.columbia.edu/doi/10.7916/d8-z3j5-9g62/download

[464] Dupree, C. H., & Fiske, S. T. (2019). Self-presentation in interracial settings: The competence downshift by White liberals. Journal of Personality and Social Psychology, 117(3), 579–604. Retrieved from https://www.americanvoiceforfreedom.org/wp-content/uploads/2018/12/Competence-Downshift-by-White-Liberals.pdf

[465] Schlenker, Barry R., John R. Chambers, and Bonnie M. Le. "Conservatives are happier than liberals, but why? Political ideology, personality, and life satisfaction." Journal of Research in Personality 46.2 (2012): 127-146. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jrp.2011.12.009

[466] Belmi, Peter, and Margaret Neale. "Mirror, mirror on the wall, who's the fairest of them all? Thinking that one is attractive increases the tendency to support inequality." Organizational Behavior and Human Decision Processes 124.2 (2014): 133-149. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.obhdp.2014.03.002

[467] Peterson, Rolfe Daus, and Carl L. Palmer. "Effects of physical attractiveness on political beliefs." Politics and the Life Sciences 36.2 (2017): 3-16. Retrieved from https://sci-hub.se/https://doi.org/10.1017/pls.2017.18

[468] Berggren, Niclas, Henrik Jordahl, and Panu Poutvaara. "The right look: Conservative politicians look better and voters reward it." Journal of Public Economics 146 (2017): 79-86. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jpubeco.2016.12.008

[469] Bardis, Panos D. "GUILT FEELINGS AMONG WHITE COLLEGE STUDENTS CONCERNING NEGRO SOCIAL INEQUALITY." International Review of Modern Sociology (1973): 168-182. Retrieved from https://sci-hub.se/https://www.jstor.org/stable/41420494

[470] Swim, Janet K., and Deborah L. Miller. "White guilt: Its antecedents and consequences for attitudes toward affirmative action." Personality and Social Psychology Bulletin 25.4 (1999): 500-514. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0146167299025004008

[471] Case, Kim A. "Raising white privilege awareness and reducing racial prejudice: Assessing diversity course effectiveness." Teaching of Psychology 34.4 (2007): 231-235. Retrieved from https://sci-hub.se/https://doi.org/10.1080/00986280701700250

[472] Powell, Adam A., Nyla R. Branscombe, and Michael T. Schmitt. "Inequality as ingroup privilege or outgroup disadvantage: The impact of group focus on collective guilt and interracial attitudes." Personality and Social Psychology Bulletin 31.4 (2005): 508-521. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0146167204271713

[473] Roberts, Robert E., et al. "The structure of ethnic identity of young adolescents from diverse ethnocultural groups." The Journal of Early Adolescence 19.3 (1999): 301-322. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0272431699019003001

[474] Carlson, Cindy, Sarika Uppal, and Ellie C. Prosser. "Ethnic differences in processes contributing to the self-esteem of early adolescent girls." The Journal of Early Adolescence 20.1 (2000): 44-67. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0272431600020001003

[475] Horowitz, M, J., Brown, A., & Cox, K. (2019). “The role of race and ethnicity in Americans’ personal lives” Pew Research Center. Retrieved from https://www.pewsocialtrends.org/2019/04/09/the-role-of-race-and-ethnicity-in-americans-personal-lives/

[476] Phinney, J.S., Cantu, C.L. & Kurtz, D.A. Ethnic and American Identity as Predictors of Self-Esteem Among African American, Latino, and White Adolescents. Journal of Youth and Adolescence 26, 165–185 (1997). Retrieved from https://sci-hub.se/https://doi.org/10.1023/A:1024500514834

[477] Phinney, Jean S. "The multigroup ethnic identity measure: A new scale for use with diverse groups." Journal of adolescent research 7.2 (1992): 156-176. Retrieved from https://sci-hub.se/https://doi.org/10.1177/074355489272003

[478] Zigerell, L. J. "Black and White discrimination in the United States: Evidence from an archive of survey experiment studies." Research & Politics 5.1 (2018): 2053168017753862. Retrieved from https://sci-hub.se/https://doi.org/10.1177/2053168017753862

[479] Forscher, Patrick S., et al. "A meta-analysis of procedures to change implicit measures." Journal of Personality and Social Psychology 117.3 (2019): 522. Retrieved from https://psyarxiv.com/dv8tu/

[480] Jencks, Christopher. "Inequality: A reassessment of the effect of family and schooling in America." (1972). Retrieved from https://b-ok.cc/book/3605784/4fce41

[481] Bouchard Jr, Thomas J. "Do environmental similarities explain the similarity in intelligence of identical twins reared apart?." Intelligence 7.2 (1983): 175-184. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(83)90027-2

[482] Bouchard, Thomas J., et al. "Sources of human psychological differences: The Minnesota study of twins reared apart." Science 250.4978 (1990): 223-228. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.2218526

[483] Rose, Richard J., and Jaakko Kaprio. "Frequency of social contact and intrapair resemblance of adult monozygotic cotwins—or does shared experience influence personality after all?." Behavior Genetics 18.3 (1988): 309-328. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF01260932

[484] Kaprio, Jaakko, Markku Koskenvuo, and Richard J. Rose. "Change in cohabitation and intrapair similarity of monozygotic (MZ) cotwins for alcohol use, extraversion, and neuroticism." Behavior Genetics 20.2 (1990): 265-276. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF01067794

[485] Lykken, David T., et al. "Does contact lead to similarity or similarity to contact?." Behavior Genetics 20.5 (1990): 547-561. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF01065871

[486] Barnes, J. C., et al. "Demonstrating the validity of twin research in criminology." Criminology 52.4 (2014): 588-626. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF01065871

[487] Wright, John Paul, et al. "Mathematical proof is not minutiae and irreducible complexity is not a theory: A final response to Burt and Simons and a call to criminologists." Criminology 53 (2015): 113. Retrieved from https://sci-hub.se/https://doi.org/10.1111/1745-9125.12059

[488] Barnes, J. C., and Brian B. Boutwell. "A demonstration of the generalizability of twin-based research on antisocial behavior." Behavior Genetics 43.2 (2013): 120-131. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-012-9580-8

[489] Trut, Lyudmila N. "Early Canid Domestication: The Farm-Fox Experiment: Foxes bred for tamability in a 40-year experiment exhibit remarkable transformations that suggest an interplay between behavioral genetics and development." American Scientist 87.2 (1999): 160-169. Retrieved from https://sci-hub.se/https://doi.org/10.2307/27857815

[490] Polderman, Tinca JC, et al. "Meta-analysis of the heritability of human traits based on fifty years of twin studies." Nature genetics 47.7 (2015): 702-709. Retrieved from https://sci-hub.se/https://doi.org/10.1038/ng.3285

[491] Schwabe, Inga, Luc Janss, and Stéphanie M. Van Den Berg. "Can we validate the results of twin studies? A census-based study on the heritability of educational achievement." Frontiers in genetics 8 (2017): 160. Retrieved from https://sci-hub.se/https://doi.org/10.3389/fgene.2017.00160

[492] Plomin, Robert, and Frank M. Spinath. "Intelligence: genetics, genes, and genomics." Journal of personality and social psychology 86.1 (2004): 112. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-3514.86.1.112

[493] Panizzon, Matthew S., et al. "Genetic and environmental influences on general cognitive ability: Is ga valid latent construct?." Intelligence 43 (2014): 65-76. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2014.01.008

[494] Christensen, Alexander P., et al. "Clever people: Intelligence and humor production ability." Psychology of Aesthetics, Creativity, and the Arts 12.2 (2018): 136. Retrieved from https://sci-hub.se/https://doi.org/10.1037/aca0000109

[495] T NAGOSHI, C. R. A. I. G., and Ronald C. Johnson. "Socioeconomic status does not moderate the familiality of cognitive abilities in the Hawaii Family Study of Cognition." Journal of Biosocial Science 37.6 (2005): 773. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0021932004007023

[496] Haworth, Claire MA, et al. "A twin study of the genetics of high cognitive ability selected from 11,000 twin pairs in six studies from four countries." Behavior genetics 39.4 (2009): 359-370. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s%2010519-009-9262-3

[497] Grant, Michael D., et al. "Does parental education have a moderating effect on the genetic and environmental influences of general cognitive ability in early adulthood?." Behavior genetics 40.4 (2010): 438-446. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-010-9351-3

[498] Kirkpatrick, Robert M., Matt McGue, and William G. Iacono. "Replication of a gene–environment interaction via multimodel inference: additive-genetic variance in Adolescents' General Cognitive Ability Increases with Family-of-Origin Socioeconomic Status." Behavior genetics 45.2 (2015): 200-214. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-014-9698-y

[499] Turkheimer, Eric, Christopher E. Beam, and Deborah W. Davis. "The Scarr-Rowe interaction in complete seven-year WISC data from the Louisville twin study: Preliminary report." Behavior genetics 45.6 (2015): 635-639. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-015-9760-4

[500] Little, Callie W., Rasheda Haughbrook, and Sara A. Hart. "Cross-study differences in the etiology of reading comprehension: A meta-analytical review of twin studies." Behavior genetics 47.1 (2017): 52-76. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-016-9810-6

[501] Hart, Sara A., Stephen A. Petrill, and Claire M. Kamp Dush. "Genetic influences on language, reading, and mathematics skills in a national sample: An analysis using the National Longitudinal Survey of Youth." Language, speech, and hearing services in schools (2010). Retrieved from https://sci-hub.se/https://doi.org/10.1044/0161-1461(2009/08-0052)

[502] Zaboski II, Brian A., John H. Kranzler, and Nicholas A. Gage. "Meta-analysis of the relationship between academic achievement and broad abilities of the Cattell-Horn-Carroll theory." Journal of school psychology 71 (2018): 42-56. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jsp.2018.10.001

[503] Woodley of Menie, Michael A., Jonatan Pallesen, and Matthew A. Sarraf. "Evidence for the Scarr–Rowe Effect on Genetic Expressivity in a Large U.S. Sample." Twin Research and Human Genetics 21.6 (2018): 495-501. Retrieved from https://sci-hub.se/https://doi.org/10.1017/thg.2018.63

[504] Thurstone, Louis Leon. Primary mental abilities. Vol. 119. Chicago: University of Chicago Press, 1938. https://www.amazon.com/Primary-Mental-Abilities-Louis-Thurstone/dp/0226801136

[505] Eysenck, H. J. "Primary mental abilities." British Journal of Educational Psychology 9.3 (1939): 270-275. Retrieved from https://sci-hub.se/https://doi.org/10.1007/978-94-011-6129-9_8

[506] Spearman, C. "Thurstone's work re-worked." Journal of Educational Psychology 30.1 (1939): 1. Retrieved from https://sci-hub.se/https://doi.org/10.1037/h0061267

[507] Guilford, Joy P. "Zero correlations among tests of intellectual abilities." Psychological Bulletin 61.6 (1964): 401-404. Retrieved from https://sci-hub.se/https://doi.org/10.1037/h0048576

[508] Alliger, George M. "Do zero correlations really exist among measures of different intellectual abilities?." Educational and Psychological Measurement 48.2 (1988): 275-280. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0013164488482001

[509] Elliott, Colin D. "The factorial structure and specificity of the British Ability Scales." British Journal of Psychology 77.2 (1986): 175-185. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.2044-8295.1986.tb01992.x

[510] Keith, Timothy Z., John H. Kranzler, and Dawn P. Flanagan. "What does the Cognitive Assessment System (CAS) measure? Joint confirmatory factor analysis of the CAS and the Woodcock-Johnson Tests of Cognitive Ability." School Psychology Review 30.1 (2001): 89-119. Retrieved from https://www.researchgate.net/profile/John_Kranzler/publication/284482210_What_does_the_Cognitive_Assessment_System_CAS_measure_Joint_confirmatory_factor_analysis_of_the_CAS_and_the_Woodcock-Johnson_Tests_of_Cognitive_Ability/links/589b7ae792851c942ddae238/What-does-the-Cognitive-Assessment-System-CAS-measure-Joint-confirmatory-factor-analysis-of-the-CAS-and-the-Woodcock-Johnson-Tests-of-Cognitive-Ability.pdf

[511] Brody, Nathan. "Construct validation of the Sternberg Triarchic abilities test: Comment and reanalysis." Intelligence 31.4 (2003): 319-329. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0160-2896(01)00087-3

[512] Humphreys, Lloyd G., Susan A. Rich, and Timothy C. Davey. "A Piagetian test of general intelligence." Developmental Psychology 21.5 (1985): 872. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0012-1649.21.5.872

[513] Devlin, Bernie, et al., eds. Intelligence, genes, and success: Scientists respond to The Bell Curve. Springer Science & Business Media, 1997. Retrieved from https://b-ok.cc/book/2296418/a43343

[514] Gustafsson, Jan-Eric. "A unifying model for the structure of intellectual abilities." Intelligence 8.3 (1984): 179-203. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(84)90008-4

[515] Woodcock, Richard W. "Theoretical foundations of the WJ-R measures of cognitive ability." Journal of Psychoeducational Assessment 8.3 (1990): 231-258. Retrieved from https://sci-hub.se/https://doi.org/10.1177/073428299000800303

[516] Carroll, John B. "The higher-stratum structure of cognitive abilities: Current evidence supports g and about ten broad factors." The scientific study of general intelligence. Pergamon, 2003. 5-21. Retrieved from https://sci-hub.se/https://doi.org/10.1016/B978-008043793-4/50036-2

[517] Gardner, Howard. Frames of mind: The theory of multiple intelligences. Hachette Uk, 2011. Retrieved from https://b-ok.cc/book/766868/b9903e

[518] Cucina, Jeffrey M., et al. "Role of mental abilities and mental tests in explaining high-school grades." Intelligence 54 (2016): 90-104. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.11.007

[519] Gardner, Howard. "Frequently asked questions—multiple intelligences and related educational topics." Retrieved February 9 (2004): 2008. Retrieved from https://howardgardner01.files.wordpress.com/2012/06/faq_march2013.pdf

[520] Waterhouse, Lynn. "Multiple intelligences, the Mozart effect, and emotional intelligence: A critical review." Educational Psychologist 41.4 (2006): 207-225. Retrieved from https://sci-hub.se/https://doi.org/10.1207/s15326985ep4104_1

[521] Visser, Beth A., Michael C. Ashton, and Philip A. Vernon. "Beyond g: Putting multiple intelligences theory to the test." Intelligence 34.5 (2006): 487-502. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2006.02.004

[522] Gardner, Howard, and Seana Moran. "The science of multiple intelligences theory: A response to Lynn Waterhouse." Educational psychologist 41.4 (2006): 227-232. Retrieved from https://emilkirkegaard.dk/en/wp-content/uploads/The-Science-of-Multiple-Intelligences-Theory-A-Response-to-Lynn-Waterhouse.pdf

[523] Rauscher, Frances H., and Sean C. Hinton. "The Mozart effect: Music listening is not music instruction." Educational Psychologist 41.4 (2006): 233-238. Retrieved from https://emilkirkegaard.dk/en/wp-content/uploads/The-Mozart-Effect-Music-Listening-is-Not-Music-Instruction.pdf

[524] Cherniss, Cary, et al. "Emotional intelligence: what does the research really indicate?." Educational Psychologist 41.4 (2006): 239-245. Retrieved from https://emilkirkegaard.dk/en/wp-content/uploads/Emotional-Intelligence-What-Does-the-Research-Really-Indicate.pdf

[525] Waterhouse, Lynn. "Inadequate evidence for multiple intelligences, Mozart effect, and emotional intelligence theories." Educational Psychologist 41.4 (2006): 247-255. Retrieved from https://emilkirkegaard.dk/en/wp-content/uploads/Inadequate-evidence-for-Multiple-Intelligences-Mozart-Effect-and-Emotional-Intelligence-Theories.pdf

[526] Alexander, Patricia A. "Evolution of a learning theory: A case study." Educational Psychologist 41.4 (2006): 257-264. Retrieved from https://emilkirkegaard.dk/en/wp-content/uploads/Evolution-of-a-Learning-Theory-A-Case-Study.pdf

[527] Gardner, Howard. "On failing to grasp the core of MI theory: A response to Visser et al." Intelligence 34.5 (2006): 503-505. Retrieved from https://emilkirkegaard.dk/en/wp-content/uploads/2.-On-failing-to-grasp-the-core-of-MI-theory-A-response-to-Visser-et-al..pdf

[528] Visser, Beth A., Michael C. Ashton, and Philip A. Vernon. "g and the measurement of Multiple Intelligences: A response to Gardner." Intelligence 34.5 (2006): 507-510. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2006.04.006

[529] Schulte, Melanie J., Malcolm James Ree, and Thomas R. Carretta. "Emotional intelligence: Not much more than g and personality." Personality and individual differences 37.5 (2004): 1059-1068. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2003.11.014

[530] O'Boyle Jr, Ernest H., et al. "The relation between emotional intelligence and job performance: A meta-analysis." Journal of Organizational Behavior 32.5 (2011): 788-818. Retrieved from https://sci-hub.se/https://doi.org/10.1002/job.714

[531] Paul, S. M. (1980). Sibling Resemblance in Mental Ability: A Review. Behavior Genetics, 10, 277–290. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF01067773

[532] Wilson, Ronald S. "Synchronies in mental development: An epigenetic perspective." Science 202.4371 (1978): 939-948. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.568822

[533] Teasdale, T. W., and David R. Owen. "Heredity and familial environment in intelligence and educational level—a sibling study." Nature 309.5969 (1984): 620-622. Retrieved from https://sci-hub.se/https://doi.org/10.1038/309620a0

[534] Visscher, Peter M., et al. "Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings." *PLoS Genet* 2.3 (2006): e41. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pgen.0020041

[535] Segal, Nancy L., and Francisca J. Niculae. "Fullerton Virtual Twin Project: Overview and 2019 Update." *Twin Research and Human Genetics* 22.6 (2019): 731-734. Retrieved from https://sci-hub.se/https://doi.org/10.1017/thg.2019.40

[536] Ritchie, Stuart J., Timothy C. Bates, and Ian J. Deary. "Is education associated with improvements in general cognitive ability, or in specific skills?." Developmental psychology 51.5 (2015): 573. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0038981

[537] Posth, Cosimo, et al. "Deeply divergent archaic mitochondrial genome provides lower time boundary for African gene flow into Neanderthals." Nature communications 8.1 (2017): 1-9. Retrieved from https://sci-hub.se/https://doi.org/10.1038/ncomms16046

[538] Hershkovitz, Israel, et al. "The earliest modern humans outside Africa." Science 359.6374 (2018): 456-459. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.aap8369

[539] Lindqvist, Charlotte, et al. "Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of polar bear." Proceedings of the National Academy of Sciences 107.11 (2010): 5053-5057. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.0914266107

[540] Mikko, Sofia, and Leif Andersson. "Low major histocompatibility complex class II diversity in European and North American moose." Proceedings of the National Academy of Sciences 92.10 (1995): 4259-4263. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.92.10.4259

[541] Wilson, Robert E., et al. "Speciation, subspecies divergence, and paraphyly in the Cinnamon Teal and Blue-winged Teal." The Condor 113.4 (2011): 747-761. Retrieved from https://sci-hub.se/https://doi.org/10.1525/cond.2011.110042

[542] Luo, Shu-Jin, et al. "Phylogeography and genetic ancestry of tigers (Panthera tigris)." PLoS Biol 2.12 (2004): e442. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pbio.0020442

[543] Peters, Jeffrey L., William Gretes, and Kevin E. Omland. "Late Pleistocene divergence between eastern and western populations of wood ducks (Aix sponsa) inferred by the 'isolation with migration' coalescent method." Molecular Ecology 14.11 (2005): 3407-3418. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1365-294X.2005.02618.x

[544] Brammah, Martin, Joseph I. Hoffman, and William Amos. "Genetic divergence between and within two subspecies of Laudakia stellio on islands in the Greek Cyclades." The Herpetological Journal 20.2 (2010): 91-98. Retrieved from https://pdfs.semanticscholar.org/4b09/5907c4eef7c941d1345a14d2516109d8f7f1.pdf

[545] Levy, Samuel, et al. "The diploid genome sequence of an individual human." PLoS Biol 5.10 (2007): e254. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pbio.0050254

[546] Te Nijenhuis, Jan, and Michael Van den Hoek. "Spearman's hypothesis tested on Black adults: A meta-analysis." Journal of Intelligence 4.2 (2016): 6. Retrieved from https://sci-hub.se/https://doi.org/10.3390/jintelligence4020006

[547] Wise, Cheryl A., et al. "Comparative nuclear and mitochondrial genome diversity in humans and chimpanzees." Molecular Biology and Evolution 14.7 (1997): 707-716. Retrieved from https://sci-hub.se/https://doi.org/10.1093/oxfordjournals.molbev.a025810

[548] Jorde, Lynn B., et al. "Microsatellite diversity and the demographic history of modern humans." Proceedings of the National Academy of Sciences 94.7 (1997): 3100-3103. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.94.7.3100

[549] Bowcock, Anne M., et al. "High resolution of human evolutionary trees with polymorphic microsatellites." Nature 368.6470 (1994): 455-457. Retrieved from https://sci-hub.se/https://doi.org/10.1038/368455a0

[550] Reinartz, G. E., et al. "Patterns of microsatellite polymorphism in the range-restricted bonobo (Pan paniscus): considerations for interspecific comparison with chimpanzees (P. troglodytes)." Molecular Ecology 9.3 (2000): 315-328. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.1365-294x.2000.00852.x

[551] O'Ryan, Colleen, et al. "Microsatellite analysis of genetic diversity in fragmented South African buffalo populations." Animal Conservation 1.2 (1998): 85-94. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1469-1795.1998.tb00015.x

[552] Uphyrkina, Olga, et al. "Phylogenetics, genome diversity and origin of modern leopard, Panthera pardus." Molecular ecology 10.11 (2001): 2617-2633. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.0962-1083.2001.01350.x

[553] Eizirik, Eduardo, et al. "Phylogeography, population history and conservation genetics of jaguars (Panthera onca, Mammalia, Felidae)." Molecular Ecology 10.1 (2001): 65-79. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.1365-294X.2001.01144.x

[554] Culver, Melanie, et al. "Genomic ancestry of the American puma (Puma concolor)." Journal of Heredity 91.3 (2000): 186-197. Retrieved from https://sci-hub.se/https://doi.org/10.1093/jhered/91.3.186

[555] Ebersberger, Ingo, et al. "Genomewide comparison of DNA sequences between humans and chimpanzees." The American Journal of Human Genetics 70.6 (2002): 1490-1497. Retrieved from https://sci-hub.se/https://doi.org/10.1086/340787

[556] Schwartz, Michael K., et al. "DNA reveals high dispersal synchronizing the population dynamics of Canada lynx." Nature 415.6871 (2002): 520-522. Retrieved from https://sci-hub.se/https://doi.org/10.1038/415520a

[557] Paetkau, David, et al. "Variation in genetic diversity across the range of North American brown bears." Conservation Biology 12.2 (1998): 418-429. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1523-1739.1998.96457.x

[558] Walker, C. W., et al. "Genetic variation and population structure in Scandinavian wolverine (Gulo gulo) populations." Molecular Ecology 10.1 (2001): 53-63. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.1365-294X.2001.01184.x

[559] García-Moreno, Jaime, et al. "Relationships and genetic purity of the endangered Mexican wolf based on analysis of microsatellite loci." Conservation Biology 10.2 (1996): 376-389. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.1523-1739.1996.10020376.x

[560] Girman, D. J., et al. "Patterns of population subdivision, gene flow and genetic variability in the African wild dog (Lycaon pictus)." Molecular Ecology 10.7 (2001): 1703-1723. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.0962-1083.2001.01302.x

[561] Kyle, C. J., and C. Strobeck. "Genetic structure of North American wolverine (Gulo gulo) populations." *Molecular Ecology* 10.2 (2001): 337-347. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.1365-294X.2001.01222.x

[562] Walker, C. W., et al. "Genetic variation and population structure in Scandinavian wolverine (Gulo gulo) populations." *Molecular Ecology* 10.1 (2001): 53-63. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.1365-294X.2001.01184.x

[563] Polziehn, Renee O., et al. "Microsatellite analysis of North American wapiti (Cervus elaphus) populations." *Molecular Ecology* 9.10 (2000): 1561-1576. https://sci-hub.se/https://doi.org/10.1046/j.1365-294x.2000.01033.x

[564] Forbes, Stephen H., et al. "Microsatellite evolution in congeneric mammals: domestic and bighorn sheep." *Molecular Biology and Evolution* 12.6 (1995): 1106-1113. Retrieved from https://www.researchgate.net/profile/Allan_Crawford/publication/51295709_Microsatellite_evolution_in_congeneric_mammals_Domestic_and_bighorn_sheep/links/566be35408ae1a797e3c8f6f/Microsatellite-evolution-in-congeneric-mammals-Domestic-and-bighorn-sheep.pdf

[565] Reinartz, G. E., et al. "Patterns of microsatellite polymorphism in the range-restricted bonobo (Pan paniscus): considerations for interspecific comparison with chimpanzees (P. troglodytes)." *Molecular Ecology* 9.3 (2000): 315-328. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.1365-294x.2000.00852.x

[566] Paetkau, David, et al. "Genetic structure of the world's polar bear populations." *Molecular Ecology* 8.10 (1999): 1571-1584. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.1365-294x.1999.00733.x

[567] Wilton, A. N., D. J. Steward, and K. Zafiris. "Microsatellite variation in the Australian dingo." *Journal of Heredity* 90.1 (1999): 108-111. Retrieved from https://sci-hub.se/https://doi.org/10.1093/jhered/90.1.108

[568] Barbujani, Guido, et al. "An apportionment of human DNA diversity." Proceedings of the National Academy of Sciences 94.9 (1997): 4516-4519. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.94.9.4516

[569] Elhaik, Eran. "Empirical distributions of F ST from large-scale human polymorphism data." PloS one 7.11 (2012): e49837. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pone.0049837

[570] Long, Jeffrey C., and Rick A. Kittles. "Human genetic diversity and the nonexistence of biological races." Human biology 81.5/6 (2009): 777-798. Retrieved from https://sci-hub.se/https://doi.org/10.3378/027.081.0621

[571] Jorde, Lynn B., et al. "The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data." The American Journal of Human Genetics 66.3 (2000): 979-988. Retrieved from https://sci-hub.se/https://doi.org/10.1086/302825

[572] Kim, K. S., et al. "Genetic variability in East Asian dogs using microsatellite loci analysis." Journal of Heredity 92.5 (2001): 398-403. Retrieved from https://sci-hub.se/https://doi.org/10.1093/jhered/92.5.398

[573] Van Hooft, W. F., A. F. Groen, and H. H. T. Prins. "Microsatellite analysis of genetic diversity in African buffalo (Syncerus caffer) populations throughout Africa." Molecular Ecology 9.12 (2000): 2017-2025. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.1365-294x.2000.01101.x

[574] Roy, Michael S., et al. "Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci." Molecular biology and Evolution 11.4 (1994): 553-570. Retrieved from https://sci-hub.se/https://doi.org/10.1093/oxfordjournals.molbev.a040137

[575] Wilson, Gregory M., et al. "Genetic variability of wolverines (Gulo gulo) from the Northwest Territories, Canada: conservation implications." Journal of Mammalogy 81.1 (2000): 186-196. Retrieved from https://sci-hub.se/https://doi.org/10.1644/1545-1542(2000)081%3C0186:GVOWGG%3E2.0.CO;2

[576] Jackson, Jennifer A., et al. "Global diversity and oceanic divergence of humpback whales (Megaptera novaeangliae)." Proceedings of the Royal Society B: Biological Sciences 281.1786 (2014): 20133222. Retrieved from https://sci-hub.se/https://doi.org/10.1098/rspb.2013.3222

[577] Lorenzen, Eline D., Peter Arctander, and Hans R. Siegismund. "High variation and very low differentiation in wide ranging plains zebra (Equus quagga): insights from mtDNA and microsatellites." Molecular ecology 17.12 (2008): 2812-2824. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1365-294X.2008.03781.x

[578] Lorenzen, Eline D., et al. "Phylogeography, hybridization and Pleistocene refugia of the kob antelope (Kobus kob)." Molecular ecology 16.15 (2007): 3241-3252. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1365-294X.2007.03382.x

[579] Jordana, Jordi, et al. "Genetic structure of eighteen local south European beef cattle breeds by comparative F-statistics analysis." Journal of Animal Breeding and Genetics 120.2 (2003): 73-87. Retrieved from https://sci-hub.se/https://doi.org/10.1046/j.1439-0388.2003.00384.x

[580] Williams, C. Lenney, et al. "Microsatellite variation in Red-winged Blackbirds (Agelaius phoeniceus)." Biochemical Genetics 42.1-2 (2004): 35-41. Retrieved from https://sci-hub.se/https://doi.org/10.1023/B:BIGI.0000012142.96374.b6

[581] Bamshad, Michael J., et al. "Human population genetic structure and inference of group membership." The American Journal of Human Genetics 72.3 (2003): 578-589. Retrieved from https://sci-hub.se/https://doi.org/10.1086/368061

[582] Guo, Guang, et al. "Genetic bio-ancestry and social construction of racial classification in social surveys in the contemporary United States." Demography 51.1 (2014): 141-172. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s13524-013-0242-0

[583] Allocco, Dominic J., et al. "Geography and genography: prediction of continental origin using randomly selected single nucleotide polymorphisms." BMC genomics 8.1 (2007): 68. Retrieved from https://sci-hub.se/https://doi.org/10.1186/1471-2164-8-68

[584] Lao, Oscar, et al. "Evaluating self-declared ancestry of US Americans with autosomal, Y-chromosomal and mitochondrial DNA." Human mutation 31.12 (2010): E1875-E1893. Retrieved from https://sci-hub.se/https://doi.org/10.1002/humu.21366

[585] Rosenberg, Noah A., et al. "Clines, clusters, and the effect of study design on the inference of human population structure." PLoS Genet 1.6 (2005): e70. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pgen.0010070

[586] Tang, Hua, et al. "Genetic structure, self-identified race/ethnicity, and confounding in case-control association studies." The American Journal of Human Genetics 76.2 (2005): 268-275. Retrieved from https://sci-hub.se/https://doi.org/10.1086/427888

[587] Witherspoon, David J., et al. "Genetic similarities within and between human populations." Genetics 176.1 (2007): 351-359. Retrieved from https://sci-hub.se/https://doi.org/10.1534/genetics.106.067355

[588] Porras-Hurtado, Liliana, et al. "An overview of STRUCTURE: applications, parameter settings, and supporting software." Frontiers in genetics 4 (2013): 98. Retrieved from https://sci-hub.se/https://doi.org/10.3389/fgene.2013.00098

[589] Rosenberg, Noah A., et al. "Genetic structure of human populations." science 298.5602 (2002): 2381-2385. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.1078311

[590] Bryc, Katarzyna, et al. "The genetic ancestry of african americans, latinos, and european Americans across the United States." The American Journal of Human Genetics 96.1 (2015): 37-53. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.ajhg.2014.11.010

[591] Salter, Frank. On genetic interests: Family, ethnicity, and humanity in an age of mass migration. Transaction publishers, 2006. Retrieved from https://ia800704.us.archive.org/22/items/OnGeneticInterestsFamilyEthnicityAndHumanityInAnAgeOfMassMigration2006ByFrankKempSalter/On%20Genetic%20Interests%20-%20Family%2C%20Ethnicity%2C%20and%20Humanity%20in%20an%20Age%20of%20Mass%20Migration%20%282006%29%20by%20Frank%20Kemp%20Salter.pdf

[592] Cofnas, Nathan. "Science is not always “self-correcting”." Foundations of Science 21.3 (2016): 477-492. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10699-015-9421-3

[593] Gottfredson, Linda S. "Egalitarian fiction and collective fraud." Society 31.3 (1994): 53-59. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF02693231

[594] Wicherts, Jelte M., et al. "Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking." Frontiers in psychology 7 (2016): 1832. Retrieved from https://sci-hub.se/https://doi.org/10.3389/fpsyg.2016.01832

[595] Hampshire, Adam, et al. "Fractionating human intelligence." Neuron 76.6 (2012): 1225-1237. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.neuron.2012.06.022

[596] Haier, Richard J., et al. "A comment on “Fractionating Intelligence” and the peer review process." Intelligence 46 (2014): 323-332. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2014.02.007

[597] Silberzahn, Raphael, et al. "Many analysts, one dataset: Making transparent how variations in analytical choices affect results.(2017)." *Retrieved from psyarxiv. com/qkwst* (2017). Retrieved from https://www.researchgate.net/publication/320041452_Many_analysts_one_dataset_Making_transparent_how_variations_in_analytical_choices_affect_results

[598] Botvinik-Nezer, Rotem, et al. "Variability in the analysis of a single neuroimaging dataset by many teams." *Nature* (2020): 1-7. Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41586-020-2314-9

[599] Munro, Geoffrey D., and Cynthia A. Munro. "“Soft” versus “hard” psychological science: Biased evaluations of scientific evidence that threatens or supports a strongly held political identity." *Basic and Applied Social Psychology* 36.6 (2014): 533-543. Retrieved from https://sci-hub.se/https://doi.org/10.1080/01973533.2014.960080

[600] Frisby, Craig L. "The Treatment of Race, Racial Differences, and Racism in Applied Psychology." Cultural Competence in Applied Psychology. Springer, Cham, 2018. 281-325. Retrieved from https://sci-hub.se/https://doi.org/10.1007/978-3-319-78997-2_13

[601] Wicherts, Jelte M., and Wendy Johnson. "Group differences in the heritability of items and test scores." Proceedings of the Royal Society B: Biological Sciences 276.1667 (2009): 2675-2683. Retrieved from https://sci-hub.se/https://doi.org/10.1098/rspb.2009.0238

[602] Wicherts, Jelte M. "Psychometric problems with the method of correlated vectors applied to item scores (including some nonsensical results)." Intelligence 60 (2017): 26-38. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2016.11.002

[603] Wicherts, Jelte M. "Ignoring psychometric problems in the study of group differences in cognitive test performance." Journal of biosocial science 50.6 (2018): 868-869. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0021932018000172

[604] Wicherts, Jelte M. "This (method) is (not) fine." Journal of Biosocial Science 50.6 (2018): 872-874. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0021932018000184

[605] Kirkegaard, Emil O. W. "Using classical test theory statistics with Jensen's method: A comment on Wicherts (2016)" Retrieved from https://rpubs.com/EmilOWK/230077

[606] Source 607 Supplementary Materials Download Link. Retrieved from https://www.tandfonline.com/doi/suppl/10.1080/1369183X.2015.1133279/suppl_file/cjms_a_1133279_sm2318.docx

[607] Zschirnt, Eva, and Didier Ruedin. "Ethnic discrimination in hiring decisions: a meta-analysis of correspondence tests 1990–2015." Journal of Ethnic and Migration Studies 42.7 (2016): 1115-1134. Retrieved from https://sci-hub.se/https://doi.org/10.1080/1369183X.2015.1133279

[608] Mitchell, Ojmarrh, and Doris L. MacKenzie. "The relationship between race, ethnicity, and sentencing outcomes: A meta-analysis of sentencing research." Final Report Submitted to the National Institute of Justice (2004). Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/208129.pdf

[609] de la Fuente, J., Davies, G., Grotzinger, A.D. et al. A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. Nat Hum Behav (2020). Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41562-020-00936-2

[610] Wu, Dong-Dong, and Ya-Ping Zhang. "Different level of population differentiation among human genes." BMC evolutionary biology 11.1 (2011): 1-7. Retrieved from https://sci-hub.se/https://doi.org/10.1186/1471-2148-11-16

[611] Wang, Eric T., et al. "Global landscape of recent inferred Darwinian selection for Homo sapiens." Proceedings of the National Academy of Sciences 103.1 (2006): 135-140. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.0509691102

[612] Lewontin, Richard C. "The apportionment of human diversity." Evolutionary biology. Springer, New York, NY, 1972. 381-398. Retrieved from https://sci-hub.se/https://doi.org/10.1007/978-1-4684-9063-3_14

[613] Harpending, Henry. "Kinship and population subdivision." Population and Environment 24.2 (2002): 141-147. Retrieved from https://sci-hub.se/https://doi.org/10.1023/A:1020815420693

[614] Oliva, Meritxell et al. "The Impact of Sex on Gene Expression Across Human Tissues." Science 369.6509 (2020): eaba3066. Retrieved from https://sci-hub.se/http://doi.org/10.1126/science.aba3066

[615] Farrer, Lindsay A., et al. "Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: a meta-analysis." Jama 278.16 (1997): 1349-1356. Retrieved from https://sci-hub.se/https://doi.org/10.1001/jama.1997.03550160069041

[616] Howell, Jennifer C., et al. "Race modifies the relationship between cognition and Alzheimer's disease cerebrospinal fluid biomarkers." Alzheimer's research & therapy 9.1 (2017): 88. Retrieved from https://sci-hub.se/https://doi.org/10.1186/s13195-017-0315-1

[617] Helgadottir, Anna, et al. "A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction." Nature genetics 38.1 (2006): 68-74. Retrieved from https://sci-hub.se/https://doi.org/10.1038/ng1692

[618] Fan, Chun Chieh, et al. "Modeling the 3D geometry of the cortical surface with genetic ancestry." Current Biology 25.15 (2015): 1988-1992. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.cub.2015.06.006

[619] Saiz, María, et al. "Genetic structure in the paternal lineages of South East Spain revealed by the analysis of 17 Y-STRs." Scientific reports 9.1 (2019): 1-9. Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41598-019-41580-9

[620] Dupanloup, Isabelle, et al. "Estimating the impact of prehistoric admixture on the genome of Europeans." Molecular Biology and Evolution 21.7 (2004): 1361-1372. Retrieved from https://sci-hub.se/https://doi.org/10.1093/molbev/msh135

[621] Bosch, Elena, et al. "High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula." The American Journal of Human Genetics 68.4 (2001): 1019-1029. Retrieved from https://sci-hub.se/https://doi.org/10.1086/319521

[622] Comas, David, et al. "Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: evidence for a strong genetic boundary through the Gibraltar Straits." Human genetics 107.4 (2000): 312-319. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s004390000370

[623] Bertoni, Bernardo, et al. "Admixture in Hispanics: distribution of ancestral population contributions in the Continental United States." Human Biology (2003): 1-11. Retrieved from https://sci-hub.se/https://doi.org/10.1353/hub.2003.0016

[624] Trzaskowski, Maciej, et al. "Genetic influence on family socioeconomic status and children's intelligence." Intelligence 42 (2014): 83-88. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2013.11.002

[625] Marceau, Kristine, et al. "The prenatal environment in twin studies: a review on chorionicity." Behavior genetics 46.3 (2016): 286-303. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-016-9782-6

[626] Supplementary materials for source 299 https://emilkirkegaard.dk/en/wp-content/uploads/chorion_data.csv

[627] Kirkegaard code (2016): https://gist.github.com/Deleetdk/eadb040c9c51f02e9203

[628] Lunde, Astrid, et al. "Genetic and environmental influences on birth weight, birth length, head circumference, and gestational age by use of population-based parent-offspring data." American journal of epidemiology 165.7 (2007): 734-741. Retrieved from https://sci-hub.se/https://doi.org/10.1093/aje/kwk107

[629] Dulal, Sophiya, et al. "Does antenatal micronutrient supplementation improve children’s cognitive function? Evidence from the follow-up of a double-blind randomised controlled trial in Nepal." BMJ global health 3.1 (2018). Retrieved from https://sci-hub.se/https://doi.org/10.1136/bmjgh-2017-000527

[630] Ritchie, Stuart J., and Elliot M. Tucker-Drob. "How much does education improve intelligence? A meta-analysis." Psychological science 29.8 (2018): 1358-1369. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0956797618774253

[631] Ritchie, Stuart J., et al. "Education is associated with higher later life IQ scores, but not with faster cognitive processing speed." Psychology and aging 28.2 (2013): 515. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0030820

[632] Der, Geoff, and Ian J. Deary. "The relationship between intelligence and reaction time varies with age: Results from three representative narrow-age age cohorts at 30, 50 and 69 years." Intelligence 64 (2017): 89-97. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2017.08.001

[633] Finn, Amy S., et al. "Cognitive skills, student achievement tests, and schools." Psychological Science 25.3 (2014): 736-744. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0956797613516008

[634] Watkins, Marley W., Pui-Wa Lei, and Gary L. Canivez. "Psychometric intelligence and achievement: A cross-lagged panel analysis." Intelligence 35.1 (2007): 59-68. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2006.04.005

[635] Watkins, Marley W., and Kara M. Styck. "A cross-lagged panel analysis of psychometric intelligence and achievement in reading and math." Journal of Intelligence 5.3 (2017): 31. Retrieved from https://sci-hub.se/https://doi.org/10.3390/jintelligence5030031

[636] Vernot, Benjamin, and Joshua M. Akey. "Resurrecting surviving Neandertal lineages from modern human genomes." Science 343.6174 (2014): 1017-1021. Retrieved from https://www.sci-hub.se/https://doi.org/10.1126/science.1245938

[637] Der, Geoff, G. David Batty, and Ian J. Deary. "The association between IQ in adolescence and a range of health outcomes at 40 in the 1979 US National Longitudinal Study of Youth." Intelligence 37.6 (2009): 573-580. Retrieved from https://sci-hub.se/10.1016/j.intell.2008.12.002

[638] Batty, G. David, et al. "IQ in early adulthood, socioeconomic position, and unintentional injury mortality by middle age: a cohort study of more than 1 million Swedish men." American journal of epidemiology 169.5 (2009): 606-615. Retrieved from https://sci-hub.se/https://doi.org/10.1093/aje/kwn381

[639] Whitley, Elise, et al. "Intelligence in early adulthood and subsequent risk of unintentional injury over two decades: cohort study of 1 109 475 Swedish men." Journal of Epidemiology & Community Health 64.5 (2010): 419-425. Retrieved from https://sci-hub.se/https://doi.org/10.1136/jech.2009.100669

[640] Deary, Ian J., et al. "Intelligence and persisting with medication for two years: Analysis in a randomised controlled trial." Intelligence 37.6 (2009): 607-612. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2009.01.001

[641] Gottfredson, Linda S., and Ian J. Deary. "Intelligence predicts health and longevity, but why?." Current Directions in Psychological Science 13.1 (2004): 1-4. Retrieved from https://sci-hub.se/https://doi.org/10.1111%2Fj.0963-7214.2004.01301001.x

[642] Gottfredson, Linda S. "Intelligence: is it the epidemiologists' elusive" fundamental cause" of social class inequalities in health?." Journal of personality and social psychology 86.1 (2004): 174. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-3514.86.1.174

[643] Heine, Steven J., et al. "What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect." Journal of personality and social psychology 82.6 (2002): 903. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-3514.82.6.903

[644] Heine, Steven J., Emma E. Buchtel, and Ara Norenzayan. "What do cross-national comparisons of personality traits tell us? The case of conscientiousness." Psychological Science 19.4 (2008): 309-313. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-9280.2008.02085.x

[645] Van de Gaer, Eva, et al. "The reference group effect: An explanation of the paradoxical relationship between academic achievement and self-confidence across countries." Journal of Cross-Cultural Psychology 43.8 (2012): 1205-1228. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0022022111428083

[646] Nye, Christopher D., et al. "Testing the measurement equivalence of personality adjective items across cultures." Journal of Research in Personality 42.6 (2008): 1524-1536. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jrp.2008.07.004

[647] Nuijten, Michèle B., et al. "Effect sizes, power, and biases in intelligence research: A meta-meta-analysis." (2018). Retrieved from https://sci-hub.se/https://doi.org/10.3390/jintelligence8040036

[648] Replication Index. “A Meta-Psychological Investigation of Intelligence Research with Z-Curve.2.0.” (2020) Retrieved from https://replicationindex.com/2020/10/03/a-meta-psychological-investigation-of-intelligence-research-with-z-curve-2-0/

[649] Source 647’s supplementary material: https://osf.io/z8emy/

[650] McDonald, Melissa M., et al. "Fertility and intergroup bias in racial and minimal-group contexts: Evidence for shared architecture." Psychological Science 22.7 (2011): 860-865. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0956797611410985

[651] Kanazawa, Satoshi. "Childhood intelligence and adult obesity." Obesity 21.3 (2013): 434-440. Retrieved from https://sci-hub.se/https://doi.org/10.1002/oby.20018

[652] Espenshade, Thomas J., Chang Y. Chung, and Joan L. Walling. "Admission preferences for minority students, athletes, and legacies at elite universities." Social Science Quarterly 85.5 (2004): 1422-1446. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.0038-4941.2004.00284.x

[653] Wilson, Edward O., and William L. Brown Jr. "The subspecies concept and its taxonomic application." Systematic zoology 2.3 (1953): 97-111. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2411818

[654] Gignac, Gilles E., and Timothy C. Bates. "Brain volume and intelligence: The moderating role of intelligence measurement quality." Intelligence 64 (2017): 18-29. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2017.06.004

[655] Plomin, Robert, et al. "Nature, nurture, and cognitive development from 1 to 16 years: A parent-offspring adoption study." Psychological Science 8.6 (1997): 442-447. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-9280.1997.tb00458.x

[656] Kan, Kees-Jan, et al. "On the nature and nurture of intelligence and specific cognitive abilities: The more heritable, the more culture dependent." Psychological Science 24.12 (2013): 2420-2428. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0956797613493292

[657] Kan, Kees-Jan. The nature of nurture: the role of gene-environment interplay in the development of intelligence. Universiteit van Amsterdam [Host], 2012. Retrieved from https://pure.uva.nl/ws/files/1689258/101363_thesis.pdf

[658] Jensen, Arthur R., and Frank CJ McGurk. "Black-white bias in ‘cultural and ‘noncultural’ test items." Personality and individual differences 8.3 (1987): 295-301. https://sci-hub.se/https://doi.org/10.1016/0191-8869(87)90029-8

[659] Jensen, Arthur R. "An examination of culture bias in the Wonderlic Personnel Test." Intelligence 1.1 (1977): 51-64. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(77)90026-5

[660] Reynolds, Cecil R., and Terry B. Gutkin. "A multivariate comparison of the intellectual performance of Black and White children matched on four demographic variables." Personality and Individual Differences 2.3 (1981): 175-180. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0191-8869(81)90021-0

[661] Jensen, Arthur R., and Cecil R. Reynolds. "Race, social class and ability patterns on the WISC-R." Personality and Individual Differences 3.4 (1982): 423-438. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0191-8869(82)90007-1

[662] Reynolds, Cecil R., and Arthur R. Jensen. "WISC-R subscale patterns of abilities of Blacks and Whites matched on Full Scale IQ." Journal of Educational Psychology 75.2 (1983): 207. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-0663.75.2.207

[663] Jensen (1973). Educability and Group Differences. Retrieved from https://b-ok.cc/book/2325266/aa3814

[664] Chuck. “A few New Analyses.” Human Varieties (2013). Retrieved from https://humanvarieties.org/2013/09/20/a-few-new-analyses/

[665] Pol, Hilleke E. Hulshoff, et al. "Genetic contributions to human brain morphology and intelligence." Journal of Neuroscience 26.40 (2006): 10235-10242. Retrieved from https://sci-hub.se/https://doi.org/10.1523/JNEUROSCI.1312-06.2006

[666] Hernstein, Richard, and Charles Murray. "The Bell Curve: Intelligence and Class Structure in American Life. 1994." (1994). Retrieved from https://emilkirkegaard.dk/en/wp-content/uploads/Richard_Herrnstein_Charles_Murray_The_Bell_Curve-OCR.pdf

[667] Chen, Ming-Huei, et al. "Trans-ethnic and ancestry-specific blood-cell genetics in 746,667 individuals from 5 global populations." Cell 182.5 (2020): 1198-1213. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.cell.2020.06.045

[668] Solloch, Ute V., et al. "Frequencies of gene variant CCR5-Δ32 in 87 countries based on next-generation sequencing of 1.3 million individuals sampled from 3 national DKMS donor centers." Human immunology 78.11-12 (2017): 710-717. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.humimm.2017.10.001

[669] Voigt, Emily A., et al. "Genetically defined race, but not sex, is associated with higher humoral and cellular immune responses to measles vaccination." Vaccine 34.41 (2016): 4913-4919. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.vaccine.2016.08.060

[670] Frisby, Craig L., and A. Alexander Beaujean. "Testing Spearman's hypotheses using a bi-factor model with WAIS-IV/WMS-IV standardization data." Intelligence 51 (2015): 79-97. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.04.007

[671] Gregory, Michael D., et al. "Neanderthal-derived genetic variation shapes modern human cranium and brain." Scientific Reports 7.1 (2017): 1-11. Retrieved from https://sci-hub.se/10.1038/s41598-017-06587-0

[672] Douet, Vanessa, et al. "Genetic influences on brain developmental trajectories on neuroimaging studies: from infancy to young adulthood." Brain imaging and behavior 8.2 (2014): 234-250. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11682-013-9260-1

[673] Huguet, Guillaume, et al. "Genome wide analysis of gene dosage in 24,092 individuals shows that 10,000 genes modulate cognitive ability." bioRxiv (2020). Retrieved from https://www.biorxiv.org/content/10.1101/2020.04.03.024554v3.full.pdf+html

[674] Hawks, John, et al. "Recent acceleration of human adaptive evolution." Proceedings of the National Academy of Sciences 104.52 (2007): 20753-20758. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.0707650104

[675] Candler, William H., and Hazel Ivey. "Cold weather injuries among US soldiers in Alaska: a five-year review." Military medicine 162.12 (1997): 788-791. Retrieved from https://sci-hub.se/https://doi.org/10.1093/milmed/162.12.788

[676] Simonti, Corinne N., et al. "The phenotypic legacy of admixture between modern humans and Neandertals." Science 351.6274 (2016): 737-741. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.aad2149

[677] SM, GARN. "Types and distribution of the hair in man." Annals of the New York Academy of Sciences 53.3 (1951): 498-507. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1749-6632.1951.tb31952.x

[678] Pollitzer, William S., and John J. Anderson. "Ethnic and genetic differences in bone mass: a review with a hereditary vs environmental perspective." The American journal of clinical nutrition 50.6 (1989): 1244-1259. Retrieved from https://sci-hub.se/https://doi.org/10.1093/ajcn/50.6.1244

[679] Araujo, Andre B., et al. "Lean mass, muscle strength, and physical function in a diverse population of men: a population-based cross-sectional study." BMC public health 10.1 (2010): 508. Retrieved from https://sci-hub.se/https://doi.org/10.1186/1471-2458-10-508

[680] Ebert, Antje D., et al. "Characterization of the molecular mechanisms underlying increased ischemic damage in the aldehyde dehydrogenase 2 genetic polymorphism using a human induced pluripotent stem cell model system." Science translational medicine 6.255 (2014): 255ra130-255ra130. Retrieved from https://sci-hub.se/https://doi.org/10.1126/scitranslmed.3009027

[681] Lara-Riegos, J. C., et al. "Diabetes susceptibility in Mayas: evidence for the involvement of polymorphisms in HHEX, HNF4α, KCNJ11, PPARγ, CDKN2A/2B, SLC30A8, CDC123/CAMK1D, TCF7L2, ABCA1 and SLC16A11 genes." Gene 565.1 (2015): 68-75. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.gene.2015.03.065

[682] Piel, Frédéric B., et al. "Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis." Nature communications 1.1 (2010): 1-7. Retrieved from https://sci-hub.se/https://doi.org/10.1038/ncomms1104

[683] Pennington, Bruce F., et al. "A twin MRI study of size variations in the human brain." Journal of cognitive neuroscience 12.1 (2000): 223-232. Retrieved from https://sci-hub.se/https://doi.org/10.1162/089892900561850

[684] Mitton, Jeffry B. "Genetic differentiation of races of man as judged by single-locus and multilocus analyses." The American Naturalist 111.978 (1977): 203-212. Retrieved from https://sci-hub.se/https://doi.org/10.1086/283155

[685] Paschou, Peristera, et al. "Ancestry informative markers for fine-scale individual assignment to worldwide populations." Journal of Medical Genetics 47.12 (2010): 835-847. Retrieved from https://sci-hub.se/10.1136/jmg.2010.078212

[686] Jin, Yumi, et al. "GRAF-pop: a fast distance-based method to infer subject ancestry from multiple genotype datasets without principal components analysis." G3: Genes, Genomes, Genetics 9.8 (2019): 2447-2461. Retrieved from https://sci-hub.se/https://doi.org/10.1534/g3.118.200925

[687] Nagar, Shashwat Deepali, Andrew B. Conley, and I. King Jordan. "Population structure and pharmacogenomic risk stratification in the United States." bioRxiv (2020). Retrieved from https://sci-hub.se/https://doi.org/10.1186/s12915-020-00875-4

[688] Fang, Huaying, et al. "Harmonizing genetic ancestry and self-identified race/ethnicity in genome-wide association studies." The American Journal of Human Genetics 105.4 (2019): 763-772. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.ajhg.2019.08.012

[689] Yuan, Jiao, et al. "Integrated analysis of genetic ancestry and genomic alterations across cancers." Cancer cell 34.4 (2018): 549-560. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.ccell.2018.08.019

[690] Burchard, Esteban González, et al. "The importance of race and ethnic background in biomedical research and clinical practice." New England Journal of Medicine 348.12 (2003): 1170-1175. Retrieved from https://sci-hub.se/https://doi.org/10.1056/NEJMsb025007

[691] Holick, Michael F., Ellen Smith, and Stephanie Pincus. "Skin as the site of vitamin D synthesis and target tissue for 1, 25-dihydroxyvitamin D3: use of calcitriol (1, 25-dihydroxyvitamin D3) for treatment of psoriasis." Archives of dermatology 123.12 (1987): 1677-1683a. Retrieved from https://sci-hub.se/https://doi.org/10.1001/archderm.1987.01660360108022

[692] Centers for Disease Control and Prevention. "Second national report on biochemical indicators of diet and nutrition in the US population." Atlanta (GA): CDC (2012). Retrieved from https://www.cdc.gov/nutritionreport/pdf/Nutrition_Book_complete508_final.pdf

[693] Bediou, Benoit, et al. "Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills." Psychological bulletin 144.1 (2018): 77. Retrieved from https://greenlab.psych.wisc.edu/wp-content/uploads/sites/280/2017/11/Bediou_et_al_PsychBull_2017.pdf

[694] Noble, Claire, et al. "The impact of shared book reading on children's language skills: A meta-analysis." Educational Research Review 28 (2019): 100290. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.edurev.2019.100290

[695] Hyytinen, Ari, et al. "Heritability of lifetime income." Helsinki Center of Economic Research Discussion Paper 364 (2013). Retrieved from https://sci-hub.se/https://doi.org/10.2139/ssrn.2253264

[696] Duncan, Greg J., Pamela A. Morris, and Chris Rodrigues. "Does money really matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments." Developmental psychology 47.5 (2011): 1263. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0023875

[697] Te Nijenhuis, Jan, et al. "Do schooling gains yield anomalous Jensen effects? A reply to Flynn (2019) including a meta-analysis." Journal of biosocial science 51.6 (2019): 917-919. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S002193201900021X

[698] Maynard, Rebecca A., and Richard J. Murnane. "The effects of a negative income tax on school performance: Results of an experiment." Journal of Human Resources (1979): 463-476. Retrieved from https://sci-hub.se/https://doi.org/10.2307/145317

[699] Blau, David M. "The effect of income on child development." Review of Economics and Statistics 81.2 (1999): 261-276. Retrieved from https://sci-hub.se/https://doi.org/10.1162/003465399558067

[700] Maynard, Rebecca A., and Richard J. Murnane. "The effects of a negative income tax on school performance: Results of an experiment." Journal of Human Resources (1979): 463-476. Retrieved from https://sci-hub.se/https://doi.org/10.2307/145317

[701] Chuck. “100 years of Testing Negro Intelligence.” Human Varieties (2013). Retrieved from https://humanvarieties.org/2013/01/15/100-years-of-testing-negro-intelligence/

[702] Faulk. Ryan. “Changes in the American Black-White IQ Gap: 1916-2016.” (2016). Retrieved from https://thealternativehypothesis.org/index.php/2016/12/22/changes-in-the-american-black-white-iq-gap-1916-2016/

[703] Nyborg, Helmuth, and Arthur R. Jensen. "Occupation and income related to psychometric g." Intelligence 29.1 (2001): 45-55. Retrieved from http://arthurjensen.net/wp-content/uploads/2014/06/Occupation-and-Income-Related-to-Psychometric-g-2001-by-Helmuth-S%C3%B8rensen-Nyborg-Arthur-Robert-Jensen.pdf

[704] Kanazawa, Satoshi. "The myth of racial discrimination in pay in the United States." Managerial and Decision Economics 26.5 (2005): 285-294. Retrieved from https://sci-hub.se/https://doi.org/10.1002/mde.1229

[705] Farkas, George, and Keven Vicknair. "Appropriate tests of racial wage discrimination require controls for cognitive skill: Comment on Cancio, Evans, and Maume." American Sociological Review 61.4 (1996): 557-560. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2096392

[706] Kirkegaard, Emil Ole William. "Cognitive meritocracy is the main cause of racial income gaps in the United States: an analysis of 5 large datasets." (2017). Retrieved from https://psyarxiv.com/qty3n/

[707] Frisby, Craig L., and A. Alexander Beaujean. "Testing Spearman's hypotheses using a bi-factor model with WAIS-IV/WMS-IV standardization data." Intelligence 51 (2015): 79-97. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.04.007

[708] Lubke, Gitta H., et al. "On the relationship between sources of within-and between-group differences and measurement invariance in the common factor model." Intelligence 31.6 (2003): 543-566. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0160-2896(03)00051-5

[709] MeasuringU "Z-Score to Percentile Calculator" Retrieved from https://measuringu.com/pcalcz/

[710] Caldwell, Kathleen L., et al. "Iodine status of the US population, National Health and Nutrition Examination Survey, 2005–2006 and 2007–2008." Thyroid 21.4 (2011): 419-427. Retrieved from https://sci-hub.se/https://doi.org/10.1089/thy.2010.0077

[711] Kang, Jae Hee, et al. "A randomized trial of vitamin E supplementation and cognitive function in women." Archives of internal medicine 166.22 (2006): 2462-2468. Retrieved from https://sci-hub.se/https://doi.org/10.1001/archinte.166.22.2462

[712] Anastasiou, Costas A., Mary Yannakoulia, and Nikolaos Scarmeas. "Vitamin D and cognition: an update of the current evidence." Journal of Alzheimer's Disease 42.s3 (2014): S71-S80. Retrieved from https://sci-hub.se/https://doi.org/10.3233/JAD-132636

[713] Lima, Aldo AM, et al. "Zinc, vitamin A, and glutamine supplementation in Brazilian shantytown children at risk for diarrhea results in sex-specific improvements in verbal learning." Clinics 68.3 (2013): 351-358. Retrieved from https://sci-hub.se/https://doi.org/10.6061/clinics/2013(03)OA11

[714] Protzko, John. "Raising IQ among school-aged children: Five meta-analyses and a review of randomized controlled trials." Developmental Review 46 (2017): 81-101. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.dr.2017.05.001

[715] Balk, Ethan M., et al. "Vitamin B6, B12, and folic acid supplementation and cognitive function: a systematic review of randomized trials." Archives of internal medicine 167.1 (2007): 21-30. Retrieved from https://sci-hub.se/https://doi.org/10.1001/archinte.167.1.21

[716] Harrison, Fiona E. "A critical review of vitamin C for the prevention of age-related cognitive decline and Alzheimer's disease." Journal of Alzheimer's Disease 29.4 (2012): 711-726. Retrieved from https://sci-hub.se/https://doi.org/10.3233/JAD-2012-111853

[717] Gogia, Siddhartha, and Harshpal S. Sachdev. "Zinc supplementation for mental and motor development in children." Cochrane Database of Systematic Reviews 12 (2012). Retrieved from https://sci-hub.se/https://doi.org/10.1002/14651858.CD007991.pub2

[718] Bhatnagar, Shinjini, and Sunita Taneja. "Zinc and cognitive development." British journal of nutrition 85.S2 (2001): S139-S145. Retrieved from https://sci-hub.se/https://doi.org/10.1079/BJN2000306

[719] Black, Maureen M. "The evidence linking zinc deficiency with children's cognitive and motor functioning." The Journal of nutrition 133.5 (2003): 1473S-1476S. Retrieved from https://sci-hub.se/https://doi.org/10.1093/jn/133.5.1473S

[720] Tupe, Rama P., and Shashi A. Chiplonkar. "Zinc supplementation improved cognitive performance and taste acuity in Indian adolescent girls." Journal of the American College of Nutrition 28.4 (2009): 388-396. Retrieved from https://sci-hub.se/https://doi.org/10.1080/07315724.2009.10718101

[721] Rico, Javier Alatorre, et al. "Efficacy of iron and/or zinc supplementation on cognitive performance of lead-exposed Mexican schoolchildren: a randomized, placebo-controlled trial." Pediatrics 117.3 (2006): e518-e527. Retrieved from https://sci-hub.se/https://doi.org/10.1542/peds.2005-1172

[722] Briefel, Ronette R., et al. "Zinc intake of the US population: findings from the third National Health and Nutrition Examination Survey, 1988–1994." The Journal of nutrition 130.5 (2000): 1367S-1373S. Retrieved from https://sci-hub.se/https://doi.org/10.1093/jn/130.5.1367S

[723] Sachdev, H. P. S., Tarun Gera, and Penelope Nestel. "Effect of iron supplementation on mental and motor development in children: systematic review of randomised controlled trials." Public health nutrition 8.2 (2005): 117-132. Retrieved from https://sci-hub.se/https://doi.org/10.1079/PHN2004677

[724] Lanphear, Bruce P., et al. "Low-level environmental lead exposure and children's intellectual function: an international pooled analysis." Environmental health perspectives 113.7 (2005): 894-899. Retrieved from https://sci-hub.se/https://doi.org/10.1289/ehp.7688

[725] of Menie, Michael A. Woodley, et al. "Are the effects of lead exposure linked to the g factor? A meta-analysis." Personality and Individual Differences 137 (2019): 184-191. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2018.09.005

[726] White, Brandi M., Heather Shaw Bonilha, and Charles Ellis. "Racial/ethnic differences in childhood blood lead levels among children< 72 months of age in the United States: a systematic review of the literature." Journal of racial and ethnic health disparities 3.1 (2016): 145-153. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s40615-015-0124-9

[727] Tsoi, Man-Fung, et al. "Continual decrease in blood lead level in Americans: United States National Health Nutrition and examination survey 1999-2014." The American journal of medicine 129.11 (2016): 1213-1218. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.amjmed.2016.05.042

[728] Ryan, Camille L., and Kurt Bauman. "Educational attainment in the United States: 2015 population characteristics." United States Census Bureau 2010 (2016): 20-578. Retrieved from https://www.census.gov/content/dam/Census/library/publications/2016/demo/p20-578.pdf

[729] Roth, Philip L., et al. "Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis." Personnel Psychology 54.2 (2001): 297-330. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1744-6570.2001.tb00094.x

[730] Plucker, Jonathan, et al. "Evaluation of the Cleveland Scholarship and Tutoring Program Technical Report 1998-2004." Director 812 (2006): 855-4438. Retrieved from http://schottfoundation.org/sites/default/files/resources/200602_Clev_Tech_Final.pdf

[731] Chrisman, Joe. "Test Score Data for Pupils in the Milwaukee Parental Choice Program." Legislative Audit Bureau (2012): Report 5. Retrieved from https://legis.wisconsin.gov/lab/reports/12-14full.pdf

[732] Wolf, Patrick, et al. "Evaluation of the DC Opportunity Scholarship Program." US Department of Education, Institute of Education Sciences (2010). Retrieved from https://ies.ed.gov/ncee/pubs/20104018/pdf/20104018.pdf

[733] Murray, Sheila, and Kim Rueben. "Racial disparities in education finance: Going beyond equal revenues." My science work (2008). Retrieved from https://www.taxpolicycenter.org/publications/racial-disparities-education-finance-going-beyond-equal-revenues/full

[734] Richwine, Jason. "The myth of racial disparities in public school funding." The Backgrounder 2548 (2011): 1. Retrieved from http://thf_media.s3.amazonaws.com/2011/pdf/bg2548.pdf

[735] Corcoran, Sean, et al. "The changing distribution of education finance, 1972–1997." Social inequality (2004): 433-465. Retrieved from https://www.russellsage.org/sites/all/files/u4/Corcoran%20et%20al.pdf

[736] Card, David, and Alan B. Krueger. "School quality and black-white relative earnings: A direct assessment." The quarterly journal of Economics 107.1 (1992): 151-200. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2118326

[737] Nagai, Althea. "Racial and Ethnic Admission Preferences at Arizona State University College of Law." Center for Equal Opportunity (2008): i-20. Retrieved from http://ceousa.org/attachments/article/541/ASU_LAW.pdf

[738] Lerner, Robert., Nagai, Althea. "Racial and Ethnic Preferences at the Three Virginia Public Law Schools." Center for Equal Opportunity (2002): i-51. Retrieved from http://ceousa.org/attachments/article/651/VALaw.pdf

[739] Armor, David. "AFFIRMATIVE ACTION AT THREE UNIVERSITIES." Virginia Association of Scholars Annual Meeting (2004): 1-10. Retrieved from http://ceousa.org/attachments/article/665/VAS%20Report.pdf

[740] Sander, Richard, and Aaron Danielson. "Thinking Hard About Race-Neutral Admissions." U. Mich. JL Reform 47 (2013): 967. Retrieved from https://repository.law.umich.edu/cgi/viewcontent.cgi?article=1116&context=mjlr

[741] Nagai, Althea. "Racial and Ethnic Preferences in Undergraduate Admissions at the University of Michigan." Center for Equal Opportunity (2006): i-31. Retrieved from http://ceousa.org/attachments/article/541/ASU_LAW.pdf

[742] Lerner, Robert., Nagai, Althea. "Racial and Ethnic Preferences in Admissions at Five Public Medical Schools." Center for Equal Opportunity (2001): i-88. Retrieved from http://ceousa.org/attachments/article/651/VALaw.pdf

[743] Nagai, Althea. "Racial and Ethnic Preferences in Undergraduate Admissions at Two Ohio Public Universities." Center for Equal Opportunity (2011): i-33. Retrieved from http://ceousa.org/attachments/article/541/ASU_LAW.pdf

[744] Lerner, Robert., Nagai, Althea. “Preferences at the Service Academies: Racial, Ethnic and Gender Preferences in Admissions to the U.S. Military Academy and the U.S. Naval Academy.” Center for Equal Opportunity (2006): 1-20. Retrieved from http://ceousa.org/attachments/article/663/ceousa-service-adademies.pdf

[745] Carnevale, Anthony P., et al. "SAT-Only Admission: How Would It Change College Campuses?." Georgetown University Center on Education and the Workforce (2019). Retrieved from https://files.eric.ed.gov/fulltext/ED600056.pdf

[746] Forbes. “How Much Help Do Millennials Get from Their Parents Paying for College?” (2017) Retrieved from https://www.forbes.com/sites/priceonomics/2017/05/18/how-much-help-do-millennials-get-from-their-parents-paying-for-college/?sh=4990b4ee7cf5

[747] 1000 Genomes Project Consortium. "A global reference for human genetic variation." Nature 526.7571 (2015): 68-74.
Paper: https://www.nature.com/articles/nature15393
Sci-hub: https://sci-hub.se/https://doi.org/10.1038/nature15393
Data: https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/

[748] Lee, James J., et al. "Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals." Nature genetics 50.8 (2018): 1112-1121. Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41588-018-0147-3

[749] Kantrowitz, Mark. "The distribution of grants and scholarships by race." FinAid. org, September 2 (2011). Retrieved from http://i.bnet.com/blogs/20110902racescholarships.pdf

[750] U.S. Department of Commerce, Census Bureau, Current Population Survey (CPS), October, 2018. See Digest of Education Statistics 2019, table 503.40. Retrieved from https://nces.ed.gov/programs/coe/pdf/coe_ssa.pdf

[751] Ang, SiewChing, Joseph Lee Rodgers, and Linda Wänström. "The Flynn Effect within subgroups in the US: Gender, race, income, education, and urbanization differences in the NLSY-Children data." Intelligence 38.4 (2010): 367-384. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2010.05.004

[752] Kirkegaard, Emil OW, et al. "Biogeographic ancestry, cognitive ability and socioeconomic outcomes." Psych 1.1 (2019): 1-25. Retrieved from https://sci-hub.se/https://doi.org/10.3390/psych1010001

[753] te Nijenhuis, Jan, Michael van den Hoek, and Elijah L. Armstrong. "Spearman's hypothesis and Amerindians: A meta-analysis." Intelligence 50 (2015): 87-92. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.02.006

[754] te Nijenhuis, Jan, Michael van den Hoek, and Joep Dragt. "A Meta-Analysis of Spearman’s Hypothesis Tested on Latin-American Hispanics, Including a New Way to Correct for Imperfectly Measuring the Construct of g." Psych 1.1 (2019): 101-122. Retrieved from https://sci-hub.se/https://doi.org/10.3390/psych1010008

[755] te Nijenhuis, Jan, et al. "Spearman's hypothesis tested on European Jews vs non-Jewish Whites and vs Oriental Jews: Two meta-analyses." Intelligence 44 (2014): 15-18. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2014.02.002

[756] Wicherts, Jelte M., Conor V. Dolan, and David J. Hessen. "Stereotype threat and group differences in test performance: a question of measurement invariance." Journal of personality and social psychology 89.5 (2005): 696. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-3514.89.5.696

[757] Flynn, James R. "Race, IQ and Jensen." (1980). Retrieved from http://arthurjensen.net/wp-content/uploads/2014/06/Race-IQ-and-Jensen-James-Flynn.pdf

[758] Race and self-esteem: meta-analyses comparing whites, blacks, Hispanics, Asians, and American Indians and comment on Gray-Little and Hafdahl (2000): 371-408. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0033-2909.128.3.371

[759] Breslau, Joshua, et al. "Specifying race-ethnic differences in risk for psychiatric disorder in a US national sample." Psychological medicine 36.1 (2006): 57. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0033291705006161

[760] Riolo, Stephanie A., et al. "Prevalence of depression by race/ethnicity: findings from the National Health and Nutrition Examination Survey III." American journal of public health 95.6 (2005): 998-1000. Retrieved from https://sci-hub.se/https://doi.org/10.2105/AJPH.2004.047225

[761] Stepler, Renee. “Hispanic, black parents see college degree as key for children’s success.” Pew Research Center (2016). Retrieved from https://www.pewresearch.org/fact-tank/2016/02/24/hispanic-black-parents-see-college-degree-as-key-for-childrens-success/

[762] U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (PFI-NHES), 2007. Retrieved from https://nces.ed.gov/pubs2012/2012026/tables/table_35.asp

[763] Mau, Wei-Cheng, and Lynette Heim Bikos. "Educational and vocational aspirations of minority and female students: A longitudinal study." Journal of Counseling & Development 78.2 (2000): 186-194. Retrieved from https://sci-hub.se/https://doi.org/10.1002/j.1556-6676.2000.tb02577.x

[764] Kuo, Yi-Lung, et al. "The intersectionality of race/ethnicity and socioeconomic status on social and emotional skills." Journal of Research in Personality 84 (2020): 103905. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jrp.2019.103905

[765] Moore, Elsie G. "Family socialization and the IQ test performance of traditionally and transracially adopted Black children." Developmental psychology 22.3 (1986): 317. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0012-1649.22.3.317

[766] Scarr, Sandra, and Richard A. Weinberg. "IQ test performance of black children adopted by white families." American Psychologist 31.10 (1976): 726. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0003-066X.31.10.726

[767] Weinberg, Richard A., Sandra Scarr, and Irwin D. Waldman. "The Minnesota Transracial Adoption Study: A follow-up of IQ test performance at adolescence." Intelligence 16.1 (1992): 117-135. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(92)90028-P

[768] Lynn, Richard. "Some reinterpretations of the Minnesota transracial adoption study." Intelligence 19.1 (1994): 21-27. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(94)90050-7

[769] Dickens, William T., and James R. Flynn. "Common ground and differences." PSYCHOLOGICAL SCIENCE-CAMBRIDGE- 17.10 (2006): 923. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-9280.2006.01804.x

[770] Levin, Michael. "Comment on the Minnesota transracial adoption study." Intelligence 19.1 (1994): 13-20. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(94)90049-3

[771] Scarr, Sandra. "About Us." DAILY FIX COFFEE (2001). Retrieved from https://web.archive.org/web/20160319081511/https://www.dailyfixcoffee.com/about_our_farm.htm

[772] Eyferth, Klaus. "Leistungen verschiedener gruppen von besatzungskindern in Hamburg-Wechsler Intelligenztest fur Kinder (HAWIK)." Archiv für die gesamte Psychologie 113 (1961): 222-241. Retrieved from https://emilkirkegaard.dk/en/wp-content/uploads/Eyferth-1961.pdf

[773] Davenport, Roy K. "Implications of military selection and classification in relation to universal military training." The Journal of Negro Education 15.4 (1946): 585-594. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2965882

[774] Thomas, Drew. "Racial IQ Differences among Transracial Adoptees: Fact or Artifact?." Journal of Intelligence 5.1 (2017): 1-18. Retrieved from https://sci-hub.se/https://doi.org/10.3390/jintelligence5010001

[775] Shuey, Audrey M. "The testing of Negro intelligence." (1966). Retrieved from https://www.gwern.net/docs/iq/1966-shuey-thetestingofnegrointelligencevol1.pdf

[776] Rowe, David C. "IQ, birth weight, and number of sexual partners in White, African American, and mixed race adolescents." Population and Environment 23.6 (2002): 513-524. Retrieved from https://sci-hub.se/https://doi.org/10.1023/A:1016313718644

[777] Lasker, Jordan, et al. "Global ancestry and cognitive ability." Psych 1.1 (2019): 431-459. Retrieved from https://sci-hub.se/https://doi.org/10.3390/psych1010034

[778] Scarr, Sandra, et al. "Absence of a relationship between degree of White ancestry and intellectual skills within a Black population." Human genetics 39.1 (1977): 69-86. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF00273154

[779] Lynn, Richard. "Skin color and intelligence in African Americans." Population and Environment 23.4 (2002): 365-375. Retrieved from https://sci-hub.se/https://doi.org/10.1023/A:1014572602343

[780] Fuerst, John GR, Richard Lynn, and Emil OW Kirkegaard. "The Effect of Biracial Status and Color on Crystallized Intelligence in the US-Born African–European American Population." Psych 1.1 (2019): 44-54. Retrieved from https://sci-hub.se/https://doi.org/10.3390/psych1010004

[781] Wilkerson, D. A., and Mary A. Morton. "Abstracts and Digests." Journal of Negro Education (1937): 210-217. Retrieved from https://lesacreduprintemps19.files.wordpress.com/2011/07/witty-and-jenks.pdf

[782] Hicks, Robert A., and Robert J. Pellegrini. "The meaningfulness of Negro-white differences in intelligence test performance." The Psychological Record 16.1 (1966): 43-46. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF03393641

[783] Wilkerson, D. A., and Mary A. Morton. "Abstracts and Digests." Journal of Negro Education (1937): 210-217. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2292268

[784] Herskovits, Melville Jean. The anthropometry of the American Negro. Vol. 11. Ardent Media, 1969. Retrieved from https://books.google.com/books?id=_9Sc1vVpJrAC&pg=PA212&lpg=PA212&dq=%22The+anthropometry+of+the+American+Negro%22+negro-white&source=bl&ots=YFvlVVkKGX&sig=b-HmFz3evk2olOde3ux29STya6I&hl=en&ei=PvgaToyQHujz0gGVhtGWBQ&sa=X&oi=book_result&ct=result&resnum=3&ved=0CCYQ6AEwAg#v=onepage&q&f=false

[785] Chuck. “Witty and Jenkins (1936).” Race Genes And Disparity (2008). Retrieved from https://abc102.wordpress.com/2008/07/13/witty-and-jenkins-1936/

[786] Coleman, Karen J., et al. "Racial-ethnic differences in psychiatric diagnoses and treatment across 11 health care systems in the mental health research network." Psychiatric Services 67.7 (2016): 749-757. Retrieved from https://sci-hub.se/https://doi.org/10.1176/appi.ps.201500217

[787] Pesta, Bryan J., et al. "Intelligence-associated Polygenic Scores Predict g, Independent of Ancestry, Parental Educational Levels, and Color among Hispanics in comparison to European, European-African, and African Americans." bioRxiv (2020). Retrieved from https://www.biorxiv.org/content/10.1101/2020.09.24.312074v2.full.pdf+html

[788] Fuerst, John, and Emil OW Kirkegaard. "Admixture in the Americas: Regional and national differences." Mankind Quarterly 56.3 (2016): 255. Retrieved from https://www.researchgate.net/publication/298214364_Admixture_in_the_Americas_Regional_and_National_Differences

[789] Kirkegaard, Emil OW, Mingrui Wang, and John Fuerst. "Biogeographic Ancestry and Socioeconomic Outcomes in the Americas: A Meta-Analysis." bioRxiv (2016): 055681. Retrieved from https://www.biorxiv.org/content/10.1101/055681v4.full.pdf+html

[790] Kirkegaard, Emil OW, and John Fuerst. "Admixture in Argentina." Mankind Quarterly 57.4 (2017): 542. Retrieved from https://www.researchgate.net/publication/317552097_Admixture_in_Argentina

[791] Templer, Donald I. "Correlational and factor analytic support for Rushton’s differential K life history theory." Personality and Individual Differences 45.6 (2008): 440-444. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2008.05.010

[792] Meisenberg, Gerhard, and Michael A. Woodley. "Global behavioral variation: A test of differential-K." Personality and individual differences 55.3 (2013): 273-278. Retrieved from https://philipperushton.net/wp-content/uploads/2015/02/iq-race-r-k-theory-rushton-meisenberg-personality-individual-differences-7-2013.pdf

[793] Templer, Donald I., and John S. Stephens. "The relationship between IQ and climatic variables in African and Eurasian countries." Intelligence 46 (2014): 169-178. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2014.06.001

[794] Scarr, Sandra, et al. "Absence of a relationship between degree of White ancestry and intellectual skills within a Black population." Human genetics 39.1 (1977): 69-86. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF00273154

[795] Fish, Jefferson M., ed. Race and intelligence: Separating science from myth. Routledge, 2001. Retrieved from https://b-ok.cc/book/1179262/e502be

[796] Nisbett, Richard E. Intelligence and how to get it: Why schools and cultures count. WW Norton & Company, 2009. Retrieved from https://b-ok.cc/book/1298274/2f6f98

[797] Kirkegaard, Emil OW. "Inequality across US counties: an S factor analysis." Open Quantitative Sociology & Political Science (2016). Retrieved from https://www.researchgate.net/publication/319870395_Inequality_across_US_counties_an_S_factor_analysis

[798] Kirkegaard, Emil O. W.. (2016). Inequality across prefectures in Japan: An S factor analysis. Open Quantitative Sociology & Political Science. Retrieved from https://www.researchgate.net/publication/299788991_Inequality_across_prefectures_in_Japan_An_S_factor_analysis

[799] Zakharia, Fouad, et al. "Characterizing the admixed African ancestry of African Americans." Genome biology 10.12 (2009): R141. Retrieved from https://sci-hub.se/https://doi.org/10.1186/gb-2009-10-12-r141

[800] Scarr, Sandra. "On Arthur Jensen's integrity." Intelligence 26.3 (1998): 227-232. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0160-2896(99)80005-1

[801] Kirkegaard, Emil OW. "IQ and socioeconomic variables in French departements: Reanalysis and new data." Mankind Quarterly 56.2 (2015): 113. Retrieved from https://www.researchgate.net/publication/290430496_IQ_and_Socioeconomic_Variables_in_French_Departements_Reanalysis_and_New_Data

[802] Kirkegaard, Emil OW. "The S factor in Brazilian states." The Winnower (2015). Retrieved from https://www.researchgate.net/publication/277591477_The_S_factor_in_Brazilian_states

[803] Kirkegaard, Emil OW. "The general socioeconomic factor among Colombian departments." The Winnower (2015). Retrieved from https://www.researchgate.net/publication/278392268_The_general_socioeconomic_factor_among_Colombian_departments

[804] Kirkegaard, Emil OW. "Inequality among 32 London boroughs: An S factor analysis." Open Quantitative Sociology & Political Science 1.1 (2016). Retrieved from https://www.researchgate.net/profile/Emil_O_W_Kirkegaard/publication/291947280_Inequality_among_32_London_Boroughs_An_S_factor_analysis/links/56a773b308ae860e02556028/Inequality-among-32-London-Boroughs-An-S-factor-analysis.pdf

[805] KIRKEGAARD, EMIL OW. "A replication of the S factor among US states using a new and larger dataset." (2015). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.902.4431&rep=rep1&type=pdf

[806] Kirkegaard, Emil OW. "The S factor in the British Isles: A reanalysis of Lynn (1979)." The Winnower (2015). Retrieved from https://www.researchgate.net/publication/275338677_The_S_factor_in_the_British_Isles_A_reanalysis_of_Lynn_1979

[807] Kirkegaard, Emil OW. "An S factor among census tracts of Boston." The Winnower (2015). Retrieved from https://www.researchgate.net/publication/277591289_An_S_factor_among_census_tracts_of_Boston

[808] Kirkegaard, Emil OW. "The S factor in China." The Winnower (2015). Retrieved from https://www.researchgate.net/publication/275338927_The_S_factor_in_China

[809] Kirkegaard, Emil OW. "Examining the S factor in Mexican states." The Winnower (2017). Retrieved from https://www.researchgate.net/profile/Emil_O_W_Kirkegaard/publication/275338746_Examining_the_S_factor_in_Mexican_states/links/58d366d5a6fdccd24d43c839/Examining-the-S-factor-in-Mexican-states.pdf

[810] Kirkegaard, E. O., and Bryan J. Pesta. "An S factor analysis on the provinces of Vietnam: Relationships with cognitive ability, ethnicity, and Latitude." Mankind Quarterly 58.4 (2018): 562-579. Retrieved from https://www.researchgate.net/profile/Emil_O_W_Kirkegaard/publication/325654239_An_S_Factor_Analysis_on_the_Provinces_of_Vietnam_Relationships_with_Cognitive_Ability_Ethnicity_and_Latitude/links/5b1adbeea6fdcca67b67152d/An-S-Factor-Analysis-on-the-Provinces-of-Vietnam-Relationships-with-Cognitive-Ability-Ethnicity-and-Latitude.pdf

[811] Kirkegaard, Emil OW. "S and G in Italian regions: Re-analysis of Lynn's data and new data." The Winnower (2015). Retrieved from https://www.researchgate.net/publication/275339328_S_and_G_in_Italian_regions_Re-analysis_of_Lynn's_data_and_new_data

[812] Lynn, Richard. Race differences in intelligence: an evolutionary analysis. Washington Summit Publishers, 2015. Retrieved from https://b-ok.cc/book/3355534/40ead2

[813] Lewis, Jason E., et al. "The mismeasure of science: Stephen Jay Gould versus Samuel George Morton on skulls and bias." PLoS Biol 9.6 (2011): e1001071. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pbio.1001071

[814] Beals, Kenneth L., et al. "Brain size, cranial morphology, climate, and time machines [and comments and reply]." Current Anthropology 25.3 (1984): 301-330. Retrieved from https://sci-hub.se/https://doi.org/10.1086/203138

[815] Harvey, I., et al. "Volumetric MRI measurements in bipolars compared with schizophrenics and healthy controls." Psychological Medicine 24.3 (1994): 689-699. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0033291700027847

[816] Jones, P. B., et al. "Cerebral ventricle dimensions as risk factors for schizophrenia and affective psychosis: an epidemiological approach to analysis." Psychological medicine 24.4 (1994): 995-1011. Retrieved from https://sci-hub.se/https://doi.org/10.1017/s0033291700029081

[817] Isamah, Nneka, et al. "Variability in frontotemporal brain structure: the importance of recruitment of African Americans in neuroscience research." PLoS One 5.10 (2010): e13642. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pone.0013642

[818] Altmann, Andre, and Janaina Mourao-Miranda. "Evidence for bias of genetic ancestry in resting state functional MRI." 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE, 2019. Retrieved from https://sci-hub.se/https://doi.org/10.1109/ISBI.2019.8759284

[819] Rushton, J. Philippe. "Cranial size and IQ in Asian Americans from birth to age seven." Intelligence 25.1 (1997): 7-20. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0160-2896(97)90004-0

[820] Rushton, J. Philippe. "Sex and race differences in cranial capacity from International Labour Office data." Intelligence 19.3 (1994): 281-294. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(94)90002-7

[821] Rushton, J. Philippe. "Cranial capacity related to sex, rank, and race in a stratified random sample of 6,325 US military personnel." Intelligence 16.3-4 (1992): 401-413. Retrieved from https://sci-hub.se/https://doi.org/10.1016/0160-2896(92)90017-L

[822] Rushton, J. Philippe. Race, evolution, and behavior: A life history perspective. Transaction Publ., 1996. Retrieved from https://b-ok.cc/book/2719026/32b5ba

[823] Tobias, Phillip V. "Brain-size, grey matter and race—fact or fiction?." American Journal of Physical Anthropology 32.1 (1970): 3-25. Retrieved from https://sci-hub.se/https://doi.org/10.1002/ajpa.1330320103

[824] Nisbett, Richard E., et al. "Intelligence: new findings and theoretical developments." American psychologist 67.2 (2012): 130. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0026699

[825] Ho, Khang-Cheng, et al. "Newborn brain weight in relation to maturity, sex, and race." Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society 10.3 (1981): 243-246. Retrieved from https://sci-hub.se/https://doi.org/10.1002/ana.410100308

[826] Mujugira, Andrew, et al. "Fetal head circumference, operative delivery, and fetal outcomes: a multi-ethnic population-based cohort study." BMC pregnancy and childbirth 13.1 (2013): 106. Retrieved from https://sci-hub.se/https://doi.org/10.1186/1471-2393-13-106

[827] Thomas, Pam, et al. "A new look at intrauterine growth and the impact of race, altitude, and gender." Pediatrics 106.2 (2000): e21-e21. Retrieved from https://sci-hub.se/https://doi.org/10.1542/peds.106.2.e21

[828] Madan, Ashima, et al. "Racial differences in birth weight of term infants in a northern California population." Journal of Perinatology 22.3 (2002): 230-235. Retrieved from https://sci-hub.se/https://doi.org/10.1038/sj.jp.7210703

[829] Drooger, Jan C., et al. "Ethnic differences in prenatal growth and the association with maternal and fetal characteristics." Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology 26.2 (2005): 115-122. Retrieved from https://sci-hub.se/https://doi.org/10.1002/uog.1962

[830] Parikh, Laura I., et al. "Fetal biometry: does patient ethnicity matter?." The Journal of Maternal-Fetal & Neonatal Medicine 27.5 (2014): 500-504. Retrieved from https://sci-hub.se/https://doi.org/10.3109/14767058.2013.820696

[831] Weigel, M. Margaret, and Maria Elena Caiza Sanchez. "Ethnic/racial disparities in the fetal growth outcomes of Ecuadorian newborns." Journal of immigrant and minority health 15.1 (2013): 198-206. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10903-011-9571-5

[832] Rushton, J. Philippe, and Elizabeth W. Rushton. "Brain size, IQ, and racial-group differences: Evidence from musculoskeletal traits." Intelligence 31.2 (2003): 139-155. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0160-2896(02)00137-X

[833] Rushton, J. Philippe. "Placing intelligence into an evolutionary framework or how g fits into the r–K matrix of life-history traits including longevity." Intelligence 32.4 (2004): 321-328. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2004.06.003

[834] Pearce, Eiluned, and Robin Dunbar. "Latitudinal variation in light levels drives human visual system size." Biology letters 8.1 (2012): 90-93. Retrieved from https://sci-hub.se/https://doi.org/10.1098/rsbl.2011.0570

[835] Templer, Donald I., and Hiroko Arikawa. "Temperature, skin color, per capita income, and IQ: An international perspective." Intelligence 34.2 (2006): 121-139. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2005.04.002

[836] Bailey, D.H., Geary, D.C. Hominid Brain Evolution. Hum Nat 20, 67–79 (2009). Retrieved from https://sci-hub.se/https://doi.org/10.1007/s12110-008-9054-0

[837] Pearl, Raymond. "The weight of the Negro brain." Science 80.2080 (1934): 431-434. Retrieved from https://sci-hub.se/https://doi.org/10.1126/science.80.2080.431

[838] Bean, Robert Bennett. "Some racial peculiarities of the Negro brain." American Journal of Anatomy 5.4 (1906): 353-432. Retrieved from https://sci-hub.se/https://doi.org/10.1002/aja.1000050402

[839] void

[840] Krueger, Patrick M., Jarron M. Saint Onge, and Virginia W. Chang. "Race/ethnic differences in adult mortality: the role of perceived stress and health behaviors." Social science & medicine 73.9 (2011): 1312-1322. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.socscimed.2011.08.007

[841] Allcott, Hunt, et al. "Food deserts and the causes of nutritional inequality." The Quarterly Journal of Economics 134.4 (2019): 1793-1844. Retrieved from https://sci-hub.se/https://doi.org/10.1093/qje/qjz015

[842] Vuoksimaa, Eero, et al. "Higher aggression is related to poorer academic performance in compulsory education." Journal of child psychology and psychiatry (2020). Retrieved from https://sci-hub.se/https://doi.org/10.1111/jcpp.13273

[843] Woodley, Michael A., Jan te Nijenhuis, and Raegan Murphy. "Is there a dysgenic secular trend towards slowing simple reaction time? Responding to a quartet of critical commentaries." Intelligence 46 (2014): 131-147. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2014.05.012

[844] Liu, Chang, Peter CM Molenaar, and Jenae M. Neiderhiser. "The impact of variation in twin relatedness on estimates of heritability and environmental influences." Behavior genetics 48.1 (2018): 44-54. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10519-017-9875-x

[845] Rindermann, Heiner. “Cognitive capitalism: Human capital and the wellbeing of nations.” Cambridge University Press, 2018. Retrieved from https://b-ok.cc/book/3601266/3cf5c9

[846] Rowe, David C., and Joseph L. Rodgers. "Expanding variance and the case of historical changes in IQ means: A critique of Dickens and Flynn (2001)." (2002): 759. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0033-295X.109.4.759

[847] Clark, Gregory. The son also rises: Surnames and the history of social mobility. Princeton University Press, 2015. Retrieved from https://b-ok.cc/book/2329536/778a87

[848] Clark, G., & Cummins, N. (2018). Nature versus nurture in social outcomes: A lineage study of 263,000 English individuals, 1750–2017. TMP https://sci-hub.se/https://doi.org/10.1007/978-3-319-16999-6_2162-2

[849] Retrieved from https://b-ok.cc/book/2220297/e6479a

[850] Metzen, Daniel. "The causes of group differences in intelligence studied using the method of correlated vectors and psychometric meta-analysis." Unpublished Master's Thesis, University of Amsterdam (2012). Retrieved from https://lesacreduprintemps19.files.wordpress.com/2013/05/metzen-2010.pdf

[851] Peper, Jiska S., et al. "Genetic influences on human brain structure: a review of brain imaging studies in twins." Human brain mapping 28.6 (2007): 464-473. Retrieved from https://sci-hub.se/https://doi.org/10.1002/hbm.20398

[852] Arendasy, Martin E., et al. "Do individual differences in test preparation compromise the measurement fairness of admission tests?." Intelligence 55 (2016): 44-56. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2016.01.004

[853] Malanchini, Margherita, et al. "Genetic factors underlie the association between anxiety, attitudes and performance in mathematics." Translational psychiatry 10.1 (2020): 1-11. Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41398-020-0711-3

[854] Reeve, Charlie L., and Silvia Bonaccio. "Does test anxiety induce measurement bias in cognitive ability tests?." Intelligence 36.6 (2008): 526-538. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2007.11.003

[855] Sommer, Markus, and Martin E. Arendasy. "Comparing different explanations of the effect of test anxiety on respondents' test scores." Intelligence 42 (2014): 115-127. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2013.11.003

[856] Sommer, Markus, and Martin E. Arendasy. "Further evidence for the deficit account of the test anxiety–test performance relationship from a high-stakes admission testing setting." Intelligence 53 (2015): 72-80. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2015.08.007

[857] Sommer, Markus, and Martin E. Arendasy. "Does trait test anxiety compromise the measurement fairness of high-stakes scholastic achievement tests?." Learning and Individual Differences 50 (2016): 1-10. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.lindif.2016.06.030

[858] Sommer, Markus, et al. "Do individual differences in test-takers' appraisal of admission testing compromise measurement fairness?." Intelligence 73 (2019): 16-29. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2019.01.006

[859] Steedle, Jeffrey T. "Keeping Your Cool: Does Test Anxiety Bias Performance on the ACT? Research Report 2018-3." ACT, Inc. (2018). Retrieved from https://files.eric.ed.gov/fulltext/ED593180.pdf

[860] Wicherts, Jelte M., and Annemarie Zand Scholten. "Test anxiety and the validity of cognitive tests: A confirmatory factor analysis perspective and some empirical findings." Intelligence 38.1 (2010): 169-178. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2009.09.008

[861] Blum, Diego, and Heinz Holling. "Spearman's law of diminishing returns. A meta-analysis." Intelligence 65 (2017): 60-66. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2017.07.004

[862] Genç, Erhan, et al. "Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence." Nature communications 9.1 (2018): 1-11. Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41467-018-04268-8

[863] Rae, Caroline, et al. "Is pH a biochemical marker of IQ?." Proceedings of the Royal Society of London. Series B: Biological Sciences 263.1373 (1996): 1061-1064. Retrieved from https://sci-hub.st/https://doi.org/10.1098/rspb.1996.0156

[864] Rae, Caroline, et al. "Brain bioenergetics and cognitive ability." Developmental neuroscience 25.5 (2003): 324-331. Retrieved from https://sci-hub.se/https://doi.org/10.1159/000073509

[865] Jung, Rex E., et al. "Biochemical markers of intelligence: a proton MR spectroscopy study of normal human brain." Proceedings of the Royal Society of London. Series B: Biological Sciences 266.1426 (1999): 1375-1379. Retrieved from https://sci-hub.se/https://doi.org/10.1098/rspb.1999.0790

[866] Flores AJ, Chavez TA, Bolger N, Casad BJ. “RETRACTED: Cardiovascular and Self-Regulatory Consequences of SES-Based Social Identity Threat.” Personality and Social Psychology Bulletin. 2019;45(5):700-714. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0146167218795157

[867] 1.Retraction Notice. Personality and Social Psychology Bulletin. November 2020. Retrieved from https://journals.sagepub.com/doi/pdf/10.1177/0146167220973962

[868] Duncan, Laramie E., and Matthew C. Keller. "A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry." American Journal of Psychiatry 168.10 (2011): 1041-1049. Retrieved from https://sci-hub.se/https://doi.org/10.1176/appi.ajp.2011.11020191

[869] Duncan, Laramie E., Alisha R. Pollastri, and Jordan W. Smoller. "Mind the gap: Why many geneticists and psychological scientists have discrepant views about gene–environment interaction (G× E) research." American Psychologist 69.3 (2014): 249. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0036320

[870] Harris, Judith Rich. No two alike: Human nature and human individuality. WW Norton & Company, 2010. Retrieved from https://b-ok.cc/book/3413368/b48414

[871] Shamosh, Noah A., and Jeremy R. Gray. "Delay discounting and intelligence: A meta-analysis." Intelligence 36.4 (2008): 289-305. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2007.09.004

[872] Toney, Jermaine. “Is There Racialized Legacy in Wealth Across Generations? Evidence from Panel Study, 1984–2013.” Working Paper. New York: The New School, 2016. Retrieved from https://ia601402.us.archive.org/3/items/source-872-toney-2016/Source872%20Toney2016.pdf https://www.aeaweb.org/conference/2017/preliminary/1758?page=4&per-page=50

[873]. McKernan, Signe-Mary., Ratcliffe, Caroline., Steuerle, Eugene., Quakenbush, Caleb., and Kalish, Emma. “Nine Charts about Wealth Inequality in America (Updated).” Urban Institute (2017). Retrieved from https://apps.urban.org/features/wealth-inequality-charts/?fbclid=IwAR1AHv7WoI5zTB2Y5dGmJS-BaI4ah71bDZsBzvHI7o8AAjQun3pq8ORwbwI https://archive.is/LksGH

[874] Shores, Kenneth, and Simon Ejdemyr. "Pulling back the curtain: Intra-district school spending inequality and its correlates." Available at SSRN 3009775 (2017). Retrieved from https://sejdemyr.github.io/docs/ejdemyr_shores_schoolineq.pdf

[875] Miller, Raegen, and Diana Epstein. "There Still Be Dragons: Racial Disparity in School Funding Is No Myth." Center for American Progress (2011). Retrieved from https://cdn.americanprogress.org/wp-content/uploads/issues/2011/07/pdf/still_be_dragons.pdf

[876] Roth, Philip L., et al. "Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis." Personnel Psychology 54.2 (2001): 297-330. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1744-6570.2001.tb00094.x

[877] Glass, Gene V., and Mary Lee Smith. "Meta-analysis of research on class size and achievement." Educational evaluation and policy analysis 1.1 (1979): 2-16. Retrieved from https://sci-hub.se/https://doi.org/10.2307/1164099

[878] Hattie, John. "The paradox of reducing class size and improving learning outcomes." International journal of educational research 43.6 (2005): 387-425. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.ijer.2006.07.002

[879] Slavin, Robert E. "Class size and student achievement: Small effects of small classes." Educational Psychologist 24.1 (1989): 99-110. Retrieved from https://sci-hub.se/https://doi.org/10.1207/s15326985ep2401_4

[880] Krueger, Alan B. "Experimental estimates of education production functions." The quarterly journal of economics 114.2 (1999): 497-532. Retrieved from https://sci-hub.se/https://doi.org/10.1162/003355399556052

[881] Nye, Barbara, Larry V. Hedges, and Spyros Konstantopoulos. "The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment." Educational Evaluation and Policy Analysis 21.2 (1999): 127-142. Retrieved from https://sci-hub.se/https://doi.org/10.3102/01623737021002127

[882] Molnar, Alex, et al. "Evaluating the SAGE program: A pilot program in targeted pupil-teacher reduction in Wisconsin." Educational Evaluation and Policy Analysis 21.2 (1999): 165-177. Retrieved from https://sci-hub.se/https://doi.org/10.2307/1164298

[883] Ehrenberg, Ronald G., et al. "Class size and student achievement." Psychological science in the public interest 2.1 (2001): 1-30. Retrieved from https://sci-hub.se/https://doi.org/10.1111/1529-1006.003

[884] U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, High School Transcript Study (HSTS), various years, 1990-2009. Retrieved from https://www.nationsreportcard.gov/hsts_2009/race_gpa.asp?tab_id=tab2&subtab_id=Tab_1

[885] Lee, Jaekyung. "Racial and ethnic achievement gap trends: Reversing the progress toward equity?." Educational researcher 31.1 (2002): 3-12. Retrieved from https://sci-hub.se/https://doi.org/10.3102/0013189X031001003

[886] U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (PFI-NHES), 2007. Retrieved from https://nces.ed.gov/pubs2012/2012026/tables/table_35.asp

[887] Resmovits, Joy. "Black Preschool Kids Get Suspended Much More Frequently Than White Preschool Kids, U.S. Survey Says." Los Angeles Times (2016). Retrieved from https://www.latimes.com/local/education/la-na-suspension-rates-preschool-crdc-20160606-snap-story.html archive https://archive.is/00wBS#selection-1997.8-1997.12

[888] Charles, Kerwin Kofi, Erik Hurst, and Nikolai Roussanov. "Conspicuous consumption and race." The Quarterly Journal of Economics 124.2 (2009): 425-467. Retrieved from https://scihubtw.tw/https://doi.org/10.1162/qjec.2009.124.2.425

[889] Skiba, Russell J., et al. "The color of discipline: Sources of racial and gender disproportionality in school punishment." The urban review 34.4 (2002): 317-342. Retrieved from https://sci-hub.se/https://doi.org/10.1023/A:1021320817372

[890] Wright, John Paul, et al. "Prior problem behavior accounts for the racial gap in school suspensions." Journal of Criminal Justice 42.3 (2014): 257-266. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jcrimjus.2014.01.001

[891] Faris, Robert. "Race, social networks, and school bullying." (2007). Retrieved from https://ia601508.us.archive.org/19/items/source-891-faris-2007/Source891%20Faris2007.pdf

[892] Hanish, Laura D., and Nancy G. Guerra. "The roles of ethnicity and school context in predicting children's victimization by peers." American journal of community psychology 28.2 (2000): 201-223. Retrieved from https://sci-hub.se/https://doi.org/10.1023/A:1005187201519

[893] Dal Borgo, Mariela. "Ethnic and racial disparities in saving behavior." The Journal of Economic Inequality 17.2 (2019): 253-283. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10888-018-9400-3

[894] Pope, Devin G., and Justin R. Sydnor. "What's in a Picture? Evidence of Discrimination from Prosper. com." Journal of Human resources 46.1 (2011): 53-92. Retrieved from https://sci-hub.se/https://doi.org/10.1353/jhr.2011.0025

[895] Chen, Stefanos. "Lending Discrimination: Black Borrowers Face Higher Hurdles, Study Shows." Huffington Post (2012). Retrieved from https://www.huffpost.com/entry/lending-discrimination-black-borrowers-face-higher-hurdles-in-lending-study_n_1300509

[896] Reserve, US Federal. "Report to the congress on credit scoring and its effects on the availability and affordability of credit." Board of Governors of the Federal Reserve System (2007). Retrieved from https://www.federalreserve.gov/boarddocs/rptcongress/creditscore/creditscore.pdf

[897] Hunter, William C. "Discrimination in mortgage lending." Chicago Fed Letter 95 (1995): 1. Retrieved from https://ia601406.us.archive.org/32/items/cfljuly1995-95-pdf/cfljuly1995-95-pdf.pdf Warning: second link is an automatic download: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&uact=8&ved=0ahUKEwio46u__bLNAhWi3YMKHVrtC_8QFgguMAI&url=https%3A%2F%2Fwww.chicagofed.org%2F~%2Fmedia%2Fpublications%2Fchicago-fed-letter%2F1995%2Fcfljuly1995-95-pdf.pdf&usg=AFQjCNF9PuYziSOcSDuUMBc37eCm0l9nSg&sig2=KTsuGeRwovKM5Auw66StIA&bvm=bv.124817099,d.amc

[898] Ross, Stephen L., and Geoffrey MB Tootell. "Redlining, the Community Reinvestment Act, and private mortgage insurance." Journal of Urban Economics 55.2 (2004): 278-297. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0094-1190(02)00508-9

[899] Berkovec, James A., et al. "Race, redlining, and residential mortgage loan performance." The Journal of Real Estate Finance and Economics 9.3 (1994): 263-294. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF01099279

[900] Cheng, Ping, Zhenguo Lin, and Yingchun Liu. "Racial discrepancy in mortgage interest rates." The Journal of Real Estate Finance and Economics 51.1 (2015): 101-120. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11146-014-9473-0

[901] Bhutta, Neil, and Aurel Hizmo. "Do minorities pay more for mortgages?." (2020). Retrieved from https://sci-hub.se/https://doi.org/10.2139/ssrn.3352876

[902] Black, Harold A., M. Cary Collins, and Ken B. Cyree. "Do black-owned banks discriminate against black borrowers?." Journal of Financial Services Research 11.1-2 (1997): 189-204. Retrieved from https://sci-hub.se/https://doi.org/10.1023/A:1007943610378

[903] Collins, William J., and Robert A. Margo. "Race and Home Ownership from the End of the Civil War to the Present." American Economic Review 101.3 (2011): 355-59. Retrieved from https://sci-hub.se/https://doi.org/10.3386/w16665

[904] Ahlbrandt Jr, Roger S. "Exploratory research on the redlining phenomenon." Real Estate Economics 5.4 (1977): 473-481. Retrieved from https://sci-hub.se/https://doi.org/10.1111/1540-6229.00846

[905] Hutchinson, Peter M., James R. Ostas, and J. David Reed. "A survey and comparison of redlining influences in urban mortgage lending markets." Real Estate Economics 5.4 (1977): 463-472. Retrieved from https://sci-hub.se/https://doi.org/10.1111/1540-6229.00845

[906] Dingemans, Dennis. "Redlining and mortgage lending in Sacramento." Annals of the Association of American Geographers 69.2 (1979): 225-239. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-8306.1979.tb01253.x

[907] Avery, Robert B., and Thomas M. Buynak. "Mortgage redlining: Some new evidence." Economic Review 17.3 (1981): 15-31. Retrieved from https://fraser.stlouisfed.org/files/docs/publications/frbclevreview/pages/1980-1984/68477_1980-1984.pdf

[908] Tootell, Geoffrey MB. "Redlining in Boston: Do mortgage lenders discriminate against neighborhoods?." The Quarterly Journal of Economics 111.4 (1996): 1049-1079. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2946707

[909] Bohrnstedt, George, et al. "School Composition and the Black-White Achievement Gap. NCES 2015-018." National Center for Education Statistics (2015). Retrieved from https://nces.ed.gov/nationsreportcard/subject/studies/pdf/school_composition_and_the_bw_achievement_gap_2015.pdf

[910] Akoben, Taj. “UNEMPLOYED BLACK WOMAN PRETENDS TO BE WHITE, JOB OFFERS SUDDENLY SKYROCKET.” retrieved from https://selfuni.wordpress.com/2013/12/29/unemployed-black-woman-pretends-to-be-white-job-offers-suddenly-skyrocket/

[911] Buss, David. Evolutionary psychology: The new science of the mind. Psychology Press, 2015. Retrieved from https://b-ok.cc/book/4974700/28568c

[912] Agan, A. Y., and S. B. Starr. "Ban the box, criminal records, and statistical discrimination: A field experiment. University of Michigan Law and Economics Research Paper No. 16-012." Ann Arbor: University of Michigan (2016). Retrieved from https://sci-hub.se/https://doi.org/10.2139/ssrn.2795795

[913] Pager, Devah. "The mark of a criminal record." American journal of sociology 108.5 (2003): 937-975. Retrieved from https://scholar.harvard.edu/files/pager/files/pager_ajs.pdf

[914] Roth, Philip L., Allen I. Huffcutt, and Philip Bobko. "Ethnic group differences in measures of job performance: A new meta-analysis." Journal of Applied Psychology 88.4 (2003): 694. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0021-9010.88.4.694

[915] Grasby, Katrina L., et al. "Little evidence that socioeconomic status modifies heritability of literacy and numeracy in Australia." Child Development 90.2 (2019): 623-637. Retrieved from https://sci-hub.se/https://doi.org/10.1111/cdev.12920

[916] Ridgeway, Greg. Analysis of racial disparities in the New York Police Department's stop, question, and frisk practices. Rand Corporation, 2007. Retrieved from https://www.rand.org/pubs/technical_reports/TR534.html

[917] Coviello, Decio, and Nicola Persico. "An economic analysis of Black-White disparities in the New York Police Department's stop-and-frisk program." The Journal of Legal Studies 44.2 (2015): 315-360. Retrieved from https://sci-hub.se/https://doi.org/10.1086/684292

[918] Persico, Nicola, and Petra Todd. "Generalising the hit rates test for racial bias in law enforcement, with an application to vehicle searches in Wichita." The Economic Journal 116.515 (2006): F351-F367. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1468-0297.2006.01126.x

[919] Grogger, Jeffrey, and Greg Ridgeway. "Testing for racial profiling in traffic stops from behind a veil of darkness." Journal of the American Statistical Association 101.475 (2006): 878-887. Retrieved from https://sci-hub.se/https://doi.org/10.1198/016214506000000168

[920] Baumgartner, Frank R., et al. "Race or Place? The Persistence of Race Effects in Police Behavior following Traffic Stops." (2018). Retrieved from https://fbaum.unc.edu/TrafficStops/Baumgartner-etal-MeasuringDisparities-20May2018.pdf

[921] Smith, Michael R., and Matthew Petrocelli. "Racial profiling? A multivariate analysis of police traffic stop data." Police quarterly 4.1 (2001): 4-27. Retrieved from https://sci-hub.se/https://doi.org/10.1177/109861101004001001

[922] Evans, William N., and Emily G. Owens. "COPS and Crime." Journal of Public Economics 91.1-2 (2007): 181-201. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jpubeco.2006.05.014

[923] Braga, Anthony A., et al. "Hot spots policing and crime reduction: an update of an ongoing systematic review and meta-analysis." Journal of experimental criminology 15.3 (2019): 289-311. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11292-019-09372-3

[924] Mohler, George O., et al. "Randomized controlled field trials of predictive policing." Journal of the American statistical association 110.512 (2015): 1399-1411. Retrieved from https://sci-hub.se/https://doi.org/10.1080/01621459.2015.1077710

[925] Skeem, Jennifer L., and Christopher T. Lowenkamp. "Risk, race, and recidivism: Predictive bias and disparate impact." Criminology 54.4 (2016): 680-712. Retrieved from https://sci-hub.se/https://doi.org/10.1111/1745-9125.12123

[926] Falk, Örjan, et al. "The 1% of the population accountable for 63% of all violent crime convictions." Social psychiatry and psychiatric epidemiology 49.4 (2014): 559-571. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s00127-013-0783-y

[927] Wu, Jawjeong. "Racial/ethnic discrimination and prosecution: A meta-analysis." Criminal Justice and Behavior 43.4 (2016): 437-458. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0093854815628026

[928] United States Federal Bureau Of Investigations Department Of Justice. "CRIME in the United States: 2004 Uniform Crime Reports Appendix IV" Retrieved from https://www2.fbi.gov/ucr/cius_04/documents/CIUS_2004_Section7.pdf

[929] United States Department Of Commerce Census Bureau. “Population Estimates Vintage 2008: National Tables”. Retrieved from https://web.archive.org/web/20120707040749/http://www.census.gov/popest/data/historical/2000s/vintage_2008

[930] United States Department Of Justice Office of Justice Programs Bureau of Justice Statistics National Crime Victimization Survey. “Criminal Victimization in the United States, 2000 Statistical Tables.” Retrieved from https://www.bjs.gov/content/pub/pdf/cvus00.pdf

[931] United States Department Of Justice Office of Justice Programs Bureau of Justice Statistics National Crime Victimization Survey. “Criminal Victimization in the United States, 2001 Statistical Tables.” Retrieved from https://www.bjs.gov/content/pub/pdf/cvus01.pdf

[932] United States Department Of Justice Office of Justice Programs Bureau of Justice Statistics National Crime Victimization Survey. “Criminal Victimization in the United States, 2002 Statistical Tables.” Retrieved from https://www.bjs.gov/content/pub/pdf/cvus02.pdf

[933] United States Department Of Justice Office of Justice Programs Bureau of Justice Statistics National Crime Victimization Survey. “Criminal Victimization in the United States, 2003 Statistical Tables.” Retrieved from https://www.bjs.gov/content/pub/pdf/cvus03.pdf

[934] United States Department Of Justice Office of Justice Programs Bureau of Justice Statistics National Crime Victimization Survey. “Criminal Victimization in the United States, 2004 Statistical Tables.” Retrieved from https://www.bjs.gov/content/pub/pdf/cvus04.pdf

[935] United States Department Of Justice Office of Justice Programs Bureau of Justice Statistics National Crime Victimization Survey. “Criminal Victimization in the United States, 2005 Statistical Tables.” Retrieved from https://www.bjs.gov/content/pub/pdf/cvus05.pdf

[936] United States Department Of Justice Office of Justice Programs Bureau of Justice Statistics National Crime Victimization Survey. “Criminal Victimization in the United States, 2006 Statistical Tables.” Retrieved from https://www.bjs.gov/content/pub/pdf/cvus06.pdf

[937] United States Department Of Justice Office of Justice Programs Bureau of Justice Statistics National Crime Victimization Survey. “Criminal Victimization in the United States, 2001 Statistical Tables.” Retrieved from https://www.bjs.gov/content/pub/pdf/cvus01.pdf

[938] United States Department Of Justice Office of Justice Programs Bureau of Justice Statistics National Crime Victimization Survey. “Criminal Victimization in the United States, 2001 Statistical Tables.” Retrieved from https://www.bjs.gov/content/pub/pdf/cvus01.pdf

[939] Mello, Steven. "More COPS, less crime." Journal of public economics 172 (2019): 174-200. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jpubeco.2018.12.003

[940] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. “Section IV: Persons arrested 2000 Table-43”. FBI CRIME STATISTICS: 13 DOES 50! Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2000/00sec4.pdf

[941] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. “Section IV: Persons arrested 2001 Table-43”. FBI CRIME STATISTICS: 13 DOES 50! Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2001/01sec4.pdf

[942] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. “Section IV: Persons arrested 2002 Table-43”. FBI CRIME STATISTICS: 13 DOES 50! Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2002/02sec4.pdf

[943] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. “Section IV: Persons arrested 2003 Table-43”. FBI CRIME STATISTICS: 13 DOES 50! Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2003/03sec4.pdf

[944] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. “Section IV: Persons arrested 2004 Table-43”. FBI CRIME STATISTICS: 13 DOES 50! Retrieved from https://www2.fbi.gov/ucr/cius_04/documents/CIUS2004.pdf

[945] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. “Section IV: Persons arrested 2005: Table-43A”. FBI CRIME STATISTICS: 13 DOES 50! Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2005

[946] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. “Section IV: Persons arrested 2006: Table-43A”. FBI CRIME STATISTICS: 13 DOES 50! Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2006

[947] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. “Section IV: Persons arrested 2007: Table-43A”. FBI CRIME STATISTICS: 13 DOES 50! Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2007

[948] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. “Section IV: Persons arrested 2008: Table-43A”. FBI CRIME STATISTICS: 13 DOES 50! Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2008

[949] Chalfin, Aaron, and Justin McCrary. The effect of police on crime: New evidence from US cities, 1960-2010. No. w18815. National Bureau of Economic Research, 2013. Retrieved from https://sci-hub.se/https://doi.org/10.3386/w18815

[950] MacDonald, John M., Jonathan Klick, and Ben Grunwald. "The effect of private police on crime: evidence from a geographic regression discontinuity design." Journal of the Royal Statistical Society: Series A (Statistics in Society) 179.3 (2016): 831-846. Retrieved from https://sci-hub.se/https://doi.org/10.1111/rssa.12142

[951] Shikishima, Chizuru, et al. "Is g an entity? A Japanese twin study using syllogisms and intelligence tests." Intelligence 37.3 (2009): 256-267. Retrieved from https://sci-hub.ee/https://doi.org/10.1016/j.intell.2008.10.010

[952] Franssen, Dennys B. "What is the significance of test-score differences? Six psychometric meta-analyses on g loadings and IQ scores: The relation of inbreeding, visual impairment, schizophrenia, epilepsy, hearing impairment, and giftedness with general intelligence." (2010). Retrieved from https://docs.google.com/file/d/0B0VDoaXaIou8MGItSF82aVpIVWM/edit archive: https://ia601404.us.archive.org/21/items/source-952-franssen-2010/Source952%20Franssen2010.pdf

[953] Kirkegaard, Emil OW. "The international general socioeconomic factor: Factor analyzing international rankings." Open Differential Psychology 8 (2014). Retrieved from https://openpsych.net/files/papers/Kirkegaard_2014e.pdf

[954] Jansen, Philip R., et al. "Genome-wide meta-analysis of brain volume identifies genomic loci and genes shared with intelligence." Nature communications 11.1 (2020): 1-12. Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41467-020-19378-5

[955] Weisburd, Sarit. "Police presence, rapid response rates, and crime prevention." Review of Economics and Statistics (2016): 1-45. Retrieved from https://sci-hub.se/https://doi.org/10.1162/rest_a_00889

[956] Rushton, J. Philippe. "No narrowing in mean Black–White IQ differences—Predicted by heritable g." (2012): 500. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0029614

[957] Nagin, Daniel S. "Deterrence in the twenty-first century." Crime and justice 42.1 (2013): 199-263. Retrieved from https://sci-hub.se/https://doi.org/10.1086/670398

[958] Edelman, Benjamin, Michael Luca, and Dan Svirsky. "Racial discrimination in the sharing economy: Evidence from a field experiment." American Economic Journal: Applied Economics 9.2 (2017): 1-22. Retrieved from https://sci-hub.se/https://doi.org/10.1257/app.20160213 (a)
https://www.hbs.edu/faculty/Publication%20Files/16-069_5c3b2b36-d9f8-4b38-9639-2175aaf9ebc9.pdf (b)

[959] Peter, Benjamin M., Desislava Petkova, and John Novembre. "Genetic landscapes reveal how human genetic diversity aligns with geography." Molecular biology and evolution 37.4 (2020): 943-951. Retrieved from https://sci-hub.se/https://doi.org/10.1093/molbev/msz280

[960] Hur, Yoon-Mi, and Timothy Bates. "Genetic and Environmental Influences on Cognitive Abilities in Extreme Poverty." Twin Research and Human Genetics 22.5 (2019): 297-301. Retrieved from https://sci-hub.se/https://doi.org/10.1017/thg.2019.92

[961] Piza, Eric L., and Vijay F. Chillar. "The Effect of Police Layoffs on Crime: A Natural Experiment Involving New Jersey's Two Largest Cities." Justice Evaluation Journal (2020): 1-19. Retrieved from https://sci-hub.se/https://doi.org/10.1080/24751979.2020.1858697
https://www.gwern.net/docs/sociology/2020-piza.pdf

[962] Darolia, Rajeev, et al. "Race and gender effects on employer interest in job applicants: new evidence from a resume field experiment." Applied Economics Letters 23.12 (2016): 853-856. Retrieved from https://sci-hub.se/https://doi.org/10.1080/13504851.2015.1114571

[963] Steffensmeier, Darrell, John Kramer, and Jeffery Ulmer. "Age differences in sentencing." Justice Quarterly 12.3 (1995): 583-602. Retrieved from https://sci-hub.se/10.1080/07418829500096151

[964] U.S. Census Bureau. Current Population Survey, Annual Social and Economic Supplement, 2005 - 2011. Table: Age by Race. CPS 7-year average - Data Collected in 2005 to 2011. Persons - All. Percentages by Race. Sums in Whole Numbers. Retrieved from https://www.census.gov/cps/data/cpstablecreator.html

[965] U.S. Census Bureau. Historical Poverty Tables: People and Families - 1959 to 2019. Table 5. Percent of People By Ratio of Income to Poverty Level. Retrieved from https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-people.html warning: direct download: https://www2.census.gov/programs-surveys/cps/tables/time-series/historical-poverty-people/hstpov5.xlsx

[966] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. Historical Data. Disaster Center. Retrieved from http://www.disastercenter.com/crime/uscrime.htm

[967] Kposowa, Augustine J., Kevin D. Breault, and Beatrice M. Harrison. "Reassessing the structural covariates of violent and property crimes in the USA: A county level analysis." British Journal of Sociology (1995): 79-105. Retrieved from https://scihubtw.tw/https://doi.org/10.2307/591624

[968] Zaw, Khaing, Darrick Hamilton, and William Darity. "Race, wealth and incarceration: Results from the National Longitudinal Survey of Youth." Race and Social Problems 8.1 (2016): 103-115. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s12552-016-9164-y

[969] Land, Kenneth C., Patricia L. McCall, and Lawrence E. Cohen. "Structural covariates of homicide rates: Are there any invariances across time and social space?." American journal of sociology 95.4 (1990): 922-963. Retrieved from https://sci-hub.se/https://doi.org/10.1086/229381

[970] void

[971] void

[972] void

[973] Holmes, Malcolm D., et al. "Ethnicity, legal resources, and felony dispositions in two southwestern jurisdictions." Justice Quarterly 13.1 (1996): 11-30. Retrieved from https://sci-hub.se/https://doi.org/10.1080/07418829600092801

[974] Spohn, Cassia, and Jerry Cederblom. "Race and disparities in sentencing: A test of the liberation hypothesis." Justice Quarterly 8.3 (1991): 305-327. Retrieved from https://sci-hub.se/https://doi.org/10.1080/07418829100091071

[975] Unnever, James D. "Direct and organizational discrimination in the sentencing of drug offenders." Social Problems 30.2 (1982): 212-225. Retrieved from https://sci-hub.se/https://doi.org/10.2307/800519

[976] Spohn, Cassia, John Gruhl, and Susan Welch. "The effect of race on sentencing: A re-examination of an unsettled question." Law & Soc'y Rev. 16 (1981): 71. Retrieved from https://sci-hub.se/https://doi.org/10.2307/3053550

[977] Hagan, John, Ilene H. Nagel, and Celesta Albonetti. "The differential sentencing of white-collar offenders in ten federal district courts." American Sociological Review (1980): 802-820. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2094896

[978] Spohn, Cassia, and Miriam DeLone. "When does race matter? An analysis of the conditions under which race affects sentence severity." Sociology of Crime, Law, and Deviance 2.1 (2000): 3-37. Retrieved from https://sci-hub.se/https://doi.org/10.1108/s1521-6136(2000)0000002005

[979] Albonetti, Celesta A. "Race and the probability of pleading guilty." Journal of Quantitative Criminology 6.3 (1990): 315-334. Retrieved from https://sci-hub.se/https://doi.org/10.1007/BF01065413

[980] Albonetti, Celesta A. "Sentencing under the federal sentencing guidelines: Effects of defendant characteristics, guilty pleas, and departures on sentence outcomes for drug offenses, 1991-1992." Law and Society Review (1997): 789-822. Retrieved from https://sci-hub.se/https://doi.org/10.2307/3053987

[981] Bushway, Shawn D., and Anne Morrison Piehl. "Judging judicial discretion: Legal factors and racial discrimination in sentencing." Law and Society Review (2001): 733-764. Retrieved from https://sci-hub.se/https://doi.org/10.2307/3185415

[982] Dixon, Jo. "The organizational context of criminal sentencing." American journal of sociology 100.5 (1995): 1157-1198. Retrieved from https://sci-hub.se/https://doi.org/10.1086/230635

[983] Engen, Rodney L., and Sara Steen. "The power to punish: Discretion and sentencing reform in the war on drugs." American Journal of Sociology 105.5 (2000): 1357-1395. Retrieved from https://sci-hub.se/https://doi.org/10.1086/210433

[984] Spohn, Cassia. Offender Race and Case Outcomes: Do Crime Seriousness and Strength of Evidence Matter? Final Activities Report. 2000. Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/184774.pdf

[985] LaFree, Gary D. "Official reactions to Hispanic defendants in the Southwest." Journal of Research in Crime and Delinquency 22.3 (1985): 213-237. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0022427885022003003

[986] Petersilia, Joan. "Racial disparities in the criminal justice system: A summary." Crime & Delinquency 31.1 (1985): 15-34. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0011128785031001002

[987] Welch, Susan, John Gruhl, and Cassia Spohn. "Dismissal, conviction, and incarceration of Hispanic defendants: A comparison with Anglos and Blacks." Social Science Quarterly 65.2 (1984): 257. Retrieved from https://search.proquest.com/openview/66f61427e67ab4f109b1a57b6c69a74c/1?pq-origsite=gscholar&cbl=1816420

[988] Beaver, Kevin M., et al. "No evidence of racial discrimination in criminal justice processing: Results from the National Longitudinal Study of Adolescent Health." Personality and Individual Differences 55.1 (2013): 29-34. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2013.01.020

[989] Mitchell, Tara L., et al. "Racial bias in mock juror decision-making: A meta-analytic review of defendant treatment." Law and human behavior 29.6 (2005): 621-637. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10979-005-8122-9

[990] Devine, Dennis J., and David E. Caughlin. "Do they matter? A meta-analytic investigation of individual characteristics and guilt judgments." Psychology, Public Policy, and Law 20.2 (2014): 109. Retrieved from https://sci-hub.se/https://doi.org/10.1037/law0000006

[991] Libgober, Brian. "Getting a Lawyer While Black: A Field Experiment." Lewis & Clark L. Rev. 24 (2020): 53. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3389279

[992] Steffensmeier, Darrell, and Chester L. Britt. "Judges' race and judicial decision making: Do black judges sentence differently?." Social Science Quarterly 82.4 (2001): 749-764. Retrieved from https://sci-hub.se/https://doi.org/10.1111/0038-4941.00057

[993] Uhlman, Thomas M. "Black elite decision making: The case of trial judges." American Journal of Political Science (1978): 884-895. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2110596

[994] void

[995] Worden, Robert E., Sarah J. McLean, and Andrew P. Wheeler. "Testing for racial profiling with the veil-of-darkness method." Police Quarterly 15.1 (2012): 92-111. Retrieved from https://sci-hub.se/https://doi.org/10.1177/1098611111433027

[996] Horrace, William C., and Shawn M. Rohlin. "How dark is dark? Bright lights, big city, racial profiling." Review of Economics and Statistics 98.2 (2016): 226-232. Retrieved from https://sci-hub.se/https://doi.org/10.1162/REST_a_00543

[997] Kalinowski, Jesse, Matthew Ross, and Stephen L. Ross. Addressing Seasonality in Veil of Darkness Tests for Discrimination: An Instrumental Variables Approach. No. 2019-028. 2019. Retrieved from https://media.economics.uconn.edu/working/2019-07.pdf

[998] Pickrell, Timothy M., and Tony Jianqiang Ye. Seat Belt Use in 2008—Race and Ethnicity Results Among Occupants Traveling With Children. No. HS-811 107. 2009. Retrieved from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811107

[999] Tattersall, Ian. "Homo sapiens". *Encyclopedia Britannica*, 12 Nov. 2020. Retrieved from https://www.britannica.com/topic/Homo-sapiens

[1000] Coleman, James S. “Equality of Educational Opportunity”. US Department of Health, Education, and Welfare, Office of Education, 1966. Retrieved from https://files.eric.ed.gov/fulltext/ED012275.pdf

[1001] Carpenter, Christopher S., and Mark Stehr. "The effects of mandatory seatbelt laws on seatbelt use, motor vehicle fatalities, and crash-related injuries among youths." Journal of Health Economics 27.3 (2008): 642-662. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jhealeco.2007.09.010

[1002] Kim, Sungyop, et al. "Analysis of teenage seat belt use: From the 2007 Missouri high school seat belt survey." Journal of safety research 40.4 (2009): 311-316. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jsr.2009.07.001

[1003] Ellis, Herman M., et al. "Achieving a credible health and safety approach to increasing seat belt use among African Americans." Journal of health care for the poor and underserved 11.2 (2000): 144-150. Retrieved from https://sci-hub.se/https://doi.org/10.1353/hpu.2010.0689

[1004] Ramchand, Rajeev, Rosalie Liccardo Pacula, and Martin Y. Iguchi. "Racial differences in marijuana-users' risk of arrest in the United States." Drug and alcohol dependence 84.3 (2006): 264-272. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.drugalcdep.2006.02.010

[1005] Langan, Patrick A. "The racial disparity in us drug arrests." Bureau of Justice Statistics (BJS) and US Dept of Justice and Office of Justice Programs and United States of America (1995). Retrieved from https://www.bjs.gov/content/pub/pdf/rdusda.pdf

[1006] Lange, James E., Mark B. Johnson, and Robert B. Voas. "Testing the racial profiling hypothesis for seemingly disparate traffic stops on the New Jersey Turnpike." Justice Quarterly 22.2 (2005): 193-223. Retrieved from https://sci-hub.se/https://doi.org/10.1080/07418820500088952

[1007] Tillyer, Rob, and Robin S. Engel. "Racial differences in speeding patterns: Exploring the differential offending hypothesis." Journal of Criminal Justice 40.4 (2012): 285-295. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jcrimjus.2012.04.001

[1008] Cherkauskas, Jennifer Marie-Calnon. "The Role of driving behavior in the debate over racially biased policing." (2011). Retrieved from https://etda.libraries.psu.edu/files/final_submissions/2838

[1009] Smith, William R., et al. North Carolina Highway Traffic Study. No. NCJ-204021. 2004. Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/204021.pdf

[1010] Olsen, Emily O'Malley, Ruth A. Shults, and Danice K. Eaton. "Texting while driving and other risky motor vehicle behaviors among US high school students." Pediatrics 131.6 (2013): e1708-e1715. Retrieved from https://sci-hub.se/https://doi.org/10.1542/peds.2012-3462

[1011] Pierson, Emma, et al. "A large-scale analysis of racial disparities in police stops across the United States." Nature human behaviour (2020): 1-10. Retrieved from https://sci-hub.se/https://doi.org/10.1038/s41562-020-0858-1

[1012] Devi, Tanaya, and Roland G. Fryer Jr. Policing the Police: The Impact of" Pattern-or-Practice" Investigations on Crime. No. w27324. National Bureau of Economic Research, 2020. Retrieved from https://sci-hub.se/https://doi.org/10.3386/w27324

[1013] Fendrich, Michael, and Timothy P. Johnson. "Race/ethnicity differences in the validity of self-reported drug use: Results from a household survey." Journal of Urban Health 82.3 (2005): iii67-iii81. Retrieved from https://sci-hub.se/https://doi.org/10.1093/jurban/jti065

[1014] Page, W. F., et al. "Urinalysis screened vs verbally reported drug use: The identification of discrepant groups." International Journal of the Addictions 12.4 (1977): 439-450. Retrieved from https://sci-hub.se/https://doi.org/10.3109/10826087709027235

[1015] Falck, Russel, et al. "The validity of injection drug users self-reported use of opiates and cocaine." Journal of Drug Issues 22.4 (1992): 823-832. Retrieved from https://sci-hub.se/https://doi.org/10.1177/002204269202200402

[1016] Feucht, Thomas E., Richard C. Stephens, and Michael L. Walker. "Drug use among juvenile arrestees: A comparison of self-report, urinalysis and hair assay." Journal of Drug Issues 24.1 (1994): 99-116. Retrieved from https://sci-hub.se/https://doi.org/10.1177/002204269402400106

[1017] Johnson, Timothy P., and Phillip J. Bowman. "Cross-cultural sources of measurement error in substance use surveys." Substance Use & Misuse 38.10 (2003): 1447-1490. Retrieved from https://sci-hub.se/https://doi.org/10.1081/JA-120023394

[1018] Ledgerwood, David M., et al. "Comparison between self-report and hair analysis of illicit drug use in a community sample of middle-aged men." Addictive behaviors 33.9 (2008): 1131-1139. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.addbeh.2008.04.009

[1019] Substance Abuse and Mental Health Services Administration, Drug Abuse Warning Network, 2011: National Estimates of Drug-Related Emergency Department Visits. HHS Publication No. (SMA) 13-4760, DAWN Series D-39. Rockville, MD: Substance Abuse and Mental Health Services Administration, 2013. Retrieved from https://www.samhsa.gov/data/sites/default/files/DAWN2k11ED/DAWN2k11ED/DAWN2k11ED.pdf

[1020] O'Malley, Patrick M., Jerald G. Bachman, and Lloyd D. Johnston. "Period, age, and cohort effects on substance use among American youth, 1976-82." American Journal of Public Health 74.7 (1984): 682-688. Retrieved from https://sci-hub.se/https://doi.org/10.2105/AJPH.74.7.682

[1021] Rubenstein, Edwin S. The color of crime: Race, crime, and justice in America. Second Edition. New Century Foundation, 2005. Retrieved from http://2kpcwh2r7phz1nq4jj237m22.wpengine.netdna-cdn.com/wp-content/uploads/2011/12/2005-Color-of-Crime-Report.pdf

[1022] Rubenstein, Edwin S. The color of crime: Race, crime, and justice in America. 2016 Revised Edition. New Century Foundation, 2005. Retrieved from https://2kpcwh2r7phz1nq4jj237m22-wpengine.netdna-ssl.com/wp-content/uploads/2016/03/Color-Of-Crime-2016.pdf

[1023] Brown, Robert A., and James Frank. "Race and officer decision making: Examining differences in arrest outcomes between black and white officers." Justice quarterly 23.1 (2006): 96-126. Retrieved from https://sci-hub.se/https://doi.org/10.1080/07418820600552527

[1024] Hixson, Lindsay., Hepler, Bradford B., and Myoung Ouk, Kim. "The White population: 2010. 2010 Census Briefs." Washington, DC: US Department of Commerce, Economics and Statistics Administration. (2011). Retrieved from https://www.census.gov/prod/cen2010/briefs/c2010br-05.pdf

[1025] Rastogi, Sonya., Johnson, Tallese D., Hoeffel, Elizabeth M., and Drewery, Malcolm P. "The Black population: 2010. 2010 Census Briefs." Washington, DC: US Department of Commerce, Economics and Statistics Administration. (2011). Retrieved from https://www.census.gov/prod/cen2010/briefs/c2010br-06.pdf

[1026] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. "Section IV: Persons arrested 2008: Table-43A". FBI CRIME STATISTICS: 13 DOES 50! Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/tables/table-43

[1027] United States Federal Bureau Of Investigations Department Of Justice. Uniform Crime Report. "Law Enforcement Officers Feloniously Killed". Table 44 - Race and Sex of Known Offender, 2001–2010. FBI CRIME STATISTICS 13 DOES 50! Retrieved from https://ucr.fbi.gov/leoka/leoka-2010/tables/table44-leok-feloniously-race-sex-known-offender-01-10.xls

[1028] Burch, Andrea M., et al. "Arrest-Related Deaths, 2003-2009-Statistical Tables." United States Department Of Justice, Office Of Justice Programs, Bureau Of Justice Statistics (2011). Retrieved from https://www.bjs.gov/content/pub/pdf/ard0309st.pdf

[1029] Cesario, Joseph, David J. Johnson, and William Terrill. "Is there evidence of racial disparity in police use of deadly force? Analyses of officer-involved fatal shootings in 2015–2016." Social psychological and personality science 10.5 (2019): 586-595. Retrieved from https://sci-hub.se/https://doi.org/10.1177/1948550618775108

[1030] Mentch, Lucas. "On Racial Disparities in Recent Fatal Police Shootings." Statistics and Public Policy 7.1 (2020): 9-18. Retrieved from https://sci-hub.se/https://doi.org/10.1080/2330443X.2019.1704330

[1031] Weisburst, Emily K. "Police use of force as an extension of arrests: Examining disparities across civilian and officer race." AEA Papers and Proceedings. Vol. 109. 2019. Retrieved from https://sci-hub.se/https://doi.org/10.1257/pandp.20191028

[1032] Johnson, David J., and Joseph Cesario. "Reply to Knox and Mummolo and Schimmack and Carlsson: Controlling for crime and population rates." Proceedings of the National Academy of Sciences 117.3 (2020): 1264-1265. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.1920184117

[1033] Johnson, David J., et al. "Officer characteristics and racial disparities in fatal officer-involved shootings." Proceedings of the National Academy of Sciences 116.32 (2019): 15877-15882. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.1903856116

[1034] Shjarback, John A., and Justin Nix. "Considering violence against police by citizen race/ethnicity to contextualize representation in officer-involved shootings." Journal of Criminal Justice 66 (2020): 101653. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jcrimjus.2019.101653

[1035] Schwartz, Gabriel L., and Jaquelyn L. Jahn. "Mapping fatal police violence across US metropolitan areas: Overall rates and racial/ethnic inequities, 2013-2017." PloS one 15.6 (2020): e0229686. Retrieved from https://sci-hub.se/https://doi.org/10.1371/journal.pone.0229686

[1036] Streeter, Shea. "Lethal force in black and white: Assessing racial disparities in the circumstances of police killings." The Journal of Politics 81.3 (2019): 1124-1132. Retrieved from https://sci-hub.se/https://doi.org/10.1086/703541

[1037] James, Lois, Bryan Vila, and Kenn Daratha. "Results from experimental trials testing participant responses to White, Hispanic and Black suspects in high-fidelity deadly force judgment and decision-making simulations." Journal of Experimental Criminology 9.2 (2013): 189-212. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s11292-012-9163-y

[1038] James, Lois, Stephen M. James, and Bryan J. Vila. "The reverse racism effect: Are cops more hesitant to shoot black than white suspects?." Criminology & Public Policy 15.2 (2016): 457-479. Retrieved from https://sci-hub.se/https://doi.org/10.1111/1745-9133.12187

[1039] James, Lois, Stephen James, and Bryan Vila. "Does the “reverse racism effect” withstand the test of police officer fatigue?." Policing: An International Journal of Police Strategies & Management (2017). Retrieved from https://sci-hub.se/https://doi.org/10.1108/PIJPSM-01-2016-0006

[1040] Menifield, Charles E., Geiguen Shin, and Logan Strother. "Do white law enforcement officers target minority suspects?." Public Administration Review 79.1 (2019): 56-68. Retrieved from https://sci-hub.se/https://doi.org/10.1111/puar.12956

[1041] Retraction of source 1033: "Retraction for Johnson et al., Officer characteristics and racial disparities in fatal officer-involved shootings". (2020). Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.2014148117

[1042] Knox, Dean, and Jonathan Mummolo. "Making inferences about racial disparities in police violence." Proceedings of the National Academy of Sciences 117.3 (2020): 1261-1262. Retrieved from https://sci-hub.se/https://doi.org/10.1073/pnas.1919418117

[1043] Wikipedia "List of cognitive biases". (2020). Retrieved from https://en.wikipedia.org/wiki/List_of_cognitive_biases

[1044] Kleck, Robert E., and Angelo Strenta. "Perceptions of the impact of negatively valued physical characteristics on social interaction." Journal of Personality and Social Psychology 39.5 (1980): 861. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-3514.39.5.861

[1045] Pew Research Center, June 27, 2016. "On Views of Race and Inequality, Blacks and Whites Are Worlds Apart". Chapter 5: "Personal experiences with discrimination". Retrieved from https://www.pewsocialtrends.org/2016/06/27/5-personal-experiences-with-discrimination/
https://www.pewsocialtrends.org/wp-content/uploads/sites/3/2016/06/ST_2016.06.27_Race-Inequality-Final.pdf

[1046] National Public Radio, Robert Wood Johnson Foundation, and Havard TH Chan School of Public Health. "Discrimination in America: Experiences and views of African Americans." (2017). Retrieved from https://cdn1.sph.harvard.edu/wp-content/uploads/sites/21/2017/10/NPR-RWJF-HSPH-Discrimination-African-Americans-Final-Report.pdf

[1047] Survey of U.S. adults conducted Feb. 29-May 8, 2016. "On Views of Race and Inequality, Blacks and Whites Are Worlds Apart". Pew Research Center. Retrieved from https://www.pewresearch.org/fact-tank/2016/06/29/roughly-half-of-hispanics-have-experienced-discrimination/

[1048] Pew Research Center. "Gender discrimination comes in many forms for today's working women". Survey conducted July. 11-Aug. 10, 2017. Retrieved from https://www.pewresearch.org/fact-tank/2017/12/14/gender-discrimination-comes-in-many-forms-for-todays-working-women/

[1049] Pew Research Center. Survey of U.S. adults conducted Aug. 8-21 and Sept. 14-28, 2017. "Wide Partisan Gaps in U.S. Over How Far the Country Has Come on Gender Equality". Page 6: "Among women, Millennials most likely to see advantages for men". Retrieved from https://www.pewsocialtrends.org/2017/10/18/wide-partisan-gaps-in-u-s-over-how-far-the-country-has-come-on-gender-equality/
https://www.pewsocialtrends.org/wp-content/uploads/sites/3/2017/10/Gender-Equality-Report-FINAL-10.18.pdf

[1050] Pew Research Center. Survey of U.S. adults conducted Jan. 22-Feb. 5, 2019. "For black Americans, experiences of racial discrimination vary by education level, gender". Retrieved from https://www.pewresearch.org/fact-tank/2019/05/02/for-black-americans-experiences-of-racial-discrimination-vary-by-education-level-gender/

[1051] American Dad: "Office Spaceman" - Season 4, Episode 14. Francine Hates Lefties. Retrieved from https://www.youtube.com/watch?v=Piay9dVkIsE&t

[1052] Emotional Thinking. “This is Amerikkka.”. 2019. Retrieved from https://twitter.com/KenidraRWoods_/status/1131471430753898496?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1131471430753898496%7Ctwgr%5E%7Ctwcon%5Es1_&ref_url=https%3A%2F%2Fideasanddata.wordpress.com%2F2019%2F05%2F31%2Flived-experience-is-not-evidence%2F

[1053] Saucier, Donald A., Carol T. Miller, and Nicole Doucet. "Differences in helping whites and blacks: A meta-analysis." Personality and Social Psychology Review 9.1 (2005): 2-16. Retrieved from https://sci-hub.se/https://doi.org/10.1207/s15327957pspr0901_1

[1054] Crosby, Faye, Stephanie Bromley, and Leonard Saxe. "Recent unobtrusive studies of Black and White discrimination and prejudice: A literature review." Psychological bulletin 87.3 (1980): 546. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0033-2909.87.3.546

[1055] Toosi, Negin R., et al. "Dyadic interracial interactions: a meta-analysis." Psychological bulletin 138.1 (2012): 1. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0025767

[1056] Pew Research Center, April 2019, “Race in America 2019”. Retrieved from https://www.pewsocialtrends.org/wp-content/uploads/sites/3/2019/04/Race-report_updated-4.29.19.pdf https://www.pewsocialtrends.org/2019/04/09/the-role-of-race-and-ethnicity-in-americans-personal-lives/

[1057] Charlesworth, Tessa ES, and Mahzarin R. Banaji. "Patterns of implicit and explicit attitudes: I. Long-term change and stability from 2007 to 2016." Psychological science 30.2 (2019): 174-192. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0956797618813087

[1058] Krysan, M., and S. Moberg. "Trends in racial attitudes. University of Illinois Institute of Government and Public Affairs." (2016). Retrieved from https://igpa.uillinois.edu/programs/racial-attitudes#section-6

[1059] Inductivist “Race, ethnicity, and attitudes toward inter-ethnic marriage”. (2010). Retrieved from https://inductivist.blogspot.com/2010/02/race-ethnicity-and-attitudes-toward.html

[1060] Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). "Understanding and using the Implicit Association Test: IV: What we know (so far) about the method". (Pp. 59-102). In B. Wittenbrink & N.S. Schwarz (Eds.). "Implicit measures of attitudes: Procedures and controbersies". New York: Guilford press. Retrieved from https://faculty.washington.edu/agg/pdf/Lane%20et%20al.UUIAT4.OCR.2007.pdf

[1061] Gawronski, Bertram, et al. "Temporal stability of implicit and explicit measures: A longitudinal analysis." Personality and Social Psychology Bulletin 43.3 (2017): 300-312. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0146167216684131

[1062] Stephanie Glen. "Test-Retest Reliability / Repeatability" From StatisticsHowTo.com: Elementary Statistics for the rest of us!. Retrieved from https://www.statisticshowto.com/test-retest-reliability/

[1063] Jost, John T. "The IAT is dead, long live the IAT: Context-sensitive measures of implicit attitudes are indispensable to social and political psychology." Current Directions in Psychological Science 28.1 (2019): 10-19. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0963721418797309

[1064] Carlsson, Rickard, and Jens Agerström. "A closer look at the discrimination outcomes in the IAT literature." Scandinavian journal of psychology 57.4 (2016): 278-287. Retrieved from https://sci-hub.se/https://doi.org/10.1111/sjop.12288

[1065] Oswald, Frederick L., et al. "Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies." Journal of personality and social psychology 105.2 (2013): 171. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0032734

[1066] Chekroud, Adam Mourad, et al. "A review of neuroimaging studies of race-related prejudice: does amygdala response reflect threat?." Frontiers in Human Neuroscience 8 (2014): 179. Retrieved from https://sci-hub.se/https://doi.org/10.3389/fnhum.2014.00179

[1067] Thompson, Jeffrey P., and Gustavo Suarez. "Exploring the racial wealth gap using the survey of consumer finances." Available at SSRN 2665627 (2015). Retrieved from https://www.federalreserve.gov/econresdata/feds/2015/files/2015076pap.pdf

[1068] Source: Thomson Reuters Datastream - Stephen Culp @ReutersCulp 1/9/2015 Retrieved from https://www.businessinsider.com/unemployment-by-race-chart-2015-1

[1069] Hamermesh, Daniel S., Katie R. Genadek, and Michael C. Burda. "Racial/ethnic differences in non-work at work." ILR Review (2017): 0019793919891429. Retrieved from http://ftp.iza.org/dp10496.pdf

[1070] Jussim, Lee, Jarret T. Crawford, and Rachel S. Rubinstein. "Stereotype (in) accuracy in perceptions of groups and individuals." Current Directions in Psychological Science 24.6 (2015): 490-497. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0963721415605257

[1071] Pew Research Center analysis of 2014-2015 American Community Survey (IPUMS). "Intermarriage in the U.S. 50 Years After Loving v. Virginia". Retrieved from https://www.pewresearch.org/fact-tank/2017/06/12/key-facts-about-race-and-marriage-50-years-after-loving-v-virginia/

[1072] Joyner, Kara, and Grace Kao. "Interracial relationships and the transition to adulthood." American Sociological Review 70.4 (2005): 563-581. Retrieved from https://sci-hub.st/https://doi.org/10.1177/000312240507000402

[1073] Bramlett MD and Mosher WD. “Cohabitation, Marriage, Divorce, and Remarriage in the United States”. National Center for Health Statistics. Vital Health Stat 23(22). 2002. Retrieved from https://www.cdc.gov/nchs/data/series/sr_23/sr23_022.pdf

[1074] Fusco, Rachel A. "Intimate partner violence in interracial couples: A comparison to White and ethnic minority monoracial couples." Journal of interpersonal violence 25.10 (2010): 1785-1800. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0886260509354510

[1075] Wakefield, Juliet Ruth Helen, et al. "The relationship between group identification and satisfaction with life in a cross-cultural community sample." Journal of Happiness Studies 18.3 (2017): 785-807. Retrieved from https://sci-hub.se/https://doi.org/10.1007/s10902-016-9735-z

[1076] Bécares, Laia, Michael E. Dewey, and Jayati Das-Munshi. "Ethnic density effects for adult mental health: systematic review and meta-analysis of international studies." Psychological medicine 48.12 (2018): 2054. Retrieved from https://sci-hub.se/https://doi.org/10.1017/S0033291717003580

[1077] Serra-Garcia, M., & Gneezy, U. (2021). Nonreplicable publications are cited more than replicable ones. *Science advances*, 7(21), eabd1705. Retrieved from https://sci-hub.se/https://doi.org/10.1126/sciadv.abd1705

[1078] Nir, Esther, and Elizabeth Griffiths. "Sentencing on the Evidence." Criminal Justice Policy Review 29.4 (2018): 365-390. Retrieved from https://scihubtw.tw/https://doi.org/10.1177/0887403416635248

[1079] Ellis, Lee, David P. Farrington, and Anthony W. Hoskin. Handbook of crime correlates. Academic Press, 2019. Retrieved from https://b-ok.cc/book/996636/ad2721

[1080] Chiricos, Theodore G. "Rates of crime and unemployment: An analysis of aggregate research evidence." Social problems 34.2 (1987): 187-212. Retrieved from https://sci-hub.se/https://doi.org/10.2307/800715

[1081] Hsieh, Ching-Chi, and Meredith D. Pugh. "Poverty, income inequality, and violent crime: a meta-analysis of recent aggregate data studies." Criminal justice review 18.2 (1993): 182-202. Retrieved from https://sci-hub.se/https://doi.org/10.1177/073401689301800203

[1082] Pratt, Travis C., and Francis T. Cullen. "Assessing macro-level predictors and theories of crime: A meta-analysis." Crime and justice 32 (2005): 373-450. Retrieved from https://sci-hub.se/https://doi.org/10.1086/655357

[1083] Nivette, Amy E. "Cross-national predictors of crime: A meta-analysis." Homicide Studies 15.2 (2011): 103-131. Retrieved from https://sci-hub.se/https://doi.org/10.1177/1088767911406397

[1084] Rufrancos, Hector, et al. "Income inequality and crime: A review and explanation of the timeâ series evidence." Sociology and Criminology-Open Access (2013). Retrieved from https://www.longdom.org/open-access/income-inequality-and-crime-a-review-and-explanation-of-the-timeseries-evidence-2375-4435.1000103.pdf

[1085] Sariaslan, Amir, et al. "Childhood family income, adolescent violent criminality and substance misuse: quasi-experimental total population study." The British Journal of Psychiatry 205.4 (2014): 286-290. Retrieved from https://sci-hub.se/https://doi.org/10.1192/bjp.bp.113.136200

[1086] U.S. Census Bureau. Current Population Survey. Annual Social and Economic Supplement, 2005 - 2011. Table: Number of related children under 18 by Race. Persons - All. Percentages by Race. Sums in Whole Numbers. CPS 7-year average - Data Collected in 2005 to 2011. Retrieved from https://www.census.gov/cps/data/cpstablecreator.html

[1087] Ruggles, Steve. "The origins of African-American family structure." American Sociological Review (1994): 136-151. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2096137

[1088] Petrosino, Anthony., Derzon, James., Lavenberg, Julia. "The role of the family in crime and delinquency: Evidence from prior quantitative reviews." Southwest Journal of Criminal Justice 6.2 (2009): 108-132. Retrieved from https://www.researchgate.net/profile/James_Derzon/publication/224049275_the_role_of_the_family_in_Crime_and_Delinquency_evidence_from_Prior_Quantitative_reviews/links/09e4150c8cfda3f140000000.pdf

[1089] Price, Cynthia, and Jenifer Kunz. "Rethinking the paradigm of juvenile delinquency as related to divorce." Journal of Divorce & Remarriage 39.1-2 (2003): 109-133. Retrieved from https://sci-hub.se/https://doi.org/10.1300/J087v39n01_07

[1090] Huang, Chien-Chung, and Lynn A. Warner. "Relationship characteristics and depression among fathers with newborns." Social Service Review 79.1 (2005): 95-118. Retrieved from https://scihubtw.tw/https://doi.org/10.1086/426719

[1091] U. S. Bureau of the Census.Current Population Reports. P23-180. "Marriage, Divorce, and Remarriage in the 1990's". U.S. Government Printing Office, Washington, DC. 1992. Retrieved from https://www2.census.gov/library/publications/1992/demographics/p23-180.pdf

[1092] Jaffee, Sara R., et al. "Predicting early fatherhood and whether young fathers live with their children: Prospective findings and policy reconsiderations." Journal of child psychology and psychiatry 42.6 (2001): 803-815. Retrieved from https://scihubtw.tw/https://doi.org/10.1111/1469-7610.00777

[1093] Montare, Alberto, and Sherle L. Boone. "Aggression and paternal absence: Racial-ethnic differences among inner-city boys." The Journal of genetic psychology 137.2 (1980): 223-232. Retrieved from https://sci-hub.se/https://doi.org/10.1080/00221325.1980.10532821

[1094] Marcus, David K., Jessica J. Fulton, and Erin J. Clarke. "Lead and conduct problems: a meta-analysis." Journal of Clinical Child & Adolescent Psychology 39.2 (2010): 234-241. Retrieved from https://sci-hub.se/https://doi.org/10.1080/15374411003591455

[1095] Stretesky, Paul B., and Michael J. Lynch. "The relationship between lead exposure and homicide." Archives of pediatrics & adolescent medicine 155.5 (2001): 579-582. Retrieved from https://sci-hub.se/https://doi.org/10.1001/archpedi.155.5.579

[1096] Stretesky, Paul B., and Michael J. Lynch. "The relationship between lead and crime." Journal of Health and Social Behavior 45.2 (2004): 214-229. Retrieved from https://sci-hub.se/https://doi.org/10.2307/3653840

[1097] Feigenbaum, James J., and Christopher Muller. "Lead exposure and violent crime in the early twentieth century." Explorations in economic history 62 (2016): 51-86. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/j.eeh.2016.03.002

[1098] Nevin, Rick. "Understanding international crime trends: the legacy of preschool lead exposure." Environmental research 104.3 (2007): 315-336. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.envres.2007.02.008

[1099] U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. (2015). Child maltreatment 2013. Retrieved from https://www.acf.hhs.gov/sites/default/files/documents/cb/cm2013.pdf

[1100] Currie, Janet, and Erdal Tekin. Does child abuse cause crime?. No. w12171. National Bureau of Economic Research, 2006. Retrieved from https://sci-hub.se/https://doi.org/10.3386/w12171

[1101] Widom, Cathy S., and Michael G. Maxfield. "An Update on the" Cycle of Violence: Research in Brief". U.S. Department of Justice, Office of Justice Programs, National Institute of Justice (2001). Retrieved from https://www.ncjrs.gov/pdffiles1/nij/184894.pdf

[1102] Maas, Carl, Todd I. Herrenkohl, and Cynthia Sousa. "Review of research on child maltreatment and violence in youth." Trauma, Violence, & Abuse 9.1 (2008): 56-67. Retrieved from https://sci-hub.se/https://doi.org/10.1177/1524838007311105

[1103] Carson, E. Ann, and Elizabeth Anderson. "Prisoners in 2014." US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, NCJ 247282 (2015): 2. Retrieved from https://www.bjs.gov/content/pub/pdf/p14.pdf

[1104] Book, Angela S., Katherine B. Starzyk, and Vernon L. Quinsey. "The relationship between testosterone and aggression: A meta-analysis." Aggression and Violent Behavior 6.6 (2001): 579-599. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S1359-1789(00)00032-X

[1105] Tricker, R., et al. "The effects of supraphysiological doses of testosterone on angry behavior in healthy eugonadal men--a clinical research center study." The Journal of Clinical Endocrinology & Metabolism 81.10 (1996): 3754-3758. Retrieved from https://sci-hub.se/https://doi.org/10.1210/jcem.81.10.8855834

[1106] O'Connor, Daryl B., et al. "Exogenous testosterone, aggression, and mood in eugonadal and hypogonadal men." Physiology & behavior 75.4 (2002): 557-566. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/s0031-9384(02)00647-9

[1107] Batrinos, Menelaos L. "Testosterone and aggressive behavior in man." International journal of endocrinology and metabolism 10.3 (2012): 563. Retrieved from https://sci-hub.se/https://doi.org/10.5812/ijem.3661

[1108] Child Trends Databank. (2018). Physical fighting by youth. Retrieved from https://www.childtrends.org/?indicators=physical-fighting-by-youth

[1109] Wang, Weijun. "Bullying among US school children: An examination of race/ethnicity and school-level variables on bullying." (2013). Retrieved from https://tigerprints.clemson.edu/cgi/viewcontent.cgi?article=2204&context=all_dissertations

[1110] Moffitt, Terrie E., et al. "A gradient of childhood self-control predicts health, wealth, and public safety." Proceedings of the national Academy of Sciences 108.7 (2011): 2693-2698. Retrieved from https://scihubtw.tw/https://doi.org/10.1073/pnas.1010076108

[1111] Minkov, Michael, and Michael Harris Bond. "Genetic polymorphisms predict national differences in life history strategy and time orientation." Personality and Individual Differences 76 (2015): 204-215. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.paid.2014.12.014

[1112] Lynn, Richard. "Racial and ethnic differences in psychopathic personality." Personality and Individual Differences 32.2 (2002): 273-316. Retrieved from https://sci-hub.se/https://doi.org/10.1016/S0191-8869(01)00029-0

[1113] Skeem, Jennifer L., et al. "Are there ethnic differences in levels of psychopathy? A meta-analysis." Law and human behavior 28.5 (2004): 505-527. Retrieved from https://sci-hub.se/https://doi.org/10.1023/B:LAHU.0000046431.93095.d8

[1114] McCoy, Wendy K., and John F. Edens. "Do black and white youths differ in levels of psychopathic traits? A meta-analysis of the psychopathy checklist measures." Journal of consulting and clinical psychology 74.2 (2006): 386. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0022-006X.74.2.386

[1115] Attention problems, inhibitory control, and intelligence index overlapping genetic factors: a study in 9-, 12-, and 18-year-old twins Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0014915

[1116] Alexander, Karl, and Stephen L. Morgan. "The Coleman Report at fifty: Its legacy and implications for future research on equality of opportunity." RSF: The Russell Sage Foundation Journal of the Social Sciences 2.5 (2016): 1-16. Retrieved from http://socweb.soc.jhu.edu/faculty/morgan/papers/Alexander_and_Morgan_2016.pdf

[1117] Willems, Y. E., et al. "The heritability of self-control: A meta-analysis." Neuroscience & Biobehavioral Reviews 100 (2019): 324-334. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.neubiorev.2019.02.012

[1118] Anokhin, Andrey P., et al. "The genetics of impulsivity: evidence for the heritability of delay discounting." Biological psychiatry 77.10 (2015): 887-894. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/j.biopsych.2014.10.022

[1119] Isen, Joshua D., Jordan C. Sparks, and William G. Iacono. "Predictive validity of delay discounting behavior in adolescence: a longitudinal twin study." Experimental and clinical psychopharmacology 22.5 (2014): 434. Retrieved from https://sci-hub.st/https://doi.org/10.1037/a0037340

[1120] Daly, Michael, et al. "Childhood self-control and unemployment throughout the life span: Evidence from two British cohort studies." Psychological science 26.6 (2015): 709-723. Retrieved from https://scihubtw.tw/https://doi.org/10.1177/0956797615569001

[1121] Duckworth, Angela L., and Martin EP Seligman. "Self-discipline outdoes IQ in predicting academic performance of adolescents." Psychological science 16.12 (2005): 939-944. Retrieved from https://scihubtw.tw/https://doi.org/10.1111/j.1467-9280.2005.01641.x

[1122] Wang, Mei, Marc Oliver Rieger, and Thorsten Hens. "How time preferences differ: Evidence from 53 countries." Journal of Economic Psychology 52 (2016): 115-135. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.joep.2015.12.001

[1123] De Ridder, Denise TD, et al. "Taking stock of self-control: A meta-analysis of how trait self-control relates to a wide range of behaviors." Personality and Social Psychology Review 16.1 (2012): 76-99. Retrieved from https://sci-hub.se/https://doi.org/10.1177/1088868311418749

[1124] Warner, John T., and Saul Pleeter. "The personal discount rate: Evidence from military downsizing programs." American Economic Review 91.1 (2001): 33-53. Retrieved from https://sci-hub.se/https://doi.org/10.1257/aer.91.1.33

[1125] Andrade, Leonardo F., and Nancy M. Petry. "White problem gamblers discount delayed rewards less steeply than their African American and Hispanic counterparts." Psychology of Addictive Behaviors 28.2 (2014): 599. Retrieved from https://sci-hub.se/https://doi.org/10.1037/a0036153

[1126] Castillo, Marco, et al. "The today and tomorrow of kids: Time preferences and educational outcomes of children." Journal of Public Economics 95.11-12 (2011): 1377-1385. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/j.jpubeco.2011.07.009

[1127] Herzberger, Sharon D., and Carol S. Dweck. "Attraction and delay of gratification." Journal of Personality 46.2 (1978): 215-227. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1467-6494.1978.tb00176.x

[1128] Kain, John F., and Kraig Singleton. "Equality of educational opportunity revisited." New England Economic Review Special issue (1996): 87. Retrieved from https://search.proquest.com/openview/768bdcc3f49478bd10c04f2a2181750f/ direct download: https://www.bostonfed.org/-/media/Documents/economic/neer/neer1996/neer396f.pdf

[1129] Daly, Mary, Bart Hobijn, and Joseph H. Pedtke. "Disappointing facts about the black-white wage gap." FRBSF Economic Letter 26 (2017): 1-5. Retrieved from https://www.frbsf.org/economic-research/files/el2017-26.pdf

[1130] Sacerdote, Bruce. "Slavery and the intergenerational transmission of human capital." Review of Economics and Statistics 87.2 (2005): 217-234. Retrieved from https://sci-hub.se/https://doi.org/10.1162/0034653053970230

[1131] Bleakley, Hoyt, and Joseph Ferrie. "Shocking behavior: Random wealth in antebellum Georgia and human capital across generations." The quarterly journal of economics 131.3 (2016): 1455-1495. Retrieved from https://europepmc.org/backend/ptpmcrender.fcgi?accid=PMC5436311&blobtype=pdf

[1132] Ager, Philipp, Leah Platt Boustan, and Katherine Eriksson. "Inter-generational transmission of wealth shocks: Evidence from the US Civil War." (2016). Retrieved from https://www.eh.net/eha/wp-content/uploads/2016/08/AgerBoustanEriksson.pdf

[1133] Chetty, Raj, et al. "Race and economic opportunity in the United States: An intergenerational perspective." The Quarterly Journal of Economics 135.2 (2020): 711-783. Retrieved from https://sci-hub.se/https://doi.org/10.1093/qje/qjz042

[1134] Bayer, Patrick, and Kerwin Kofi Charles. "Divergent paths: A new perspective on earnings differences between black and white men since 1940." The Quarterly Journal of Economics 133.3 (2018): 1459-1501. Retrieved from https://sci-hub.se/https://doi.org/10.1093/qje/qjy003

[1135] Fairlie, Robert W., and William A. Sundstrom. "The emergence, persistence, and recent widening of the racial unemployment gap." ILR Review 52.2 (1999): 252-270. Retrieved from https://sci-hub.se/https://doi.org/10.1177/001979399905200206

[1136] Pew Research Center analysis of Home Mortgage Disclosure Act data. "Blacks and Hispanics face extra challenges in getting home loans". 2017. Retrieved from https://www.pewresearch.org/fact-tank/2017/01/10/blacks-and-hispanics-face-extra-challenges-in-getting-home-loans/

[1137] Klein, Stephen P., and John E. Rolph. "Relationship of offender and victim race to death penalty sentences in California." Jurimetrics J. 32 (1991): 33. Retrieved from https://sci-hub.se/https://doi.org/10.2307/29762240

[1138] Corzine, Jon., Codey, Richard J., and Roberts, Joseph J Jr. New Jersey. Death Penalty Study Commission. New Jersey death penalty study commission report. New Jersey Death Penalty Study Commission, 2007. Retrieved from https://www.njleg.state.nj.us/committees/dpsc_final.pdf

[1139] Walsh, Anthony, and Virginia Hatch. "Ideology, Race, and the Death Penalty:" Lies, Damn Lies, and Statistics" in Advocacy Research." Journal of Ideology 37.1 (2017): 2. Retrieved from https://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=1006&context=ji

[1140] Paternoster, Raymond, et al. An empirical analysis of Maryland's death sentencing system with respect to the influence of race and legal jurisdiction. University of Maryland, College Park, 2003. Retrieved from https://www.aclu-md.org/sites/default/files/field_documents/md_death_penalty_race_study.pdf

[1141] Jennings, Wesley G., et al. "A critical examination of the "White victim effect" and death penalty decision-making from a propensity score matching approach: The North Carolina experience." Journal of Criminal Justice 42.5 (2014): 384-398. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jcrimjus.2014.05.004

[1142] Beaver, Kevin M., et al. "No evidence of racial discrimination in criminal justice processing: Results from the National Longitudinal Study of Adolescent Health." Personality and Individual Differences 55.1 (2013): 29-34. Retrieved from https://sci-hub.st/https://doi.org/10.1016/j.paid.2013.01.020

[1143] van Ginkel, Joost R., et al. "Young offenders caught in the act: A population-based cohort study comparing internationally adopted and non-adopted adolescents." Children and Youth Services Review 95 (2018): 32-41. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/j.childyouth.2018.10.009

[1144] Zhang, Yuanting, and Jennifer Van Hook. "Marital dissolution among interracial couples." Journal of Marriage and Family 71.1 (2009): 95-107. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1741-3737.2008.00582.x

[1145] Bratter, Jenifer L., and Rosalind B. King. ""But will it last?": Marital instability among interracial and same-race couples." Family Relations 57.2 (2008): 160-171. Retrieved from https://sci-hub.se/https://doi.org/10.1111/j.1741-3729.2008.00491.x

[1146] Wang, Hongyu, Grace Kao, and Kara Joyner. "Stability of interracial and intraracial romantic relationships among adolescents." Social Science Research 35.2 (2006): 435-453. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.ssresearch.2004.10.001

[1147] Fu, Xuanning. "An interracial study of marital disruption in Hawaii: 1983 to 1996." Journal of Divorce & Remarriage 32.3-4 (2000): 73-92. Retrieved from https://sci-hub.se/https://doi.org/10.1300/J087v32n03_04

[1148] Bratter, Jenifer L., and Karl Eschbach. "'What about the couple?'Interracial marriage and psychological distress." Social Science Research 35.4 (2006): 1025-1047. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.ssresearch.2005.09.001

[1149] Gignac, Gilles E. "Dynamic mutualism versus g factor theory: An empirical test." Intelligence 42 (2014): 89-97. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/j.intell.2013.11.004

[1150] Doebler, Philipp, and Barbara Scheffler. "The relationship of choice reaction time variability and intelligence: A meta-analysis." *Learning and Individual Differences* 52 (2016): 157-166. Retrieved from https://sci-hub.st/https://doi.org/10.1016/j.lindif.2015.02.009

[1151] Carroll, John B. "The higher-stratum structure of cognitive abilities: Current evidence supports g and about ten broad factors." The scientific study of general intelligence (2003): 5-21. Retrieved from https://sci-hub.se/https://doi.org/10.1016/B978-008043793-4/50036-2

[1152] Jewsbury, Paul A., Stephen C. Bowden, and Kevin Duff. "The Cattell–Horn–Carroll model of cognition for clinical assessment." *Journal of Psychoeducational Assessment* 35.6 (2017): 547-567. Retrieved from https://sci-hub.se/https://doi.org/10.1177/0734282916651360

[1153] Kan, Kees-Jan, Han LJ van der Maas, and Stephen Z. Levine. "Extending psychometric network analysis: Empirical evidence against g in favor of mutualism?." *Intelligence* 73 (2019): 52-62. Retrieved from https://sci-hub.st/https://doi.org/10.1016/j.intell.2018.12.004

[1154] Dubois, Julien, et al. "A distributed brain network predicts general intelligence from resting-state human neuroimaging data." *Philosophical Transactions of the Royal Society B: Biological Sciences* 373.1756 (2018): 20170284. Retrieved from https://sci-hub.se/https://doi.org/10.1098/rstb.2017.0284

[1155] Cox, S. R., et al. "Structural brain imaging correlates of general intelligence in UK Biobank." *Intelligence* 76 (2019): 101376. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2019.101376

[1156] Gur, Ruben C., et al. "Structural and functional brain parameters related to cognitive performance across development: Replication and extension of the parieto-frontal integration theory in a single sample." *Cerebral Cortex* 31.3 (2021): 1444-1463. Retrieved from https://sci-hub.se/https://doi.org/10.1093/cercor/bhaa282

[1157] Wlaszczyk, Agata, et al. "Predicting Fluid Intelligence from Structural MRI Using Random Forest regression." *Challenge in Adolescent Brain Cognitive Development Neurocognitive Prediction*. Springer, Cham, 2019. Retrieved from https://sci-hub.se/https://doi.org/10.1007/978-3-030-31901-4_10

[1158] Malanchini, Margherita, et al. "Pathfinder: A gamified measure to integrate general cognitive ability into the biological, medical and behavioural sciences." *bioRxiv* (2021). Retrieved from https://www.biorxiv.org/content/10.1101/2021.02.10.430571v1

[1159] Vazsonyi, Alexander T., Jakub Mikuška, and Erin L. Kelley. "It's time: A meta-analysis on the self-control-deviance link." *Journal of Criminal Justice* 48 (2017): 48-63. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.jcrimjus.2016.10.001

[1160] D'Acunto, Francesco, et al. *IQ, expectations, and choice*. No. w25496. National Bureau of Economic Research, 2019. Retrieved from https://sci-hub.se/https://doi.org/10.3386/w25496

[1161] Grinblatt, Mark, Matti Keloharju, and Juhani T. Linnainmaa. "IQ, trading behavior, and performance." *Journal of Financial Economics* 104.2 (2012): 339-362. Retrieved from https://sci-hub.st/https://doi.org/10.1016/j.jfineco.2011.05.016

[1162] Coyle, Thomas R. "Non-g factors predict educational and occupational criteria: More than g." *Journal of Intelligence* 6.3 (2018): 43. Retrieved from https://scihubtw.tw/https://doi.org/10.3390/jintelligence6030043

[1163] Burt, S. Alexandra, et al. "Twin Differences in Harsh Parenting Predict Youth's Antisocial Behavior." *Psychological Science*(2021): 0956797620968532. Retrieved from (not)

[1164] Black, D. A., Kolesnikova, N., Sanders, S. G., & Taylor, L. J. (2013). Are children "normal"?. The review of economics and statistics, 95(1), 21-33. Retrieved from https://scihubtw.tw/https://doi.org/10.1162/REST_a_00257

[1165] Lovenheim, M. F., & Mumford, K. J. (2013). Do family wealth shocks affect fertility choices? Evidence from the housing market. *Review of Economics and Statistics*, *95*(2), 464-475. Retrieved from https://scihubtw.tw/https://doi.org/10.1162/REST_a_00266

[1166] Protzko, J., Krosnick, J., Nelson, L. D., Nosek, B. A., Axt, J., Berent, M., ... & Schooler, J. (2020). High Replicability of Newly-Discovered Social-behavioral Findings is Achievable. Retrieved from https://psyarxiv.com/n2a9x

[1167] Simonsohn, Uri. "Just post it: The lesson from two cases of fabricated data detected by statistics alone." Psychological science 24.10 (2013): 1875-1888. Retrieved from https://scihubtw.tw/https://doi.org/10.1177/0956797613480366

[1168] Smith Jr, N., & Cumberledge, A. (2020). Quotation errors in general science journals. Proceedings of the Royal Society A, 476(2242), 20200538. Retrieved from https://scihubtw.tw/https://doi.org/10.1098/rspa.2020.0538

[1169] Panagiotou, O. A., & Ioannidis, J. P. (2012). Primary study authors of significant studies are more likely to believe that a strong association exists in a heterogeneous meta-analysis compared with methodologists. Journal of clinical epidemiology, 65(7), 740-747. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/j.jclinepi.2012.01.008

[1170] Einhorn, Hillel J. "Expert measurement and mechanical combination." Organizational behavior and human performance 7.1 (1972): 86-106. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/0030-5073(72)90009-8

[1171] Christensen-Szalanski, J. J., & Bushyhead, J. B. (1981). Physicians' use of probabilistic information in a real clinical setting. Journal of Experimental Psychology: Human perception and performance, 7(4), 928. Retrieved from https://sci-hub.se/https://doi.org/10.1037/0096-1523.7.4.928

[1172] Lee, K. L., Pryor, D. B., Harrell Jr, F. E., Califf, R. M., Behar, V. S., Floyd, W. L., ... & Rosati, R. A. (1986). Predicting outcome in coronary disease statistical models versus expert clinicians. The American journal of medicine, 80(4), 553-560. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/0002-9343(86)90807-7

[1173] Bordalo, Pedro, et al. "Diagnostic expectations and stock returns." The Journal of Finance 74.6 (2019): 2839-2874. Retrieved from https://scihubtw.tw/https://doi.org/10.1111/jofi.12833

[1174] Braun, P. A., & Yaniv, I. (1992). A case study of expert judgment: Economists' probabilities versus base-rate model forecasts. Journal of behavioral decision making, 5(3), 217-231. Retrieved from https://scihubtw.tw/https://doi.org/10.1002/bdm.3960050306

[1175] DellaVigna, S., & Pope, D. (2016). Predicting Experimental Results: Who Knows What? Retrieved from https://scihubtw.tw/https://doi.org/10.3386/w22566

[1176] Ruger, T. W., Kim, P. T., Martin, A. D., & Quinn, K. M. (2004). The supreme court forecasting project: Legal and political science approaches to predicting supreme court decision making. Columbia Law Review, 1150-1210. Retrieved from https://scihubtw.tw/https://doi.org/10.2307/4099370

[1177] Protzko, J. (2020). Kids These Days! Increasing delay of gratification ability over the past 50 years in children. Intelligence, 80, 101451. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/j.intell.2020.101451

[1178] Hutcherson, C., Sharpinskyi, K., Varnum, M. E., Rotella, A., Wormley, A., Tay, L., & Grossmann, I. (2021). The pandemic fallacy: Inaccuracy of social scientists' and lay judgments about COVID-19's societal consequences in America. Retrieved from https://psyarxiv.com/g8f9s/

[1179] Mandel, D. R., & Barnes, A. (2014). Accuracy of forecasts in strategic intelligence. Proceedings of the National Academy of Sciences, 111(30), 10984-10989. Retrieved from https://scihubtw.tw/https://doi.org/10.1073/pnas.1406138111

[1180] Tetlock, P., & Mellers, B. (2014). Judging political judgment. Proceedings of the National Academy of Sciences, 111(32), 11574-11575. Retrieved from https://scihubtw.tw/https://doi.org/10.1073/pnas.1412524111

[1181] Tetlock, P. E. (2009). Expert political judgment. Princeton University Press. Retrieved from https://b-ok.cc/book/1101127/5d24c5

[1182] void

[1183] void

[1184] void

[1185] void

[1186] void

[1187] Pontille, D., & Torny, D. (2015). From manuscript evaluation to article valuation: the changing technologies of journal peer review. Human Studies, 38(1), 57-79. Retrieved from https://scihubtw.tw/https://doi.org/10.1007/s10746-014-9335-z

[1188] Baldwin, M. (2018). Scientific autonomy, public accountability, and the rise of "peer review" in the Cold War United States. Isis, 109(3), 538-558. Retrieved from https://scihubtw.tw/https://doi.org/10.1086/700070

[1189] Caplan, B. (2019). The Case Against Education. Princeton University Press. Retrieved from https://b-ok.cc/book/6038223/4aac1c

[1190] Light, A., & Strayer, W. (2004). Who receives the college wage premium? Assessing the labor market returns to degrees and college transfer patterns. Journal of Human Resources, 39(3), 746-773. Retrieved from https://scihubtw.tw/https://doi.org/10.2307/3558995

[1191] Kane, T. J., & Rouse, C. E. (1993). Labor market returns to two-and four-year colleges: is a credit a credit and do degrees matter? (No. w4268). National Bureau of Economic Research. Retrieved from https://scihubtw.tw/https://doi.org/10.2307/2118190

[1192] Raudenbush, S., & Kasim, R. (1998). Cognitive skill and economic inequality: Findings from the National Adult Literacy Survey. Harvard Educational Review, 68(1), 33-80. Retrieved from https://scihubtw.tw/https://doi.org/10.17763/haer.68.1.1j47150021346123

[1193] Arkes, J. (1999). What do educational credentials signal and why do employers value credentials?. Economics of Education Review, 18(1), 133-141. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/S0272-7757(98)00024-7

[1194] Riddell, C. W. (2008). Understanding ‘sheepskin effects’ in the returns to education: The role of cognitive skills. Canada: Department of Economics, University of Toronto. Retrieved from http://www.clsrn.econ.ubc.ca/hrsdc/papers/Paper%20no.%202%20-%20Craig%20Riddell%20-%20Sheepskin%20Effects.pdf

[1195] Frazis, H. (1993). Selection bias and the degree effect. Journal of Human Resources, 538-554. Retrieved from https://scihubtw.tw/https://doi.org/10.2307/146159

[1196] Te Nijenhuis, J., Choi, Y. Y., van den Hoek, M., Valueva, E., & Lee, K. H. (2019). Spearman’s hypothesis tested comparing Korean young adults with various other groups of young adults on the items of the Advanced Progressive Matrices. Journal of biosocial science, 51(6), 875-912. Retrieved from https://scihubtw.tw/https://doi.org/10.1017/S0021932019000026

[1197] Card, D. (1999). The causal effect of education on earnings. Handbook of labor economics, 3, 1801-1863. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/S1573-4463(99)03011-4

[1198] Sandewall, Ö., Cesarini, D., & Johannesson, M. (2014). The co-twin methodology and returns to schooling—Testing a critical assumption. Labour Economics, 26, 1-10. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/j.labeco.2013.10.002

[1199] Bound, J., & Solon, G. (1999). Double trouble: on the value of twins-based estimation of the return to schooling. Economics of Education Review, 18(2), 169-182. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/S0272-7757(98)00048-X

[1200] Neumark, D. (1999). Biases in twin estimates of the return to schooling. Economics of Education Review, 18(2), 143-148. Retrieved from https://scihubtw.tw/https://doi.org/10.1016/S0272-7757(97)00022-8

[1201] Angrist, J. D., & Pischke, J. S. (2014). Mastering 'Metrics: The path from cause to effect. Princeton University Press. Retrieved from https://b-ok.cc/book/5286393/bc7d83

[1202] Bowles, S., Gintis, H., & Osborne, M. (2001). The determinants of earnings: A behavioral approach. Journal of economic literature, 39(4), 1137-1176. Retrieved from https://scihubtw.tw/https://doi.org/10.1257/jel.39.4.1137

[1203] Pew Research Center, January, 2019. Public’s 2019 Priorities: Economy, Health Care, Education and Security All Near Top of List. Retrieved from https://www.pewresearch.org/politics/wp-content/uploads/sites/4/2019/01/PP_2019.01.24_political-priorities_FINAL.pdf

[1204] Avdeev, Stanislav. Zero Returns To Higher Education: Evidence From A Natural Experiment. Higher School of Economics Research Paper No. WP BRP 236 (2020). Retrieved from https://wp.hse.ru/data/2020/09/22/1584505319/236EC2020.pdf

[1205] Hérault, N., & Zakirova, R. (2015). Returns to education: accounting for enrolment and completion effects. Education Economics, 23(1), 84-100. Retrieved from https://scihubtw.tw/https://doi.org/10.1080/09645292.2013.805184

[1206] Flores-Lagunes, A., & Light, A. (2010). Interpreting degree effects in the returns to education. Journal of Human Resources, 45(2), 439-467. Retrieved from https://scihubtw.tw/https://doi.org/10.3368/jhr.45.2.439

[1207] void

[1208] void

[1209] Altonji, J. G., & Dunn, T. A. (1996). The Effects of Family Characteristics on the Return to Education. The Review of Economics and Statistics, 692-704. Retrieved from https://sci-hub.se/https://doi.org/10.2307/2109956

[1210] Floyd, R. G., Shands, E. I., Rafael, F. A., Bergeron, R., & McGrew, K. S. (2009). The dependability of general-factor loadings: The effects of factor-extraction methods, test battery composition, test battery size, and their interactions. *Intelligence*, 37(5), 453-465. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2009.05.003

[1211] Major, J. T., Johnson, W., & Bouchard Jr, T. J. (2011). The dependability of the general factor of intelligence: Why small, single-factor models do not adequately represent g. *Intelligence*, 39(5), 418-433. Retrieved from https://sci-hub.se/https://doi.org/10.1016/j.intell.2011.07.002

[1212] Hoff, P. D. (2009). A first course in Bayesian statistical methods (Vol. 580). New York: Springer. Retrieved from https://b-ok.cc/book/609128/35746a

[1213] Pearl, J., Glymour, M., & Jewell, N. P. (2016). Causal inference in statistics: A primer. John Wiley & Sons. Retrieved from https://b-ok.cc/book/2664651/adcbf6

[1214] Glynn, A. N., & Kashin, K. (2017). Front-door versus back-door adjustment with unmeasured confounding: Bias formulas for front-door and hybrid adjustments with application to a job training program. Journal of the American Statistical Association, 113(523), 1040-1049. Retrieved from https://sci-hub.se/https://doi.org/10.1080/01621459.2017.1398657

[Epic] BLM. “False Information: Reviewed By Independent Fact Checkers”. 2020. Retrieved from https://imgur.com/a/DFVuCoH (https://imgur.com/7Cf14Iz)

Temp bottom of list

Note: 667, 673, 706, 752, 787, 848, 1163,