Intelligence 70 (2018) 73–83

Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell

Reflections about intelligence over 40 years

James R. Flynn

University of Otago, New Zealand

ARTICLE INFO

Keywords:
Flynn effect
Potency of environment
The nature of *g*
Genes and human autonomy

ABSTRACT

In 1977, I read Arthur Jensen about how the weakness of environment within groups suggested the same weakness between groups. This meant, in his mind, that no environmental explanation of the IQ gap between black and white was possible. It raised problems, in my mind, about whether we could ignore the role of environment between generations.

In 1987, I showed that there had been massive IQ gains (later called the "Flynn Effect"). Jensen (1998) responded that most of these gains were not true intelligence (or *g*) gains. In 2000, Dickens and Flynn showed that environment could be powerful between groups despite its weakness within groups. By 2007, I could show that even though IQ gains were not *g* gains, this did not undermine their significance and actually raised questions about the significance of *g*. Today, I can offer a range of insights: how to measure the factors that enhance IQ over time; how parents try to be fair to their children; how age affects cognitive competence; what signs environment gives of its ever-present causal potency; how the genetic determination (that twin studies imply) leaves room for human autonomy.

1. Introduction

I offer a summary of what I have learned about intelligence between 1977 and 2017. First, I will detail what I learned from my long debate with Arthur Jensen. He was irreplaceable in the sense that you often learn more from a thinker who challenges opinions you tend to take for granted. They shake you out of your dogmatic slumbers – force you to defend assumptions made unreflectively that cry out for clarification.

1.1. Jensen's arithmetic

In 1969, Arthur Jensen became the center of controversy when his article "How much can we boost IQ and academic achievement?" appeared in the *Harvard Educational Review* (Jensen, 1969). The controversy became more intense a few years later, when he reiterated the hypothesis "that something between one-half and three-fourths of the average IQ difference between American Negroes and Whites is attributable to genetic factors" (Jensen, 1973, p. 363). At Berkeley, the reaction took the form of several weeks of violent demonstrations and protests continued to flare periodically until his death in 2012. It may be that no one is entirely free of racial bias. But I never detected any in my conversations with Jensen. He did have a trait scholars should emulate: he never wavered in stating what he saw as the truth. He used to quote Mahatma Gandhi, one of his heroes, who pledged that he would always say in public what he believed in private.

I first read Jensen in 1977. I was emotionally involved in the question of the role of genes and environment in explaining the black-white IQ gap, but one of Jensen's arguments unsettled me on a far more fundamental level. He appeals to the twin studies, particularly those of identical twins (who have the same genes) but who were separated at birth and raised in differing environments. If environment was the sole cause of IQ differences, they should grow up with IQs no more alike than randomly selected members of the general population. If genes were the sole cause, they should grow up with identical IQs. The twin studies and other kinship studies showed that genes were far more powerful than environment.

Jensen captured the results in an equation: h^2 = genes at 0.80 + between-families environment at 0.12 + micro-environment at 0.08. The symbol h^2 stands for the extent to which IQ is heritable, meaning how it breaks down in terms of genes and environment. As you can see, genes are very potent and environment very weak. He emphasized that this equation applied to individual IQ differences *within* the same group. Environment's role in explaining *between*-group differences in average IQ (between black and white or between one generation and the next) was weaker still.

Between-families environment refers to all environmental differences that distinguish one family from another, for example, differences in SES. SES stands for socio-economic status, whether you belong to the upper, or middle, or lower class, which affects everything from the quality of the mother's womb when the baby is gestating, to poverty, to

E-mail address: jim.flynn@otago.ac.nz.

https://doi.org/10.1016/j.intell.2018.06.007
Received 20 April 2018; Received in revised form 27 June 2018; Accepted 27 June 2018
Available online 06 August 2018

nutrition, to neighborhood, to how many years of schooling you get, to what job awaits you. Micro-environment refers to the luck of life-history: whether you happen to be the first or last child (the first is thought to get more attention), whether you are dropped on your head and suffer brain damage, whether you are crippled by an auto accident, and so forth. Jensen argues that it would be things like SES that would provide the real explanation of between-group IQ differences. After all, both and black and white are alike (or at least very much alike) in that both have children born in a sequence (rather than all blacks being born at the same time as twins) and suffer tragedies that inflict permanent cognitive damage. Therefore, Jensen felt justified in taking the value for between-family environment (0.12) and using it to test the plausibility of environmental explanations of the black-white IQ gap.

He pointed out that explaining 12% of IQ variation between individuals was equivalent to a correlation between environment and IQ of about 0.33. Take this on faith as dictated by the mathematics of a normal curve of IQ scores. The square root of the percentage explained equals the correlation; and the square root of 0.12 = 0.34. If that were so, the ratio between the environmental gap between two groups and the magnitude of IQ explained would be about 3 to 1. Jensen put the IQ gap between black and white at 15 points or one standard deviation. Therefore, to explain the entire gap, the average black's environment would have to be 3 standard deviations below the average white's environment. If you look at a table of values under a normal cure, it will tell you what that means: the average black environment would have to be so bad that only 0.2% of whites fell below it. Or conversely the average white environment would have to be so good that only 0.2% of blacks fell above it.

As Jensen said, how can anyone argue this? Are only the small percent of whites that sleep on the streets on skid row below the average black environment? Are only the few black professors in universities above the average white environment? It looks as if the relevant environment is so weak between groups that no plausible environmental explanation is possible. This assumes that the environmental factors between groups are much the same as those within groups, but they appear to be: just as blacks differ from whites by SES and years of schooling, so blacks differ from one another by SES and years of schooling and whites differ from one another by SES and years of schooling.

And here is his clinching argument: if these differences much affect the IQ gap between black and white, they should also make a big individual difference within blacks and within whites. For example, if years of schooling make a huge IQ difference between groups, they ought to make a big IQ difference within groups. But this flies in the face of the twin studies: these all show that things like schooling are extremely weak within-group factors. Thus an environmental explanation of between-group differences in average IQ simply cannot get off the ground. Jensen's equation settles the question. This analysis froze our minds about between-group differences for almost 30 years, between 1970 and 2000.

There were refinements. It turned out that it was a mistake to lump all twin studies together. When you plotted them by the age at which the twins were tested, there was a progression from environment strong among preschoolers to very weak by adulthood. By 1980, Jensen (1980) had altered his equation. At the age of four, h^2 equaled 0.35 for genes, 0.40 for between–family environment, and 0.25 for micro-environment. By age 20, the equation read 0.75 for genes, 0.00 for within-family environment, and still 0.25 for micro-environment (Jensen, 1980). It was also acknowledged that between-family environment and “luck of life history” are not water-tight compartments. If blacks and whites differ as to family size, more blacks than whites will be born as “neglected” later children. If children in poverty homes are more likely to be dropped and violence more likely as you go down the SES scale, blacks are more likely to suffer brain damage.

But none of this really threatens Jensen's argument. Even if we split the total remaining environment at adulthood in half, allow that half of the 0.25 is relevant to group differences, only 0.125 of variance emerges – virtually identical to the 0.12 he originally assumed. Therefore, Jensen's later books (1980 and 1998) retained the argument. Others remained under its spell, witness Herrnstein and Murray (1994) in their analysis of group differences in *The bell curve*.

1.2. First impressions (1977)

When I first read Jensen in 1977, I was naïve being trained as a moral philosopher. I had already read an article by Jack Tizard (1975). It demonstrated one exception to Jensen's “clinching argument”: his case that his equation could bridge the gap that separated within-groups and between-groups. Height differences within a generation were mainly determined by genetic differences from one individual to another and differences in nutrition played a very minor role. And yet, between generations (between 1909 and 1952), London schoolchildren had made huge height gains: enhanced nutrition over that time was the only plausible explanation.

However, I wanted to apply a method I have always used when confronted with a social science model: how does it tally with my experience of my own life (sport plays a role here) and my awareness of history. I decided to take an environmental factor, years of schooling, and assess what role it had played in the lives of my father (last generation) and in the lives of my brother and I (same generation).

My father was born in 1885 and like most Irish-American boys at that time, left school after 8 years and went into factory work. The average of the birth dates of my brother and I was about 1929 and we both completed 21 years of education (got a Ph.D.). I imagined that we all had identical genes for cognition. In fact, I shared only half my genes with both but the assumption of equal quality is not so far fetched. Although uneducated, my father did the *New York Times* crossword puzzle in ink to show off the fact that he made no mistakes. My brother won awards as a Chemist that are at least as prestigious as those I have won as a “psychologist” or philosopher.

The notion that genes had rendered schooling into a weak factor *both* within and between generations seemed absurd. Within a generation, the “identical” genes of my brother and I might roughly predict our years of schooling (21 each) and perhaps near identical IQ scores. So as for *predictability*, our genes accounted for much and an environmental factor counted for little – in the sense that schooling added near zero predictability to genes. But did zero predictability entail zero *causal potency* – as if our 21 years at school had done nothing to add to our adult cognitive performance? And what of my father? In his day, perhaps his genes also helped to predict his 8 years of schooling (the median was only 6 years then) and his eventual IQ. But certainly, his eight years in a southern Missouri primary school had a causal impact on him that differed from what Catholic University and the University of Chicago did to us.

I should add that Jensen never claimed that an individual's years of schooling was dictated solely by the quality of his or her genes. Keep in mind his equation includes the effects of micro-environment or the luck of life history. If my brother or I had been blinded, this might well have caused a difference in our years of schooling.

Nonetheless, to equate within- and between-generation schooling differences was bizarre. If someone had only 8 years of schooling within my age cohort, we would suspect that he had genes that signaled mental retardation. To interpret the schooling gap between my father and I as showing that his genes made him suspect of mental retardation was another matter. If Jensen's analysis equated the two, it led us astray.

The focus on schoolings was prescient in that I later read Luria. He taught me how potent schooling was in altering our minds so that we would do well on IQ tests. He had interviewed isolated rural people in Russia in the 1920s. These were people who were like Americans in 1900 with little formal education.

First interview: Fish and crows (Luria, 1976, p. 82).

Q: What do a fish and a crow have in common?

A: A fish lives in water. A crow flies. If the fish just lies on top of the water, the crow could peck at it. A crow can eat a fish but a fish can't eat a crow.

Q: Could you use one word for them both?

A: If you call them both "animals", that wouldn't be right. A fish isn't an animal and a crow isn't either. A crow can eat a fish but a fish can't eat a bird. A person can eat a fish but not a crow.

Note how differently these people classify the world than we do. They exploit the world to their advantage and therefore, focus on what differentiates things: the most important thing for them is how different fish and crows are and they are reluctant to lump them together. We have become used to the categories modern science gives us to understand things. We have developed what I call new "habits of mind". We are ready to ignore differences and consequently lump fish and crows together as animals, dogs and ourselves as mammals, monkeys and ourselves as primates. When asked what dogs and rabbits have in common we say they are both mammals. They tend to say that you use dogs to hunt rabbits. None of our modern abstract concepts can actually be perceived in the concrete world. We have a whole new pair of spectacles that they lack. The Wechsler IQ subtest called Similarities is all about classification.

Second interview: Camels and German cities (Luria, 1976, p. 122).

Q: There are no camels in Germany; the city of B is in Germany; are there camels there or not?

A: I don't know, I have never seen German villages. If B is a large city, there should be camels there.

Q: But what if there aren't any in all of Germany?

A: If B is a village, there is probably no room for camels.

Note how reluctant people once were to use logic to reason about hypothetical situations (what *if* Germany was a country without camels?). Even when the logical conclusion is suggested, the man tries to turn it into something that describes a concrete situation (perhaps the village is too small for camels). In everyday life, whether you have camels is the important question; not using logic on words or symbols that posit a possible situation that has no reference to anything you have encountered in the real world. We have the kind of "habits of mind" that takes the hypothetical seriously no matter how far it is from anything we "know" and we have had plenty of practice in using logic to determine its consequences. Every bit of science or social science you learn at school puts forward hypotheses to be tested in terms of logic and novel experiences. The Raven's IQ test is entirely about using logic to order symbols that are "valued" for their own sake and have no concrete reference. Formal schooling prepares our minds to do IQ tests successfully, while in 1900 people found their contents and the problems they posed alien.

Luria also held the key to the fact that the very content of what is demanded in school has changed. Genovese (2002) compared the exams the state of Ohio gave to 14-year-old schoolchildren between 1902 and 1913 with those they gave between 1997 and 1999. The former tested for in-depth knowledge of culturally valued information; the latter expected only superficial knowledge of such information and tested for understanding complex relationships between concepts. The former were likely to ask you to name the capitals of the (then) 46 to 48 states. The later exams tended to ask you why the largest city of a state was rarely the state capital (rural members dominated state legislatures, hated the big city, and bestowed the capital on a rural town). Genovese (2002, p. 101) concludes: "These findings suggest that there have been substantial changes in the cognitive skills valued by Ohio educators over the course of the 20th century."

1.3. Tasks (1980)

First impressions are all very well. But a real answer to Jensen had to be fortified by scholarship. I jotted these tasks down in my notebook. (1) How large was the IQ gap between the generations? Was it equivalent to the black-white IQ gap and was it caused by environment? (2) Was there some mechanism that operated between generations that was absent within generations? (3) If you confuse the dynamics between genes and environment that distinguish IQ differences within- and between- generations, what price do you pay? (4) Even if environment causes huge IQ differences between generations, might not the kind of cognitive progress it engenders be quite different from enhanced intelligence – different in a way that qualifies the potency of environment? (5) Can we explain the mystery of environment losing and gaining potency from one context to another? (6) Even if that can be solved, does not the fact that genes predict our cognitive performance within a generation lead to a sort of genetic determinism that abolishes the autonomy of the individual?

1.4. IQ gains over time (1984–1987)

A few years later, I documented what became called the "Flynn effect". The 20th century had been dominated by massive IQ gains from one generation to another. Americans had gained 14 IQ points on the standard IQ tests (Stanford-Binet, Wechsler) between 1932 and 1976 (Flynn, 1984); and 14 nations had made massive gains on a whole range of IQ tests, the largest on Raven's Progressive Matrices (Flynn, 1987). Jensen (1980) had called Raven's the purest measure of intelligence. This phenomenon now covers at least 34 nations (Flynn, 2012) and is accepted by all scholars. The 21st century may well be different, with gains tailing off or reversing in some nations beginning in 1995, although not in the US (Flynn and Shayer, 2017).

On the face of it, these gains showed that extending Jensen's equation from within- to between-groups was not valid. Recall that he argued that to explain a one SD average IQ difference between groups, you had to posit a 3 SD environmental gap between them. Dutch males gained 20 points on Raven's between 1952 and 1982 (Flynn, 1987). This is equivalent to 1.33 SDs of IQ, which implied that from one generation to another Dutch social progress amounted to 4 SDs of environmental quality ($1.33 \times 3 = 4$). However implausible it seemed to posit a 3 SD environmental difference between black and white, a 4 SD gap between one generation and the next seemed beyond belief.

If you look at a table of values under a normal cure, it will tell you what that would mean. The 1952 Dutch environment would have to be so bad that only 3 ten-thousandths of 1% of 1982 Dutch fell below it. Or conversely, the 1982 Dutch environment would have to be so good that only 3 ten-thousandths of 1% of 1952 Dutch lay above it. I must warn you that Jensen (1998) had a reply that we will soon confront. He contends that although the IQ gap between black and white at a given time signals a true intelligence difference, only a fragment of the IQ gain between generations was an intelligence gain: one small enough that moderate environmental progress could explain it.

That aside, Jensen (1998) made a legitimate point: he despaired when those who knew about the "Flynn effect" used it as a "mantra" to equate black and white for genes for intelligence. Even if the gap between the generations is an intelligence gap, and even if it larger than the intelligence gap between the races, providing an environmental explanation for the former cannot substitute for providing an environmental explanation for the latter. It is quite possible that the environmental factors that separate the generations are quite unlike those that separate the races. However, it did make the environmental case more plausible. Thanks to IQ gains over time, black Americans in 2002 averaged at 104.31 when normed on the whites of 1947–48 (Flynn, 2008, Box 12). Given that whites gain at 0.3 IQ points per year (Flynn, 2012), this implies that the blacks of 2002 matched the whites of 1962 (0.3×14.37 years = 4.31; 1947.5 + 14.37 = 1961.87). That the black American environment lagged 40 years behind the average white environment was far more plausible than the light-year environmental gap Jensen's arithmetic posited.

1.5. The Dickens/Flynn model (2000)

Dickens sent Flynn a model that solved the paradox of environment feeble within a generation and environment huge between generations. This model was built on earlier contributions (Scarr & McCartney, 1983) and has been evidenced by recent research (Briley & Tucker-Drob, 2013).

Flynn noted a flaw that led Dickens to elaborate the model. In addition, I wanted to apply my usual method: could we find something in everyday life that exemplified how it worked. I suggested the immense improvement in the world record for the mile between generations (clearly environmental) contrasted with the fact that genes were highly significant in who could run the best mile within a generation. He suggested the much better example of basketball. The Dickens-Flynn model (Dickens & Flynn, 2001a, 2001b) introduced the concept of the *individual multiplier* to explain the potency of genes within an age cohort and the *social multiplier* to explain the potency of environment between generations.

John and Joe are identical twins separated at birth. Their identical genes make them both taller and quicker than average to the same degree. John goes to school in one city, plays basketball a bit better on the playground, enjoys it more, practices more than most, catches the eye of the grade-school coach, plays on a team, goes on to play in high school where he gets really professional coaching. Joe goes to school in a city a hundred miles away. However, precisely because his genes are identical to John's, precisely because he is taller and quicker than average to exactly the same degree, he is likely to have a very similar life history. A powerful feedback loop is operating here: good genes put you in a better environment, that better environment further boosts your basketball performance, that puts you into an even better environment, and so forth. A gene-caused advantage that may be very small at birth is multiplied into a huge basketball advantage by adulthood. This we call the individual multiplier. It is gene driven and it is operating within an age cohort.

Now imagine that John is born in 1945 and his “identical twin” is born in 1955. The invention of TV gave basketball a mass audience and increased the pay a professional player could expect. Basketball also had the advantage that ghetto blacks without access to playing fields could play it on a small concrete court. Wider and keener participation raised the general skill level: you had to shoot more and more accurately to excel. That higher average performance fed back into play: those who learned to shoot with either hand became the best – and then they became the norm – which meant you had to be able to pass with either hand to excel – and then that became the norm – and so forth. Every escalation of the average population performance raised individual performance, which escalated the average performance further, and you get a huge escalation of basketball skills in a single generation.

A new powerful feedback loop is operating: an upgraded environment upgrades play, and the rising average quality of play forces every individual to improve to match it, and that establishes an even higher average quality of play, and that puts you in an even better environment, and that environment further boosts your basketball performance, creating a better still environment, and so forth. This we call the social multiplier. It is environmental change driven and it is operating between cohorts. I experienced this personally. After a few years of inactivity, my Catholic Youth Organization (CYO) basketball team came back to play the current team. They killed us. They had all sorts of skills we lacked, they could shoot with either hand, could pass with either hand, do fade-away jump shots. I doubt that any of them had superior genes for basketball. Rather it was the passage of time that had given them a basketball environment a world away from our own.

I take it that it is easy to apply this to the realm of cognition. Within a cohort, genetic quality tends to dictate how you respond at school, how hard you work, whether you join a book club, how you will do in high school, what university you attend – your genetic quality will eventually tend toward a matching quality of environment for cognition. Genes predict both environment and IQ and environment alone predicts very little. Between cohorts spaced over time, different forces operate.

Since the industrial revolution began, social change has caused new cognitive exercise (recall Luria), more schooling, more cognitively demanding work, and more cognitively demanding leisure. These environmental factors initially triggered a mild rise in average performance, but this rise was greatly magnified by feedback mechanisms and over a century average IQ escalated.

As the average years of schooling rose, the rising mean itself became a powerful engine in its own right as people chased it to keep up. The reasons why people did so are as varied as the people themselves. For some, it is mere imitation: parents who see other parents keeping their children in school longer tend to keep their own children in school longer. Whatever the mechanics of the behavior, there has been an education explosion, from everyone having an average of 6 years of schooling, to most people having 8 years of education, to most people having a high school education, to more than half of people having some tertiary experience. The children need a better and better education to keep matching the rising average—particularly if they are to reap the benefit of thousands of new cognitively demanding jobs, which are better paid. As formal schooling frees people's minds from the concrete world to use logic on abstractions, performance on Raven's Progressive Matrices soars (Flynn, 2009).

Now we can see just why Jensen's equation could not bridge the gap that separates within- from between-generation. Two different multipliers operate. His equation catches the context in which environment is truly feeble as a predictor, a context in which similar genes mean similar IQs - thanks to the individual multiplier. To use it in a context in which environmental change is overwhelmingly dominant in raising the average IQ, thanks to the social multiplier, is simply to use the wrong arithmetic. The context makes all the difference. For example, the speed at which one runs in preparing to throw a javelin could be caught in an equation in which speed plays a minor role compared with factors such as technique and strength. But it would be absurd to use that same equation as indicative of the role of speed versus technique (not over-striding) and strength (upper body development) when modeling something else, perhaps running the hundred-yard dash. A new equation would be needed that fit the new circumstances in which speed would have a far greater role. I will soon suggest that when we confront between-group differences in average IQ, we should use something called sociological arithmetic.

1.6. Thoughts on Lewontin and arithmetic (circa 2000)

The penalty for using the wrong arithmetic is best illustrated by the Lewontin debate. In 1966, Richard Lewontin coauthored articles with J. L. Hubby that helped set the stage for the field of molecular evolution. In 1970, on the grounds that Jensen was a genetic determinist, Lewontin called his doctrine “Jensenism,” likening it to a heresy called “Jansenism,” which Pope Innocent X had condemned as a doctrine of “total depravity, predestination, and limited atonement” (as quoted in Lewontin, 1970, p. 2). Jensen quite rightly resented the use of rhetoric to load the dice on an issue that should be debated purely in terms of logic and evidence.

Between and within-group differences: three alternatives

	Within-group 1	Between-groups	Within-group 2	Rationale
Jensen	Family – School Work - Leisure	Family – School Work - Leisure	Family – School Work - Leisure	Between-group factors can be no more potent than they are within groups
Lewontin	No X = absence of zinc for none	**X = presence or absence of zinc**	X = absence of zinc for all	X has no potency within either group but has total potency between groups
Flynn	Family – School Work - Leisure	**Family –School Work - Leisure**	Family – School Work - Leisure	Between-group factors can be far more potent than they are within groups

Fig. 1. Between and within-group differences: three alternatives.

1.6.1. *Two bags of seed*

The substance of the debate was whether or not Jensen had made a case that a large part of the IQ gap between blacks and whites was genetic. As a geneticist, an expert in disentangling genes and environment, and as a professor at Harvard, Lewontin's views carried great weight. Lewontin (1970, 1976) rejected the corollary to Jensen's equation: that whatever environmental factor was weak within two groups must also be weak as a causal factor explaining an average difference between the groups. However, he evaded Jensen only by positing a spurious difference between groups. He posited a Factor X – and Factor X did provide an exception to Jensen's rule. But he left Jensen's rule intact for *every variable other than* Factor X. Lewontin never saw that whereas twin-study arithmetic had a role within groups, it was sociological arithmetic that was appropriate when we posit factors between-groups.

His scenario: Take two populations within which genetic differences determine all individual differences for a trait. That means that genes would determine the total within-group variance for the trait. Still, if there were a deleterious environmental factor that afflicted every member of the first group equally but was totally absent within the second group, then the average trait difference between the two groups would be entirely environmental.

Lewontin posited a quantity of seeds that differed from one another in terms of genes that promote height. Distribute the seeds randomly into two bags, and each bag will have genes of the same average quality. Plant one batch in a potting mix that is both optimal and equal for every plant. That means that genes will determine all height differences *within* that population. After all, how can environment affect height if it is the same for all seeds? Plant the other batch in another potting mix that is optimal, except that it lacks zinc (good for plants). With identical environments, genes will determine all height differences within that population as well. However, the average difference in height between the two populations would be entirely due to environment (the presence or absence of zinc).

Such an environmental difference between two groups came to be called Factor X (something entirely missing in one and omnipresent in the other). Needless to say, in the real world, no one could find a Factor X. Any factor that varied between the black and white groups also varied within black and white groups (e.g., family quality, schooling quality, jobs, leisure quality – even the impact of discrimination). Thus, despite trying to answer Jensen's case that the IQ gap between races could not be explained environmentally, Lewontin made Jensen's triumph inevitable.

All Jensen had to do was revise his dictum as follows: A within-group factor can differentiate between two groups for a trait but only under totally unrealistic conditions; if it does *anything* to differentiate individuals within either group (something true of all real-world factors), then it can have no between-group potency. If Lewontin wanted to really counter Jensen, he would have to do what we have done: he would have to take the real-world environmental variables that operate within-groups and show that *they*, despite their weakness there, could have great potency in explaining average differences between groups.

1.6.2. *Alternatives to factor X*

Look at Fig. 1. It presents Jensen's approach to within- and between-group IQ differences, Lewontin's model, and my approach. Jensen, of course, has the usual sociological factors of family, school, work, and leisure very weak within groups and sticks to his corollary that they must also be very weak between groups (as if the difference between my father's and my education meant little). Lewontin goes him one better by accepting that the usual real-world factors may have no potency at all within groups (and therefore none between groups). This leaves him with no recourse save a Factor X as the total explanation of IQ differences between-groups. Thanks to the Dickens-Flynn model, I ignore Jensen's dictum entirely; and posit that the usual environmental factors are more potent between- than within-groups; and I make them large enough so that they may well explain a huge gap in average IQ.

2. Sociological arithmetic

Without breaking the laws of arithmetic, we are now liberated to use the usual sociological factors to offer an environmental explanation of between-group IQ gaps. I have used them to argue that black Americans are less adapted to the dominant culture than white Americans (Flynn, 2008, Chapters 2–4). I do not believe that my case should end the debate. We need a far better analysis of the nature of black American subculture. All I am arguing here is that a sociological thesis should be examined on its merits. Models that prove mathematically that no such case is possible should be set aside. As Kraus, Rucker, and Richeson (2017) show, there is a massive difference between the true disadvantage blacks suffer in terms of socio-economic variables and the disadvantage perceived by the American public.

Let us return to the social factors that separate generations. It is easy to see the kind of factors that would be important but can we measure them? As for the importance of parent-child interaction in the family, the literature is huge. As for formal schooling, Luria tells us that it accustoms you to classify (as in the similarities subtest) and use logic on abstractions (as in Raven's Progressive Matrices). As for jobs, we know that in nations in which working life is prolonged, there is less cognitive decay (Adam, Bonsang, Germain, & Perelman, 2007). As for leisure, keeping fit is good not just for your body but also for your brain (Altenburg, Chinapaw, & Singh, 2016).

Recently, we have got the first evidence that tells us how to measure the relative potency of social variables, at least in a particular nation over a certain period. Staff, Hogan, and Whalley (2014) show how much the IQ of one cohort differs from the IQ of another cohort *according to age* – both cohorts having been drawn from the same population. The two Scottish Lothian Birth Cohorts were born in 1921 and 1936, respectively. Both were tested on Raven's Progressive Matrices. The later cohort outscored the earlier by 3.7 IQ points at age 11 (a rate of 0.247 points per year); and by 16.5 IQ points at the age of 77 (a rate of 1.100 points per year). Data from 11 international samples suggest that the rate for 11 year-olds doubles by the ages of 17.5 to 24 (Flynn, 2012, Box 11). This suggests a missing value for those ages of about 0.500 per year for Scotland.

As for the various factors that cause IQ gains, the data above imply the following weighting:

1. Better home environment (and perhaps better quality of school) caused gains at a rate of about 0.25 points per year. Note that the *years* of school would be the same for both cohorts at age 11, which would act as a leveler.
2. More years of school in the second cohort was worth 0.50 points per year. By young adulthood (ages 17.5–24), many more had finished high school and had gone on to university.
3. More cognitively demanding jobs benefitted mature adults (ages 30–60), but all we can say here is that fragmentary data suggest that this would not be less than 0.50 points per year.
4. Vastly enhanced conditions of old age in the second cohort (exercise, diet, medical care, stimulation) were worth 1.10 points per year.

Following the Dickens-Flynn model, this summary assumes that the current environment eventually obliterates the effects of past environments, and thus the gains are not additive but rather replace one another. Measurements of this sort will vary from nation to nation and time to time and be rough. They lack the mathematical elegance of Jensen's equation but they apply to the real world. I accept Aristotle's advice: it is a mistake to demand more exactness than the subject matter admits.

2.1. *g* and the pursuit of true intelligence (2007–2012)

Jensen (1998) accepted that there have been massive IQ gains between the generations, and he accepted that these were mainly caused by environment. But what if these were not true intelligence gains but rather, gains on particular skills, gains caused by things like enhanced education, gains that were logically distinct from intelligence? He did not deny that a small fraction of IQ gains might be real, say those caused by an environmental factor that directly affected the brain (greatly improved nutrition), but these were not large enough to suggest that environment had much potency between groups (between generations). This left intact his corollary that environmental potency can be no stronger between groups than it is within groups. Or at least, the factors of family, school, work, and leisure were no more powerful "intelligence-wise" between generations than they were within a generation.

Jensen (1998, p. 320–333) calls the huge socially caused IQ gains "hollow". He also suggested a method of testing whether IQ gains were true intelligence gains, which he called the "method of correlated vectors." It determines whether IQ gains over time are "*g*" gains, which means we must introduce the concept of general intelligence or *g*.

2.1.1. The method of correlated vectors

The concept of *g* is similar to full-scale IQ but subtly different. When we compare the performances of individuals on a variety of cognitive subtests, say the 10 or 11 subtests that make up the Wechsler, there is a "positive manifold." This means that those who do better than average on a certain subtest tend to do better than average on them all. Factor analysis is a technique that measures this tendency and its result is called *g*. This immediately gives rise to calculating the degree to which better performance on a particular subtest predicts better overall performance, which is called the subtest's *g* loading.

A pattern emerges: The more complex the items of the subtest, the more predictive it is. For example, the digit-span-forward task involves simply repeating the sequence of random numbers read out to the testaker, and has a low *g* loading (low predictability of total performance). The digit-span-backwards task asks the test-taker to repeat the numbers in the reverse order from that in which they were read out, a clearly more complex task, and it has a higher *g* loading or higher predictability (Jensen, 1980).

When Jensen uses the method of correlated vectors, he is comparing the size of IQ gains with the hierarchy of *g* gains. First, one ranks the ten Wechsler subtests in order of the size of their *g* loadings; then one ranks two groups in terms of the size of the score advantage one group has over the other subtest by subtest; and finally, one determines whether the two tally. In the case of IQ gains, the two groups are two generations separated by time. Assume a later group outscores an earlier group more on the digit-span-backwards task (high *g*) than on the digit-span-forward task (low *g*). Then you have the beginnings of a *g pattern*. This sends the message that the more complex the task, the more the later group increases its advantage over the early one.

But what if there is no tendency for the score differences between the two groups to tally with *g* loadings (e.g., all subtest score differences are much the same). There may even be an "anti-*g* pattern" (e. g., the lower their *g* loadings, the larger the advantage that the later group has on subtests). In such cases, there is no indication of a positive manifold, but rather a neutral or negative one. Then you can say the score differences are non-*g* differences.

However large they may be, IQ gains over time flunk the test set by the method of correlated vectors. Usually, the magnitudes of the subtest gains one generation enjoys over the preceding generation have no correlation with the *g* loadings of the subtests (te Nijenhuis & van der Flier, 2013). That is why Jensen calls them "hollow". Helter-skelter gains on the subtests are simply a matter of mastering specific cognitive tasks and do not show that true general intelligence is rising. We can see his point: common sense seems to dictate that an intelligent person opens up a greater gap on the average person, the more complex the task.

2.1.2. Why IQ gains are not *g* gains

Social change causes fluctuations in the average IQ as generations succeed one another. Society causes the development or atrophy of cognitive skills in terms of its own priorities. It is not like someone aspiring to be a chess grandmaster that wants to best others on the most complex skill possible. Whether the magnitude of subtest gains tally with the subtest's *g* loading (complexity) simply does not count.

Maguire, Woolett, and Spiers (2006) did a wonderful study comparing London taxi drivers (who must find their way around a huge city) with London bus drivers (who follow fixed routes). The minds of the taxi drivers did far more mental exercise in terms of practicing map-reading skills. Years on the job correlated with an enlarged hippocampus only in taxi drivers, and the enlargement was greater the longer they had served. The hippocampus is the area of the brain that is most active when you do map reading and it reacted much like a muscle. When you do weightlifting your biceps enlarge, when you do a lot of map-reading your hippocampus gets bigger. The inference is that even the brain fluctuates not only in accord with how nutrition alters over time but also with how work and leisure alter over time.

Throughout the 20th century, as driving cars became more and more of a leisure-time activity, there was more cognitive exercise of the map-reading sort. We cannot prove it by evidence as yet, but the hippocampus of the average person grew, and now it should be reducing thanks to the introduction of automatic guidance systems in cars. I predict that the next study of London taxi drivers will show a reversal of hippocampus enlargement.

I suspect that map reading has a relatively low *g* loading (there seem to be no studies). Over the 20th century in America, the IQ subtest that shows the highest gains (for adults) is vocabulary, which has a very high *g* loading (Flynn, 2012). This was because as people filled more cognitively demanding jobs, they needed more education and did more on-the-job verbalizing. The relative gains on these two skills had everything to do with changing social priorities, and nothing to do with *g*. If society has a greater need for more tradesmen than for people with analytic ability, if it wants more carpenters and fewer physicists, it will get them. The social utility of skills trumps the fact that one is more or less complex than the other. The proof that society operates in this way is precisely this: that the skills it enhances over time are *not* correlated with their *g* loadings.

2.1.3. Practical significance

IQ gains over time have great practical significance no matter whether the helter-skelter pattern of skill trends passes the test of correlated vectors or not. Their trends match what new cognitive priorities the industrial revolution demanded over time (more schooling, more hot-housing of preschoolers so they will do well in school, more demanding jobs). These skill changes play the role of cause as well as effect. Society cannot advance unless its people adjust. Evidence is now accumulating concerning the causal potency of non-*g* IQ gains.

Coyle and Pillow (2008) show that the cognitive skills measured by the Scholastic Aptitude Test predict university grades even after *g* has been removed. Ritchie, Bates, and Deary (2015) show that the association of education with improved cognitive performance is not mediated by *g*; rather education directly affects specific IQ subtests. Meisenberg (2014) argues that we are accumulating over time “cognitive human capital” that is interdependent with economic growth.

Woodley (2012b) shows that the historical trend of IQ gains (which again are *not* correlated with *g*) both parallels and predicts the growth in gross domestic product (GDP) per capita in Western nations over the past 10 decades or so (correlation = 0.930). Indeed, Woodley et al. (2017) and Woodley, Sarraf, Peñaherrera-Aguirre, Fernandes, & Becker (2018) have recently published perhaps the strongest evidence that declines in *g* are accompanied by enhanced specialized abilities that have real world significance. These build on his earlier thesis that a slower “Life History Speed” explains the “Flynn Effect” (Woodley, 2012a).

Nisbett (2015) cites the example of Ireland. Between 1970 and 1985, the Irish went from educational parity with Great Britain to well above. They boosted years of education by 2.3 years whereas Britain rose by only 0.77 years. Ireland multiplied those with tertiary education by six times while Britain did so by only three times. In 1970, per capita GDP was virtually the same for both. In 2005, it was $65,292 (U.S. dollars) in Ireland and $38,149 in the United Kingdom.

The Irish were supposed to be constrained by an IQ as low as that of American Blacks (Eysenck, 1971). Lynn believed much of their low performance was genetic, stemming from selective migration of brighter Irish going abroad. However, when you select out the studies from Lynn (2006) and Lynn and Vanhanen (2002) that have large samples, and supplement them with the 2009 Irish performance on PISA (a comprehensive Program for International Assessment by the OECD), a pattern emerges: 1971, IQ 87; circa 1991 (4 studies), IQ 94; 2000, IQ 95; 2009, IQ 100. All Irish IQ estimates are based on British norms set at 100 at the time. This appears to be a classic case of society demanding more intellectual exercise followed by rising IQ interdependent with economic growth.

IQ gains have upgraded our reasoning in another area. Analytic gains do more than correlate with educational and scientific progress. They promote moral progress, for example, they accustom us to arguments that put racism and sexism on the defensive (Flynn, 2013). The key factors are taking the hypothetical seriously and using logic to universalize moral principles.

In 1955, when Martin Luther King, Jr., began the Montgomery bus boycott, young men and women, home from college, had dialogues with their parents. Mine went as follows. Question: “What if you woke up tomorrow and had turned Black?” Reply: “That is the dumbest thing you have ever said. Who do you know that turned Black overnight?” My father was simply not willing to take the hypothetical seriously. Fathers in less developed areas of the Middle East sometimes follow a code that dictates killing a daughter because she has been raped – for the sake of “family honor”. One might ask, “What if you had been knocked unconscious and sodomized?”. However, if the father sees moral maxims as concrete things, impervious to change, rather than as general principles subject to logic, he will dismiss the question as irrelevant (Flynn, 2013).

2.1.4. What about the brain?

Throughout his career, Jensen (1972, 1973, 1980, 1998) believed that *g* had a special significance because it measured something in the brain that played a key role in intelligent behavior. He spent much of his life (until the day he died) studying reaction times, how quickly people could respond to a visual or auditory stimulus as an indication of the efficiency of their brains in terms of “neural speed”. He (rightly) thought it important how quick their brain was in terms of passing information from one neuron to another along the projections (axons) that connect them. He found that the speed of reaction times had a modest correlation with IQ as measured by *g*-loaded tests such as Raven's.

Let us assume that brain research eventually shows that Jensen was correct. We can actually measure how fast information travels between neurons (neural speed). We find that it travels faster for bright people than average people, that the more complex the task the greater the advantage conferred, and thus that neural speed does have a high correlation with *g*. What I dispute is that this lends *g* (or *g* gains over time) some special significance that overshadows gains on specific cognitive skills when they do not constitute *g* gains. That is, when they do not pass the test of correlated vectors, which generally they do not. What I dispute is that the latter are somehow “hollow” compared to the former (Jensen, 1998).

After all, *g* is not unique in having some sort of physiological substratum in the brain. As we have seen, this is true of all cognitive skills, for example the hippocampus has an important role in memory and map reading (Maguire et al., 2006). When society asks us to increase our use of any skill over time, the brain responds. This is likely true of every cognitive skill, not just those with a high *g*-loading like working memory but also those with a low *g* loading like rote memory. Every increased use rewires neural connections and hence, the speed of transmission: “Over the long term, repeated practice causes the brain to assign extra neurons to the task, much as a computer assigns more memory for a complex program. The assignment of these additional neurons is more or less on a permanent basis” (Sousa, 2011).

In sum, it makes no difference whether the skill is analysis (like Raven's Progressive Matrices), or acquiring the meaning of words, or classification, or comprehension of the world, or gaining information, or spatial visualization. *All* of the relevant neural connections are enhanced by use, which prompts this question: Why do extra neurons not count (i.e., become hollow) *unless* the skill with the greatest *g* loading enjoys more thickening than skills with lesser *g* loadings? In other words, both environmental factors (like more cognitive exercise) and genetic factors (like a better inheritance) rewire the brain. Even if it is true that the method of correlated vectors shows that helter-skelter gains on IQ tests rewire the brain differently than *g* gains, the gains in neural efficiency are real.

2.1.5. What about the individual?

Within an age cohort, a boy sits a Wechsler IQ test and shows a non-*g* pattern on its 10 subtests. He scores the equivalent of 15 IQ points above the average on each subtest. He gets 115 on Vocabulary (high *g* loading) just as he does on Digit Span Forward (low *g* loading). Therefore, he gets an IQ of 115 (actually it would be a bit higher), which puts him at the 84th percentile of his peer group. According to Jensen, we should take him aside and tell him that he has opened up no intelligence gap on the average person. He merely has a helter-skelter collection of specialized cognitive skills probably due to better education.

What a pity: his better education might have resulted in a genuine intelligence advantage by elevating his scores in the order of the cognitive complexity of the tasks. Nevertheless: it would be interesting to compare his college grades with those of boys who got a 115 IQ in the orthodox way. If he did as well, what exactly is the cash value of our deflating conversation with him?

Let those who believe in the peculiar causal potency of *g* show that IQ subtest differences are impotent when the *g* pattern is absent. Flynn, te Nijenhuis, and Metzen (2014) put the final nail in the coffin. We compared the Wechsler subtests scores of typical subjects with those who suffered from iodine deficiency, prenatal cocaine exposure, fetal alcohol syndrome, and traumatic brain injury. The typical subjects were higher on every subtest. However, the magnitude of their advantages by subtest had zero correlation with the size of the subtest *g* loadings. It is difficult to deny that the typical subjects had a significant “intelligence” advantage over the four comparison groups. In sum, helter-skelter or piecemeal advantages on subtests are causally potent, whether between generations or between individuals.

2.1.6. What about words?

You can try to use words to do what only evidence can do. Over the years, each issue of the journal *Intelligence* has shown a gradual tendency toward my account of what causes IQ gains over time, culminating in a special issue devoted to the Flynn effect (*Intelligence,* 2013, vol. 41, no. 6). However, five scholars still use language I reject. Dutton, Bakhiet, Ziada, Essa, and Blahmar (in press) say the following: “Pietschnig and Voracek (2015) show the most persuasive of the proposed explanations for the Flynn Effect, a slower Life History Speed (LHS) has led to greater specialization, increased education, and the increased use of analytical thinking.”. That is perfect. But they also say: “Dutton, Van der Linden, and Lynn (2016) have argued that this environmentally induced increase has occurred in tandem with evidence . . . that general intelligence has been falling.”

By general intelligence, they mean *g*. They are referring to evidence that since the late 19th century, there have been signs of things like dysgenic mating (Woodley, te Nienhuis, & Murphy, 2013). People with more education have been having fewer children than people with less education. Therefore, if there is any correlation between genes for intelligence and education (and of course there is), the average quality of genes for intelligence has been falling. This tendency toward reduced IQ has been swamped ever since it began by huge enviromentally induced gains. But what I wish to note is that the word “intelligence” has been reserved to describe the genetic decline, and that it has been withheld from the environmental increase.

No one would talk this way about any other trait. If there had been dysgnetic mating for speed (slower people having a few more children than fast people), and yet due to the jogging craze, the average person could run a faster mile, no one would say that “speed” had been falling. If there had been dysgenic mating for height, and yet thanks to greatly improved nutrition, people were on average taller than they used to be, no one would say that “height” had been falling. All these scholars mean is that the environmental trend toward higher IQ is accompanied by oppostite genetic trends. So why not say just that? An eccentric use of the word “intelligence” adds nothing to the information conveyed but it is confusing. This is the last stand of *g*: language that implies that *g* trends are true intelligence trends and IQ gains are not. Still I have no desire to censor the vocabulary of my colleagues as long as I know what they mean. All that is left of the Chesire Cat is the smile (but see Box 1).

I want to make it clear that although enriched environment dominated the 20th century, IQ gains are not destined to persist like the law of gravity. Factors that were immediate triggers of IQ gains included more adults per child in the home, more and better schooling, more people at university, more cognitively demanding jobs, and better health and conditions of the aged. There are signs that these are beginning to show diminishing returns.

We cannot lower the number of children in the home further unless we fail to reproduce ourselves. Already 52% of U.S. adults have some tertiary experience, the economy is producing at least as much undemanding service work as professional jobs, and so forth. Today, The Netherlands shows a mix: static environment for preschoolers (no IQ gains on the previous generation), high school in mild decline (a small IQ loss at these ages), a few more cognitively demanding jobs (a small IQ gain in adulthood), the aged still getting a much better environment (huge IQ gains on the previous generation). Judged according to samples of 18-year-olds, the Nordic nations entered a period of IQ decline about 1995 (Flynn & Shayer, 2017). How much of this decline is due to the gradual deterioration of our genes? That is at work, of course. But most of the decline may be due to a sudden downward turn in the quality of environment.

2.2. The environment never goes away (2016–2017)

The Dickens-Flynn model explains how environment can be potent between generations (and between other groups) even though it is weak

Box 1

When *g* is demystified, it can pose some interesting hypotheses. Traits with a higher *g* loading (those that are more complex) are more affected by inbreeding depression than simpler cognitive traits (Rushton, 1995). Breeding with close relatives (say first cousins) is harmful. If the genomes of the mating pair are too similar, and both have a recessive (potentially harmful) gene, then during sexual reproduction, the two are more likely to match and handicap the offspring. This poses the following hypothesis: IBD affects those areas/networks of the brain that are the seat of complex thinking more than it threatens those that are the seat of less complex thinking (e.g., the hippocampus as the substratum of map-reading).

This hypothesis can also be generalized: IBD threatens areas of the brain the more they are recent rather than earlier evolutionary acquisitions. Our brains were capable of rote memory earlier than they were capable of abstract analysis. Fernandes, Woodley, and te Nijenhuis (2014) argue that in terms of phylogeny, *g* is indeed a much more recent addition to the primate brain. An alternative hypothesis is that *g* affects the brain in general (shows a net loss of thickness of all neural pathways) and thus if IBD affects the whole brain, *g* is most affected.

Table 1
Comparison of genetic proportion of variance between Flynn's Stanford-Binet Vocabulary analysis and Dutch kinship estimates.

My ages	Ave. Cor. By age	% Var. family	% Var. uncommon	% Var. genes (mine)	% Var. genes (Dutch)	Dutch ages
4	0.806	64.96	18.00	17.04	22	5
6.75	0.562	31.57	18.00	50.43	40	7
9.25	0.426	18.13	18.00	63.87	54	10
11.5	0.290	8.42	18.00	73.58	85	12
14.5	0.267	7.12	18.00	74.88	–	–
18	0.121	1.46	18.00	80.54	82	18
20–24	0.121	1.46	18.00	80.54	–	–
25–29	0.073	0.53	18.00	81.47	88	26

Source: Flynn, J. R., 2016, page 63: with permission of the Syndics of The Cambridge University Press.

as a predictor of IQ within an age cohort. This contrast comes to life when we examine exactly what happens as children age. Here we must distinguish between environment losing *predictive* potency and environment losing *causal* potency. The reader unfamiliar with the twin studies may be skeptical even about predictive potency. She may doubt that by early adulthood, the quality of environment so exactly matches the quality of genes that genes alone are sufficient to predict both environment and IQ – so that environment alone has no predictive role.

I was somewhat skeptical until I devised my own method, the age-table method of measuring the march of genes (Flynn, 2016). I found that my method was so close to the twin studies as to validate them. Table 1 compares the results. The correspondence at the age of 18 convinced me. I found that the residual influence of between-family environment had dwindled to 1.45% and that genes accounted for 80.54%; the twin studies found that the latter were at 82%. For future reference, Table 1 also shows that at adulthood there is still a substantial proportion of variance due to "uncommon environment" (the good or bad luck of life history).

My method has not been greeted with enthusiasm (Matthews & Turkheimer, 2017). Naturally, I believe I have a crushing rebuttal. My method is supposed to be flawed yet its results tally with those of the twin studies. It is as if someone had used a "flawed" method that duplicated all of Newton's predictions. Surely we would have to concede that it would be useful whenever Newtonian data were missing.

However, if you wish, my method can be ignored. As Mathews & Turkheimer grant, the twin studies *alone* vouch for the key fact: as children age, genes become sufficient to predict both environment and IQ – and environment alone loses its predictive role. I am going to argue that as children age, what really happens is that genes and environment merely become better and better correlated; and that environment retains its full causal potency in the process.

2.3. Ages 2–3 and the fairness factor

Most parents try to be fair. They may be aware that two of their children seem to differ in genetic promise, but they try to give both the same quality of environment (reading the same number of books to them both, etc.). Note what this entails. If both children get the same environment, the less promising child of necessity gets an environment that enhances his or her genetic capacity more than it enhances that of the more gifted child (Flynn, 2016, pp. 73–78).

Imagine a chariot called cognitive performance (P) pulled by two horses, one called genes (G) and the other called environment (E). The first image in Fig. 2 shows the fairness factor at work for 2- to 3-year-olds. Genes and environment are tugging performance in "opposite" directions. Actually, at these ages, environment is more powerful and dictates a path that makes genes a bad predictor of performance (Bouchard, 2013).

2.3.1. Age 10 and the rising correlation

By age 10, school, teachers, and peers (the current environment) have begun to swamp the effects of family. Rather than a shared home environment being predictive of virtually all math performance as in preschoolers, the current environment (school, teachers, peers) has taken over. Its thrust is to encourage (not discourage) the correlation between quality of genes and quality of environment (Flynn, 2016, pp. 42–61). By age 10, there is no longer a negative correlation between genes and environment; indeed, the correlation is robust. But there are still residual family effects, so the horses are not pulling performance in quite the same direction. Nonetheless. genes have become the better predictor of the path of the sleigh.

By now the individual multiplier has kicked in. It is a process of reciprocal causality. Better genes cause better performance at mathematics; this performance causes a better environment (teacher attention, the math club); this causes even more enhanced performance; which causes a better environment still (a top stream). All three factors (G, E, P) are causally potent. It is just that as the correlation between

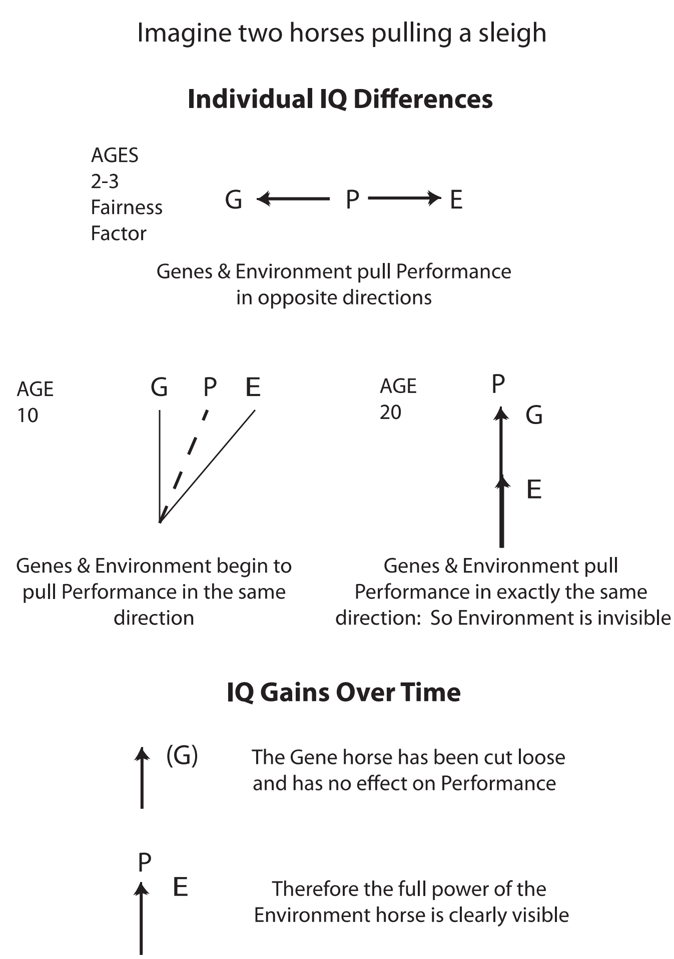

Fig. 2. Imagine two horses pulling a sleigh.

Box 2

At the bottom end of the spectrum, autonomy clearly counts. Some decades ago, *we decided* to help those who suffer from intellectual disability. We shifted them from an environment that merely matches their genes to a superior environment by giving them quality tutoring or mainstreaming. Flynn (2016, p. 15) shows that they attain a mean IQ 7 points (0.47 SD) higher than an environment matched to their genetic potential. They may therefore get a better job than they would otherwise, one that challenges them enough to keep their bonus. Our choice made a difference.

At the other end of the spectrum, we have no evidence but can posit an ideal case that warrants investigation. Assume that: (a) the genetic promise of a person who left school would put him at an IQ of 130 (top 2%) if he had a matching environment (also top 2%); and (b) thanks to bad luck or personal choice he is actually stuck in an average environment. Then his current environment is worth 100 or 30 IQ points below his "target" on the environmental scale. If the correlation between environment and IQ is 0.33, he has lost 10 IQ points (30 × 0.33). The correlation of 0.33 is dictated by 10% of IQ variance explained by autonomy (the square root of variance explained equals the correlation, and the square root of 0.10 is 0.33). Therefore, if going to a university affords him a commensurate environment, his IQ will go from 120 to 130. His voluntary action has allowed him to leapfrog from the 90th to the 98th percentile, or over four-fifths of the people who were above him. As for evidence, plenty of people with high IQs are in a bad environment. Board-certified neuropsychologist Dr. Robert Novelly tested a bouncer's IQ: "Chris is the highest individual that I have ever measured in 25 years of doing this" (Sager, 2017). Scholars should find some that have gone back to university and test them.

Note that even in the unlikely case that twin studies showed no variance to be explained by luck and choice, autonomy could still be operative. Thousands of people take a pill to alleviate chronic anxiety and thus attain a match between the quality of their genes and environment (a choice they would not make if they were Christian Scientists). In other words, autonomous choices can eliminate the evidence of their existence in twin studies. The fact that that there is a luck/choice component of cognitive variance is *proof* of the existence of these factors but its size is not an *index* of the magnitude of these factors.

the quality of genes and the quality of environment approach "perfection," the footprints of environmental causality are erased. This is because genes have undermined its role as a predictor: genes predict both quality of environment and IQ performance. It is if the gene horse has taken over the lead in pulling the sleigh and the environment horse docilely follows. However, as the individual multiplier rolls on throughout life the environment is causally active. It has to be or *reciprocal causality* would stop.

My son is now Professor of Pure Mathematics at Oxford. At the age of seven, he came to me with questions about infinity. He said, "There are an infinite number of numbers, but there are an infinite number of even numbers; so one kind of infinity has twice as many members as another.". I knew little math but was good at it. I pointed out that this means you could do arithmetic with kinds of infinity. If you subtracted the infinity of even numbers from the infinity of all numbers, you got the infinity of odd numbers. But I did not, at that point, disappear as an environmental factor, as if I had jumped out of a window. I (plus his teachers) recognized his genetic promise and began to make sure that the quality of his genes was matched with an environment of similar quality. We began to make sure that both his genes and his environment would pull his sleigh in the same direction.

2.3.2. Age 20 and the correlation made (almost) perfect

Family effects are most persistent for Vocabulary but even there, by age 20, the correlation between the excellence of current common environment and genes is almost perfect (Flynn, 2016, 64–67). Genetic differences between individuals from the same cohort have enslaved current environment to follow in their wake (except the micro-environment of course—see below). Therefore, the two horses are pulling cognitive performance in the same direction but the two are far more potent than one would be. Has environment lost its causal potency simply because the gene horse has stripped it of its predictive potency?

2.3.3. IQ gains over time

The answer lies in the phenomenon of huge IQ gains over time. From one generation to another, genetic trends are weak and environmental trends dominate. It is as if the gene horse has been cut loose from the sleigh; and environment is the only horse left; and it is causing IQ gains at the rate of 9 points per generation. The last image in Fig. 2 illustrates this. It too can be deceptive. When environmental trends swamp genetic trends, it looks as if the gene horse is entirely absent. It is really still there and if something like dysgenic mating is taking place, that trend will take over entirely once environmental enhancement ceases.

3. Genes and human autonomy (2017)

Returning to the relation between genes and environment within a cohort, we have ignored something important. At all ages, there is an unshared environment that explains about 20% of IQ variance (Flynn, 2016, p.22: Haworth et al., 2010). This means that it always remains uncorrelated with either genes, or the current environment, or the combination of the two. Identical twins can have bad luck even before birth. A majority of them share the same placenta and one can be better nourished than the other. No matter how excellent your genes or environment, you can be dropped on your head. At 18, you can be drafted out of an environment commensurate with your genes and submerged into the inferior environment of cutting throats in the forest. There is also good luck; for example, you happen to marry a partner who lifts you into a new peer group that gives you a current environment that is superior to the one commensurate with your genes. Therefore, this portion of IQ variance is often described as the result of luck.

But it is wrong to think of this 20% as mere luck. We do not know how to partition it, but at least half of it (10%) may reflect human autonomy. When an academic volunteers for the army, he has *chosen* an inferior environment. When someone who did poorly at school and is trapped in a humdrum environment decides to go back to university, he has *chosen* a superior environment. Many such decisions are, of course, influenced by genes for IQ and thus fall into the genetic portion of IQ variance. But the mere existence of uncommon environment variance signals that some of them are a matter of good luck or good choice. Box 2 spells out the extent to which an individual may benefit from his or her choice.

Despite the dominant role of genes in pulling the chariot, human autonomy is alive and well (Flynn, 2016, pp. 23–26). Genes set limits but within those limits, your fate is in your hands. When I say "20 percent of IQ variance" is uncorrelated with both genes and current environment, this must be qualified to refer to advanced nations. In many nations (e.g., Syria with its wars, Sudan with its famines), life is much less secure, and the frequency of chance events that determine one's fate is much higher.

Welcome home

I have acknowledged my debt to Jensen, Bill Dickens, and Dick Nisbett. I also profited from Thomas Bouchard (on twin studies), Steve Ceci (who showed that IQ tests often mis-measure skills people demonstrate in everyday life), Roberto Colom (on brain physiology), Howard Gardner (who reminded me that a humane society should value everyone for their unique skills), George Oesterdiekhoff (who bought Luria to my attention), Sandra scar (who anticipated features of the Dickens-Flynn model), and Robert Sternberg (who showed that IQ tests do not capture all of the cognitive skills prioritized by our society). My own contribution seems very modest. What thunderous truisms after 40 years of thought: *g* cannot legislate what IQ differences are significant, environment never goes away, arithmetic must be suited to what is there to be measured. To be fair to myself, I have set out a theory of intelligence (integrating brain physiology, individual differences, and social trends, or the BIDS theory) that I consider an advance on its predecessors (Flynn, 2016)

However, I have a sense of homecoming after a long journey. Have we not always known that determining whether the black-white IQ gap is environmental turns on only one thing, whether we can identify environmental differences potent enough to explain it? Have we not always known that the emergence of industrial society was a cultural revolution that changed everything—a new stage that made unprecedented demands on our brains and our minds and our competencies? The Piagetians saw it—what was the matter with us?

Declaration of conflicting interests

The author declares no conflicts of interest with respect to the authorship or the publication of this article.

References

Adam, S., Bonsang, E., Germain, S., & Perelman, S. (2007). Retirement and cognitive reserve: A stochastic frontier approach applied to survey data (CREPP Working Paper 2007/04) (Retrieved from) http://www.crepp.ulg.ac.be/papers/crepp-wp200704.pdf.

Altenburg, T. M., Chinapaw, M. P., & Singh, A. S. (2016). Effects *of* one versus two bouts of moderate intensity physical activity on selective attention during *a* school morning in Dutch primary schoolchildren*:* A randomized controlled trial. *Journal of Science Medicine Sport, 19*, 820–824.

Bouchard, T. J. (2013). The Wilson effect: The increase of heritability with age. *Twin Research and Human Genetics, 16*, 923–930.

Briley, D. A., & Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science, 24*, 1704–1713.

Coyle, T. R., & Pillow, D. R. (2008). SAT and ACT predict college GPA after removing "g". *Intelligence, 36*, 719–729.

Dickens, W. T., & Flynn, J. R. (2001a). Great leap forward: A new theory of intelligence. *New Scientist,* 44–47.

Dickens, W. T., & Flynn, J. R. (2001b). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*, 346–369.

Dutton, E., Bakhiet, S., Ziada, K., Essa, Y. & Blahmar, T. (in press). A Flynn effect in Khartoum, the Sudanese capital, 2004–2016. Intelligence.

Dutton, E., Van der Linden, D., & Lynn, R. (2016). The negative Flynn effect: A systematic literature review. *Intelligence, 59*, 163–169.

Eysenck, H. (1971). *Race, intelligence, and education.* London, England: Temple Smith.

Fernandes, H. B. F., Woodley, M. A., & te Nijenhuis, J. (2014). Differences in cognitive abilities among primates are concentrated on *G*: Phenotypic and phylogenetic comparisons with two meta-analytical databases. *Intelligence, 46*, 311–322.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*, 29–51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*, 171–191.

Flynn, J. R. (2008). *Where have all the liberals gone? Race, class, and ideals in America.* Cambridge, England: Cambridge University Press.

Flynn, J. R. (2009). *What is intelligence? beyond the Flynn effect.* Cambridge, England: Cambridge University Press.

Flynn, J. R. (2012). *Are we getting smarter? Rising IQ in the twenty-first century.* Cambridge, England: Cambridge University Press.

Flynn, J. R. (2013). *Intelligence and human progress: The story of what was hidden in our genes. London, England.* Elsevier.

Flynn, J. R. (2016). *Does your family make you smarter? Nature, nurture, and human autonomy.* Cambridge, England: Cambridge University Press.

Flynn, J. R., & Shayer, M. (2017). IQ decline and Piaget: Does the rot start at the top? *Intelligence, 66*, 12–121.

Flynn, J. R., te Nijenhuis, J., & Metzen, D. (2014). The *g* beyond Spearman's *g*: Flynn's paradoxes resolved using four exploratory meta-analyses. *Intelligence, 44*, 1–10.

Genovese, J. E. (2002). Cognitve skills valued by educators: Historic content analysis of testing in Ohio. *Journal of Educational Measurement, 96*, 101–114.

Haworth, C. M., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J., van Beijsterveldt, C., & Plomin, R. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry, 15*, 1112–1120.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class in American life.* New York: The Free Press.

Jensen, A. R. (1969). How much can we boost I.Q. and academic achievement? *Harvard Educational Review, 33*, 1–123.

Jensen, A. R. (1972). *Genetics and education.* New York, NY: Harper & Row.

Jensen, A. R. (1973). *Educability and group differences.* London, England: Methuen.

Jensen, A. R. (1980). *Bias in mental testing.* London: Methuen.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

Kraus, M. K., Rucker, J. M., & Richeson, J. A. (2017). Americans misperceive racial economic equality. *Proceedings of the National Academy of Sciences of the United States of America, 114*, 10324–10331.

Lewontin, R. C. (1970). Race and intelligence. *Science and Public Affairs: Bulletin of the Atomic Scientists, 24*(3), 2–8.

Lewontin, R. C. (1976). Further remarks on race and the genetics of intelligence. In N. J. Block, & G. Dworkin (Eds.). *The IQ controversy* (pp. 107–112). New York, NY: Pantheon.

Lynn, R. (2006). *Race differences in intelligence.* Whitefish, MT: Washington Summit Publishers.

Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations.* Westport, CT: Praeger/Greenwood.

Maguire, E. A., Woolett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. *Hippocampus, 16*, 1091–1101.

Matthews, L. C., & Turkheimer, E. (2017). *Flynn, the age-table method, and a metatheory of intelligence. Does your family make you smarter: Nature, nurture, and human autonomy.* Cambridge: Cambridge University Press Studies in History and Philosophy of Biological and Biomedical Sciences, 2016 (35–40).

Meisenberg, G. (2014). Cognitive human capital and economic growth in the 21st century. In T. Abrahams (Ed.). *Economic growth in the 21st century: New research* (pp. 49–106). New York, NY: Nova Publishers.

te Nijenhuis, J., & van der Flier, H. (2013). Is the Flynn effect on *g*? A meta-analysis. *Intelligence, 41*, 802–807.

Nisbett, R. (2015). *Mindware: Tools for smart thinking.* New York, NY: Farrar, Straus and Giroux.

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909-2013). *Perspectives on Psychological Science, 10*, 282–306.

Ritchie, S. J., Bates, T. C., & Deary, I. J. (2015). Is education associated with improvements in general cognitive ability, or in specific skills? *Developmental Psychology, 51*, 573–582.

Rushton, J. P. (1995). *Race, evolution, and behavior: A life history perspective.* Piscataway NJ: Transaction Books.

Sager, M. (2017 January). The smartest man in America. *Esquire, 30*, 2007.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype-environment effects. *Child Development, 54*, 424–435.

Sousa, A. (2011 December 19). Does practice really make perfect? *How the brain learns*.

Staff, R. T., Hogan, M. J., & Whalley, L. J. (2014). Aging trajectories of fluid intelligence in late life: The influence of age, practice, and childhood on Raven's progressive matrices. *Intelligence, 47*, 194–201.

Tizard, J. (1975 April 24). Race and IQ: The limits of probability. *New Behavior,* 6–9.

Woodley, M. A. (2012a). A life history model of the Lynn-Flynn effect. *Personality and Individual Differences, 53*, 152–156.

Woodley, M. A. (2012b). The social and scientific temporal correlates of genotypic intelligence and the Flynn effect. *Intelligence, 40*, 189–204.

Woodley, M.A., Figueredo, A.J., Sarraf, M.A., Hertler, S., Fernandes, H.B.F., & Peñaherrera-Aguirre, M. (2017). The rhythm of the west: A biohistory of the modern era, AD 1600 to the present. Journal of Social Political and Economic Studies, monograph series, no. 37. Washington DC: Scott Townsend Press.

Woodley, M. A., Sarraf, M. A., Peñaherrera-Aguirre, M., Fernandes, H. B. F., & Becker, D. (2018). What caused over a century of decline in general intelligence? Testing predictions *from the* genetic selection *and* neurotoxin hypotheses. *Evolutionary Psychological Science*. https://doi.org/10.1007/s40806-017-0131-7.

Woodley, M. A., te Nienhuis, J., & Murphy, R. (2013). Were the Victorians cleverer than us? The decline in general intelligence estimated from a meta-analysis of the slowing of simple reaction time. *Intelligence, 41*, 843–850.