Back

8 page(s) will be printed.

Record: 1

Title: Solving the IQ Puzzle.

Authors: Flynn, James R.[1]

Source: Scientific American Mind; 2007, Vol. 18 Issue 5, p24-31, 8p, 2 graphs, 1c

Document Type: Article

Author Affiliations: [1]Professor emeritus at the University of Otago in New Zealand and recipient of the university's Gold Medal for Distinguished Career Research

Full Text Word Count: 3938

ISSN: 1555-2284

Accession Number: 27094172

Database: Academic Search Premier

Solving the IQ Puzzle

The 20th century saw the "Flynn effect" — massive gains in IQ from one generation to another. Now Flynn explains why

On a rather dull Saturday in November 1984, I found a bombshell in my letterbox. I had received data from a distinguished Dutch researcher and saw immediately that Dutch males had made enormous IQ gains in a single generation. Today similar findings have occurred in almost 30 nations — in every country for which we have data. IQ escalation may not persist, but it has dominated the 20th century. That is enough to create a crisis of confidence. Either the children of today are far brighter than their parents, or at least in some circumstances, IQ tests are not good measures of intelligence. Paradoxes begin to multiply. Only now can we resolve them — and doing so illuminates the nature of intelligence as well as the gulf that separates our minds from those of our ancestors.

Intelligence and the Atom

Understanding intelligence is like understanding the atom: we need to know not only what holds its components together but also what splits them apart. What binds the components of intelligence together is the general intelligence factor, or g; what acts as an atom smasher is cognitive trends measured over time. The best IQ test to exemplify both these forces is the Wechsler Intelligence Scale for Children, or WISC, which has been used from 1947 through today.

The WISC's 10 subtests measure various cognitive skills. The Similarities subtest measures one's ability to perceive what things have in common; Vocabulary, whether you have accumulated the words used in everyday life; Information, your store of general information; Arithmetic, your ability to solve mathematical problems. People who are above average on one subtest tend to excel on them all. Therefore, we speak of a general intelligence factor. A mathematical technique called factor analysis measures the tendency of performance on a wide variety of cognitive tasks to be intercorrelated, and the construct called g is the quantified result.

A good performer typically exceeds the average person's results on some cognitive tasks more than others. These tasks tend to be those that are more cognitively complex, which reinforces the claim that g measures general intelligence. The WISC subtests can be ranked in terms of their g loadings. That simply means you rank them from the subtest on which high-IQ people beat the average person by the most down to the subtest on which they excel the least.

There is nothing mysterious about various traits or tasks having different g loadings. Musical people tend to be higher above average on the piano than on the drums. A talented chef is more likely to outdo the average person in the delicate task of whipping up a soufflé than in the simpler undertaking of scrambling eggs. The former is more complex than the latter and, therefore, is a better test of excellence in cooking.

Trends over Time

If general intelligence has increased over time, we would expect gains on each of the 10 WISC subtests to tally with their g loadings. But when we turn to IQ gains, we find something surprising: discrepancies between the magnitude of subtest gains and subtest g loadings. Similarities and Information have much the same g loadings, yet the former shows gains 12 times the size of the latter. Remember cooking. If skills improved over time, it would be amazing if the g loadings were ignored — for example, if there was an unexpected cooking gain in scrambling eggs but no gain in making soufflés.

Recent IQ gains show a chaotic pattern: 24 points on Similarities, whereas Vocabulary, Arithmetic and Information cluster around a mere three-point gain over 55 years [see box on opposite page]. The WISC gives not only subtest scores but also a summary judgment on intelligence, called Full Scale IQ. Its gains are huge, amounting to about 18 points. Raven's Progressive Matrices, which asks students to find the next step in a series of pictures, is also an

important test in analyzing IQ trends. Because American data are scant, I have offered a conservative estimate of a five-point gain per decade based on comparative data. How can our recent ancestors have been so unintelligent compared with ourselves? Even worse, British data suggest we have to extend the trend all the way back to 1900.

Now that I have explained the basic concepts behind the IQ boom, I can present the four paradoxes that it creates. Three arise out of the pattern and magnitude of IQ gains. The fourth also involves what we thought we knew about genes and environment.

PARADOX 1: The Factor Analysis Paradox

The patterns of IQ gains on the WISC subtests bear little relation to factor loadings. How can intelligence be both unitary (as it appears in factor analysis) and multiple (per the trends over time)? The key to this paradox is that factor analysis occurs in a static setting in which individuals are compared with social change held constant. IQ trends over time, however, take place in a dynamic setting in which social change alters cultural priorities, including which conceptual skills get greatest emphasis.

At any given time, for example, factor analysis would show that sprints and the high jump have large and similar g loadings, which is to say that people who have springy legs do well at both. But over time, young people may find sprinting romantic and the high jump boring. Performance on the first will escalate, and performance on the second will remain static. The correlation between the two events conceals the fact that there is little functional relation between the skills they require. You do not maximize your high-jump performance by sprinting toward the bar at top speed because you would mistime your jump. Improvement over time on the first is perfectly compatible with no improvement on the second.

To explain the IQ patterns, we need a functional analysis of what has elevated various cognitive skills over time. The rise of science has engendered a sea change in two respects: it has taught us that classifying the world using the categories of science is just as important as manipulating the world; and it has freed logic from the concrete, allowing us to work on abstractions with no concrete referents. In the early 20th century, a typical syllogism would have been: "Basset hounds are good at hunting rabbits. That is a basset hound. Therefore, I will use that dog when I hunt." Today we are far more likely to say the following: "Only mammals bear their young alive. Rabbits and dogs both bear their young alive. Therefore, they are both mammals."

If asked what dogs and rabbits have in common, a boy in 1900 would have said, "You use dogs to hunt rabbits." A boy in 2007 will say, "They are both mammals." It would never have occurred to someone a century ago to offer something so trivial. Who cares that dogs and rabbits are both mammals? What is important is what things are useful and under one's control.

The Similarities subtest of the WISC shows impressive gains throughout the past century because it gives zero for the hunting, or utilitarian, answer and full marks for the mammal, or classifying, answer. Subtests such as Vocabulary and Information are quite different. They sample the core vocabulary and general information needed in everyday life, and therefore the transition from the concrete to the abstract has left them largely unaffected. The other IQ test that shows sizable gains is Raven's Progressive Matrices. These increases are no longer mysterious. To do well, you must find it second nature to use logic to deal with abstract patterns — that is, you must perceive logical sequences in a series of shapes, something that is abetted by a modern culture that is more visually oriented.

It is easy to misunderstand the relation between Similarities and Raven's. Factor analysis of a wide range of mental tests showed that scores on these two have more in common than those of any other pair of tests. And now, both tests show the same huge gains over time. Nevertheless, the two tests are like sprints and the high jump, with almost nothing functional in common. The reason they correlate and their gains are so similar is that when a person benefits from seeing the world through scientific spectacles, he or she gets two distinct advantages. One is the liberation of logic from the concrete to analyze the abstract, which raises the Raven's score. The other is the transition from viewing the world as something to classify rather than merely to utilize, which raises the Similarities score. The same people are likely to enjoy both

these benefits much to the same degree. But they relate to two quite separate cognitive tasks nonetheless.

Factor analysis also shows that both Arithmetic and Raven's have high g loadings for a common factor. This fact has encouraged the notion that mathematical thinking and the cognitive problems posed by Raven's are functionally related. After all, Raven's problems demand that you see logical relations between shapes on the spot (without a previously learned method for doing so). Mathematics requires dealing with nonverbal material to master new proofs. Therefore, it seems sensible to teach young children Raven's-type problems so that they will become better mathematics problem solvers. Many U.S. schools have been doing just that since 1991.

Nevertheless, the large gains on Raven's and the virtually nonexistent gains on Arithmetic show that there cannot be a strong functional relation between the two. For nonmathematicians, mathematics is less a logical enterprise than a separate reality that obeys laws at variance with those of the natural world. Just as infants explore the natural world, children must explore the world of mathematics and become familiar with its "objects" through self-discovery. Raven's-type tasks make no contribution to that whatsoever.

Our first paradox is resolved. At any particular time, factor analysis will extract a robust g factor. Intelligence appears unitary, and the major cognitive skills are all highly intercorrelated. Over time, social reality reveals cognitive skills swimming freely of g, so intelligence appears multiple. If you want to see g, stop the film and extract a snap shot; you will not see it while the film is running. Society does not do factor analysis; it is a juggernaut that flattens factor loadings and imposes its own priorities.

2 PARADOX The Intelligence Paradox?

Gains in Full Scale IQ and Raven's suggest that our parents are some nine to 15 points duller than we are and that our children are nine to 15 points brighter. These gaps between generations should be noticeable in conversation and everyday life. Otherwise, must we not ask ourselves whether IQ gains really are intelligence gains?

But that is the wrong question. It implies all-or-nothing cognitive progress, whereas the 20th century has seen striking exceptions to the general trend. Look again at the box on page 26: the WISC subtests that show small gains are those most relevant to school-taught subjects. It is illuminating to compare their trends with those for the National Association of Educational Progress (NAEP) tests, often called the nation's report card.

From 1971 to 2002, fourth and eighth graders made a reading gain equivalent to almost four IQ points. By the 12th grade the gain dropped off to almost nothing. If we focus on WISC trends from 1972 to 2002, we see that schoolchildren made no gain in their store of general information and only minimal vocabulary gains. Therefore, although today's children may learn to master pre-adult literature at a younger age, they are no better prepared for reading more demanding adult literature. You cannot enjoy War and Peace if you have to run to the dictionary or encyclopedia every other paragraph.

From 1973 to 2000, fourth and eighth graders made mathematics gains equivalent to almost seven IQ points. The gain fell off at the 12th grade, this time literally to nothing. Increasing numbers of children have been mastering computational skills at younger ages. But the WISC Arithmetic subtest measures both computational skills and something extra. For example, consider this problem: "If four toys cost six dollars, how much do seven cost?" Many who can do straight paper calculations cannot diagnose the two operations required: that you must first divide and then multiply. Others cannot do mental arithmetic involving fractions.

My hypothesis is that children have mastered calculating skills at an earlier age but have made no progress in acquiring mathematical reasoning skills. Reasoning skills are essential for higher mathematics. Therefore, by the 12th grade the failure to develop enhanced mathematical problem-solving strategies begins to bite.

We now know why children today do not put their grandparents to shame in conversation.

Assume we hear a recent high school graduate chatting with his grandfather (who also finished high school) about a novel they both read the week before. There is no reason to believe either would have to make any allowance for the obtuseness of the other. If we were to discover essays on current affairs they both wrote shortly after graduation, there is no reason to believe that either would strike us as inferior to the other in terms of vocabulary or supply of general information.

3 PARADOX The Mental Retardation Paradox

Paradox three refers to our more remote ancestors, the Americans of 1900. If we put the average American of today at 100, the Americans of 1900 had a mean IQ of 50 to 70, which seems to signal a plague of mental retardation. We now know why we need draw no such inference. Our ancestors were no less intelligent; it is just that their intelligence was anchored in everyday reality. And it is an inability to cope with everyday life that characterizes someone who truly suffers from mental retardation.

The Vineland Adaptive Behavior Scale tells us coping skills remained stable during a period of rapid IQ gains. The performance of today's children (ages seven to 18) was compared with that of a random sample of children tested in 1984. Children had made no gains on the Communication and Socialization subtests. They had actually lost ground on a Daily Living Skills subtest. (It had obsolete items, such as "sews or hems clothes.")

The fact that we have not become more intelligent since 1900 does not imply that massive IQ gains over time are trivial. We can use abstractions, logic and the hypothetical to attack the formal problems that arise when science liberates thought from concrete situations. Since 1950 we have become much more ingenious in going beyond previously learned rules to solve problems on the spot [see box on opposite page].

PARADOX 4

The Genes and Environment Paradox

When identical twins are separated at birth and raised apart, they grow up to have IQs much more alike than randomly selected individuals would have. The obvious explanation is their identical genes, and these studies are taken as evidence that genes are potent and the environment is feeble. Yet massive IQ differences between one generation and another seem to signal the existence of environmental factors of enormous potency. Our fourth paradox asks, How can solid evidence show that environment is negligible (kinship studies) and powerful (IQ gains) at the same time?

Consider the identical twins John and Joe, who were separated at birth. Both live in an area that is basketball-mad. Their identical genes make them both taller and quicker than average to the same degree. John goes to school in one city, where he 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 and goes on to compete in high school, where he gets professional-style coaching. Joe goes to school in a city a few hundred miles away. Because his genes are identical to John's, and because he is taller and quicker than average to the same degree, he is likely to have a similar life history.

In other words, a genetic advantage that may have been quite modest at birth has a great effect on eventual basketball skills as they get matched with better environments — and genes thereby get "credit" for the potency of powerful environmental factors such as more practice, team play and professional coaching.

Now imagine one child who is born with a slightly higher aptitude than another child. Which of them will tend to like school, be encouraged, start haunting the library, get into top-tier classes and attend university? And if that child has a separated identical twin who has much the same academic history, what will account for their similar adult IQs? Not identical genes alone —

rather the ability of those identical genes to co-opt environments of similar quality will be the missing piece of the puzzle.

Genes have "profited" from seizing control of strong feedback loops that operate between performance and environment. A gene-based performance advantage causes a more-homework-done environment; the latter magnifies the academic performance advantage, which upgrades the environment further by leading to entry into a top-level class; this in turn magnifies the performance advantage once again, which facilitates access to a good university environment. These feedback loops have such an influence on the fate of individuals that my collaborator William T. Dickens of the Brookings Institution and I call them "individual multipliers."

There is also a "social multiplier." The industrial revolution in the late 19th and early 20th centuries demanded additional years of education. When a grade school education became the norm, everyone with middle-class aspirations wanted a high school diploma. When a high school diploma became the norm, everyone began to want a bachelor's degree. Economic progress created a middle class with new expectations about stimulating children intellectually, performing highly paid professional jobs in which they would be expected to think for themselves, and enjoying more cognitively demanding leisure activities. No one wants to seem deficient as a parent, unsuited for promotion, boring as a companion. Everyone responds to the new milieu by enhancing their performance, which pushes the average higher; they respond to that new average, which pushes the average higher still. Result: a dramatic escalation of cognitive skills in a single generation.

Within a generation, genetic differences drive feedback processes; between generations, environmental trends drive feedback processes. What looks potent depends on whose hand is on the throttle.

IQ gains have not inoculated people against credulity. Look at the number who believe in creationism and astrology.

A Hidden Trend and the Future

IQ gains have not inoculated people against credulity. Abstract categories and analysis can be used to defend nonsense rather than sense. Look at the number of people who believe in creationism, flying saucers and astrology.

Yet recent history has seen a second trend. The language of educated people has been enriched by words that can greatly enhance critical acumen. These terms each stand for a cluster of concepts that chart a method of analysis applicable to social and moral issues. I refer to concepts such as market (which became current in 1776), percentage (1860), natural selection (1864), control group (1875), random sample (1877), naturalistic fallacy (1903), charisma effect (1922), placebo (1938) and falsification (1959).

Thanks to division of the universities into specialties, no graduate is trained to use more than a fragment of these terms. The full potential of IQ gains over time goes unrealized. Because universities could have better educated their students at any time over the past century, improved performance in the 21st century is far from certain.

Adapted from What Is Intelligence? by James R. Flynn. Cambridge University Press, 2007.

(These large gaps in IQ between generations should be noticeable in conversations and in everyday life.)

FAST FACTS

Probing the Nature of Intelligence

During the 20th century, unexpected and massive gains on IQ tests — the Flynn effect — appeared in almost 30 countries, all of the nations for which data exist. Puzzlingly, the gains on subtests, which measure distinct components of intelligence, varied in a seemingly chaotic pattern.

- 2>> The results set off a crisis in intelligence research. Either the children of today are far brighter than their parents, or at least in some circumstances, IQ tests are not good measures of intelligence. Paradoxes began to multiply.
- 3>> The solutions to the paradoxes tell us something new about the nature of intelligence and what society must do to foster critical thinking.

Sample IQ Test Questions

Below are examples of the types of questions students answer on the 10 subtests of the WISC and Raven's Progressive Matrices.

WISC

Information	On what continent is Argentina?
Arithmetic	If four toys cost six dollars, how much do seven cost?
Vocabulary	What does "debilitating" mean?
Comprehension	Why are streets usually numbered in order?
Picture Completion	Indicate the missing part from an incomplete picture.
Block Design	Use blocks to replicate a two-color design.
Object Assembly	Assemble puzzles depicting common objects.
Coding	Using a key, match symbols with shapes or numbers.
Picture Arrangement	Reorder a set of scrambled picture cards to tell a story.
Similarities	In what way are dogs and rabbits alike?

RAVEN'S

Find the missing piece from the six pictured below.

IQ Gains and the Real World

Professional work roles enhance the ability to be innovative. They could hardly do that unless innovation was necessary to perform professional duties. Because society needs more and more people to do managerial, technical and professional jobs, gains in the ability to think on the spot rather than just follow rules (as measured by a test called Raven's Progressive Matrices [see box on page 28]) have social significance.

First-born children have more analytical interests. Reduced family size means that a higher percentage of children in recent years are first-born. Enhancing cognitive skills becomes a prerequisite for being a good parent. Parents must take "hypothetical" questions seriously — that is, they need to answer rather than dismiss the eternal string of "whys."

Video games and electronic games enhance problem solving in visual and symbolic contexts. Note the cognitive demands of games such as Tetris (spatial geometry), Myst (engineering

riddles) and Grand Theft Auto (mapping). Enhanced problem-solving skills have become necessary to fully enjoy our leisure activities. Chess grand masters are getting younger, yet the standard of play in tournaments continues to rise.

A generation ago TV programs such as I Love Lucy, Dragnet and Starsky and Hutch required virtually no concentration to follow. Beginning in 1981 with Hill Street Blues, single-episode dramas began to weave together as many as 10 threads into their plotlines. The hit drama 24 connects the lives of 20 or more characters, each with a distinct story. — J.F.

GRAPH

PHOTO (COLOR)

(Further Reading)

• *The Rising Curve: Long-Term Gains in IQ and Related Measures. Edited by Ulric Neisser. American Psychological Association, 1998.*

• *What Is Intelligence: Beyond the Flynn Effect. James R. Flynn. Cambridge University Press, 2007.*

~ ~ ~ ~ ~ ~ ~ ~ ~

By James R. Flynn

James R. Flynn is professor emeritus at the University of Otago in New Zealand and recipient of the university's Gold Medal for Distinguished Career Research. He has been named Scientist of the Year by the International Society for Intelligence Research and is a distinguished associate of the Psychometrics Center at the University of Cambridge.

The Long Rise of IQs

Gains are measured in IQ points (adopting the usual convention of setting the standard deviation at 15). One IQ test, the Wechsler Intelligence Scale for Children (WISC), was updated three times, which yields estimates of gains over three periods, collectively covering 1947 to 2002. Although there are no reliable U.S. data for Raven's Progressive Matrices, another IQ test, I have put gains conservatively at 0.5 IQ point per year. (This rate is the lowest for any developed nation for which we have data.)

— J.F.

GRAPH

Back