Brains vs. Bias: Crash Course Psychology #24

Crash Course: Psychology
https://youtube.com/watch?v=75g4d5sF3xI
https://nerdfighteria.info/v/75g4d5sF3xI

In your lifetime, you have probably stared down any number of ability tests and course exams and PSATs, SATs, ACTs, GREs, GCSEs, whatever you got in your country. Humans, it seems, really get a thrill out of measuring,

Unfortunately, as you saw last week, historically, we have been kind of bad at that. Today, we think of intelligence as determined by a series of factors related to genetics, environment, education, perhaps even randomness itself, some aspects of which may correlate with belonging to a particular social group, and others not. The key here, though, is that we don't fully understand how or how much some of these factors work.

How do elements like personal history and conditions like poverty, access to education, stress, even nutrition affect someone's scores on cognitive tests? And if a group of people share some of these conditions, how will they respond--both as individuals and as a group--to certain potentially biased intelligence tests? In the end, the irony is that in our ongoing effort to measure human intelligence, most of what we've learned is simply what we don't know.

[Intro]

What is a piano? Which one of these things is least like the others? Juice is to glass as hand is to what? Which one of these numbers does not belong in the series? Bernice had x number of jelly beans. She ate one, then gave half of what was left to Bruno, then she ate another and gave half of the remainder to her dog. Now, she only has five beans. How many did she--ahheh uhhhh.

These questions are similar to what you'd find on today's most widely-used intelligence tests. The Wechsler Adult Intelligence Scale, or WAIS, and the Wechsler Intelligence Scale for Children, or WISC. Originally published by psychologist David Wechsler in 1955, the current edition of the exam consists of fifteen different sub-tests that assess things like vocabulary, similarities between objects and concepts, and patterns in letters and numbers.

Cognitive tests usually fall into one of two categories: achievement, or the kind that reflect what you've learned, and aptitude - the kind that's supposed to predict your ability to learn something new. So the WAIS and the WISC are aptitude tests, and your final exam at the end of your math class is an achievement test.

So, how do we know if an intelligence test, or any other test for that matter, is actually any good? Well, today, we do have some standards. To be widely accepted, a test must hit three important marks. It has to be standardized, reliable, and valid. Standardization is basically about comparability. Whether you answer 15 or 50 questions correctly on a test actually means very little until you can compare those scores against how others performed. So, to achieve meaningful comparisons, test-makers must first give the test to a representative sample group, which sets a standard by which to compare future test-takers. You've probably heard of a bell-curve. Whether you're measuring height or mental aptitude or love of the Beatles, it's often assumed that everyone you're measuring will fall into what's called a normal pattern: most scores will fall in the mid-range, while fewer hit the extremes. And actually, it's those extremes that intelligence tests are most widely and effectively used for. It might help an educator identify a gifted student who totally blows the roof off a test, but they're also useful in helping clinicians determine who might have a disability or be facing some specific barrier. With victims of traumatic brain injury or stroke, for example, a WAIS test can do a nice job of sussing out whether a patient who's struggling with language actually has a problem remembering and identifying words, or if they're just having a hard time processing the information quickly.

But these tests should be regarded more skeptically when it comes to issues that are either way more specific or way more broad. Like, they won't be able to answer questions along the lines of, "Will Jesse get into Harvard?" or "Are women smarter than men?" They're just not designed to do that. And in any case, simply knowing where you fall on a normal curve on a standardized test doesn't mean much if the test is poorly designed, so along with being standardized, a good test has high reliability, meaning it yields dependably consistent results. One way to determine this is to have people take the same test a second time, or some similar version of it. If the two performances resemble one another -- if the scores correlate -- then the test is thought to have good reliability.

And the third requirement for the testing trifecta is simply validity, or the extent to which a test measures or predicts what it's supposed to. And there are different kinds of validity -- for instance, if I take the WAIS IQ test and my scores accurately predict how good my grades in college will be, that's a simple kind of predictive or criterion validity. On the other hand, if I take the test and my scores correlate strongly with my results on another similar cognitive test, like the Stanford Binet, that falls under the broad category of construct validity. The key is that all of these are ways to see if a test measures what it claims to measure. But the stickiest wicket of them all is what we make of the test scores themselves.

We've all heard plenty about the influence of both nature and nurture in psychology, so the big question is: do our genetics influence our intelligence, or does our environment? And this is an easy question to answer because they both do. And that is important. If the history of intelligence testing has taught us anything, it's that assuming everyone is smart in the same way and for the same reason can lead to disastrously bad conclusions. So let's look at the scientific evidence, and the best place for that is the wealth of twin and adoption studies, which have been fascinatingly helpful in illustrating how genetics and environment can both influence intelligence.

For example, research has shown that identical twins who were raised together have the highest rate of similarity in intelligence scores of any group. Fraternal twins who share only half the genes tend to be much less similar in their scores, even when raised in the same home. Likewise, neuroimaging studies show that certain brain regions, like those associated with language, are structurally similar between identical twins, and show similar activity while doing the same kinds of mental tasks. The brains of fraternal twins raised in the same home are very similar in some areas, but less so in others. But identical twins who are raised together have the same brain, at least, according to neuroimaging scans. Some studies even showed that identical twins raised apart from each other show higher intelligence correlation than fraternal twins raised together. And maybe even more interesting is that these intelligence correlations actually increase over time.

In one mega-study of eleven thousand twin pairs in four countries, that correlation kept increasing from middle-childhood to adolescence to young adulthood and continued to increase through adulthood. Similar research has looked at adopted children and compared their scores with those of their adopted siblings, parents, and biological parents. And the results can be kind of surprising, because as adopted kids grow up, their mental similarities to their adoptive families actually get smaller over time until there's virtually no correlation by adulthood. Instead, they become more similar in terms of mental aptitude to their biological parents over time, even if they never met. In other words: genes appear to matter. You could take a hundred kids and raise them in the exact same way, and as adults, they'd still have different aptitudes. But, does this mean that when it comes to intelligence, we're all nature and no nurture?

Well, luckily, and somewhat obviously, no. Life experiences and

environment also matter. One sad example of how early environments affect children can be found in the work of American psychologist J. McVicker Hunt in a destitute Iranian orphanage in the 1970s. Conditions were really, really bad; infants received minimal care, and whatever attention they did get was on a routine schedule, never in response to whether they were cooing or crying or anything else. Basically, they were being raised with no cause and effect between their behaviors and the responses of their caregivers, and as a result, they didn't learn how to communicate. With no stimulation or social response, the kids were just kind of passive, vacant lumps. Deprivation was essentially trumping any inborn intelligence. So, Hunt started a program. He trained caregivers to actually talk to infants, to teach them how to mimic sounds and actions, and eventually, sounds and words from their language. The results were tremendous. The kids started to learn really quickly, and basically just came alive. While it's an extreme example, Hunt's research showed how malleable early childhood intelligence can be, especially in disadvantaged and stressful conditions. So you can see that environment and heredity interact to affect intelligence, and that some tricky implications can come out of that conclusion. But that's hardly the only controversy when it comes to how we view and measure intelligence.

There have been some sensational headline-courting studies of genetic and social influences that have suggested that fundamental differences in intelligence may exist between genders and races, but many of these studies are tangled up in questions of how potential testing bias may affect performance. Basically, if a test inadvertently measures differences caused by cultural experiences or social factors instead of what we might call "innate intelligence", then you might say that the test is biased.

Extreme example: in the past, immigrants to the US were classified as "feeble minded" if they couldn't answer distinctly American questions, like "Who was the first American president?" or "What's a milkshake?". Today, concerns about bias focuses on differences among members of the same general culture. Say, a poor, rural kid who might be plenty smart, but will test low if questions involve urban, upper-class concepts like taking taxis and drinking tea out of china cups or the rules of tennis.

So the questions themselves can skew performance results, but who administers the test can also affect outcomes. Women tend to do better with a fellow female administrator, and African Americans often score higher if their test is given by an African American instructor. And the risk of bias may even fall to the test-takers' own expectations. For example, many studies have found that if you give a math test to equally capable men and women, but just before starting, you tell the subjects that women usually score lower than men, you actually negatively affect the women's performance. This self-fulfilling concern that you might mess up and inadvertently fulfill some negative stereotype is called stereotype threat. It was first described by social psychologist Claude Steele and Joshua Aronson, and it's been demonstrated frequently across a whole host of interesting studies.

Now, we've only scratched the surface of this mess that is intelligence testing. An important thing to remember next time you ace or bomb a test is that you are far more complicated and nuanced than any test score. Don't let a number puff you up or drag you down, and don't let it define you. We all have room for self-improvement. We are all full of infinite surprising potential. Ah, and answers to the questions I asked earlier: A piano is a musical instrument played using a keyboard, banana is the least similar to the others, juice is to class as hand is to glove, the number two does not belong in the series, and Bernice began with twenty-three jellybeans.

Today, your intelligent mind learned how we currently use WAIS and WISC tests to measure intelligence, and how important it is that a test be standardized, reliable, and valid. We also looked at how genetics, environment, testing bias, and stereotype threat can affect IQ test performance. Thank you for watching, especially to all of our Subbable subscribers who make Crash Course possible. To find out how you can become a supporter, just go to Subbable.com/CrashCourse.

This episode was written by Kathleen Yale, edited by Blake de Pastino, and our consultant is Dr. Ranjit Bhagwat. Our director and editor is Nicholas Jenkins, the script supervisor is Michael Aranda, who is also our sound designer, and the graphics team is Thought Café.