Measuring Student Potential —with Genetics

How DNA analysis may be able to predict education attainment

By MICHAEL J. PETRILLI

BACK IN 1965, when he proposed the Elementary and Secondary Education Act to Congress, President Lyndon B. Johnson wrote, "Every child must be encouraged to get as much education as he has the ability to take."

Today, that language might feel a bit dated (and gendered)—instead, we might say we're aiming for "all children to achieve their full academic potential." But the same idea is lurking inside these seemingly anodyne statements, and it is both common sense and something those of us in education policy don't like to say out loud.

Ability matters *and varies* from human to human. As a result, we don't all have the same ability "to take" the same amount of education.

While this might seem obviously true, particularly to anyone who has ever raised more than one child, is it scientifically true? Is innate ability, written in our genes, a real thing? If so, can we measure it?

Answering in the affirmative is Dalton Conley, a Princeton professor whose fascinating new book, *The Social Genome,* explores the interaction of our genetic code and the social environment. He demolishes what he calls "blank-slatism"—the assertion that variation in individual human traits, behaviors, and outcomes is caused entirely by our life circumstances, including social class. But he also takes down the "hereditarians," who argue that everything is predetermined by our genes. It's not nurture versus nature, he argues, but nature *and* nurture both, linked by an intricate dance whereby our genes seek out environments to fully express themselves.

MARK SMITH

Nobody in education needs convincing that "nurture" factors like the home environment kids grow up in, the type of neighborhood they live in, and the friends they keep have an impact on schooling outcomes. Indeed, for decades researchers testing the effects of new interventions have tried to control for these factors, so correlated they are to success or failure in school. But what may be harder to accept is that our DNA has a major impact on our education outcomes as well.

We don't tend to resist this line of thinking when it comes to physical attributes. Tall parents beget tall children. Kids who are athletically prodigious often have moms and dads who were superstars on the field, too. When it comes to cognition and other school-related skills, however, we grow sheepish about the role of our genes—and for good reason. We've all studied the horrors that resulted from the eugenics movement of a century ago, culminating in the gas chambers at Auschwitz and beyond. And we've all read or seen science fiction novels and movies, from *Brave New World* to *Minority Report* to *Gattaca,* that warn us of a future when genetic predispositions are taken to be determinative, or when tinkering with our genes to create tailor-made superhumans becomes the dystopian norm. (This is indeed already happening at fertility clinics worldwide to some extent, as parents select for preferred traits.)

Conley worries about all this, too, but can't deny what science is teaching us: Our genetic code has a big impact on many human traits, behaviors, and outcomes, and we're getting better at measuring the relationship.

Introducing the Education PGI

As Conley patiently explains in his book—and patiently explained to me in an interview—ever since the completion of the Human Genome Project in 2003, researchers have been racing to nail down the connection between individual genes and diseases, conditions, and attributes. What quickly became clear was that, in almost every case, variations in human attributes aren't caused by variations in single genes, but in small variations across thousands of genes.

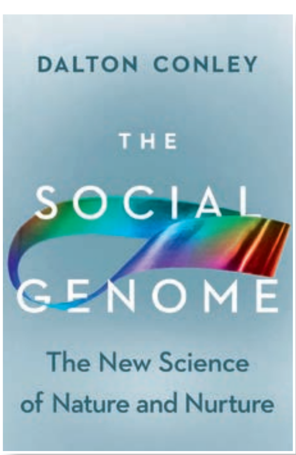
There's no one gene for height, for example. But remarkably, researchers have been able to use advanced computational methods to identify the thousands of genes that together predict someone's height and can calculate a "polygenetic index" (or PGI) for that attribute. This number measures an individual's likelihood of exhibiting a specific trait—or, in Conley's words, is "a single number for a single trait that sums the effects of thousands of different variations in our DNA across our genome observed in the population."

So when it comes to height, today anyone can have their "height PGI" calculated based on a genetic analysis of a cheek swab. Their genetic data will be fed into a complex algorithm and the score will predict their height within about one inch. (The difference comes down to the environment and depends on someone getting adequate nutrition—in the womb and beyond—and avoiding serious childhood accidents and illnesses.) We can talk, then, about whether someone has achieved their full genetic potential in terms of height.

Our genetic code has a big impact on many traits, behaviors, and outcomes, and we're getting better at measuring the relationship.

Might we also someday be able to determine whether someone has achieved their full genetic potential in terms of education attainment? Harkening back to President Johnson, could we find out how many children "take" as much education as their ability allows? That is the tantalizing possibility presented by the "education PGI," a similar calculation designed to predict someone's eventual level of education based on their genetic code.

The first education PGI calculations were released in 2013 and were very imprecise, estimating only 3 percent of the variation in education attainment from person to person. A substantially updated version released in 2022 can explain 16 percent. To put that in context, consider this: Studies of twins indicate that education attainment is about 40 percent heritable, so we could describe education PGI as explaining about 40 percent of that 40 percent.



Our genetic code impacts traits, behaviors, and outcomes, Conley observes.

That's still very noisy data, much too imprecise to predict individual outcomes—especially those in the middle of the education PGI distribution. However, a PGI's predictive power is far stronger for individuals in the top or bottom 10 percent, and there we find insights from education PGI. As Conley writes, someone in the bottom 10 percent of the ranking has about a 7 percent likelihood of completing a four-year degree, while an individual in the top tenth has about a 71 percent chance of graduating from college.

Avoiding Dystopian Uses

Like virtually all polygenetic indices, education PGI can only be used for non-Hispanic whites—a major limitation to its utility. Currently, most of the DNA samples used in the research come from this one population and studies indicate they don't work for people of other races. But it's not hard to imagine researchers building similar indices for all major racial groups, as well as these polygenic computations growing more accurate over time.

Researchers are also examining the degree to which education PGI—again, which focuses on education attainment—overlaps with the PGI for particular traits. We already know there's significant overlap with cognition, which is not surprising. But it might also overlap with certain personality traits as well as what we call "noncognitive" skills, like self-regulation, organization, and persistence. This may help settle some debates in education about what really matters when it comes to school success—brains, knowledge acquisition, perseverance, and so forth. Granted, some of those things that matter may be more malleable than others.

If these improvements come to pass, the education PGI could have significant predictive power, perhaps akin to third-grade test scores today. But as the science fiction genre teaches us, it's not hard to imagine the great harm that could follow.

Conley worries, in particular, about the potential for teacher bias to subject kids with low index scores to the soft bigotry of low expectations (see "The Power of Teacher Expectations," *research*, Winter 2018). In our interview, Conley pointed me to the Pygmalion experiment from the 1960s, whereby teachers treated some kids differently after being told they were geniuses—and those kids did better (and their peers worse) as a result. Students also might hold self-limiting views if they were informed that they had a low education PGI. Researchers like Conley might try to warn the rest of us that PGIs are probabilistic and not perfect soothsayers, but those warnings could fall on deaf ears.

Another urgent concern is whether education PGI is picking up on social class versus innate talent. For example, taller people tend to have a higher education PGI—not because height makes you smarter, but because both height and education attainment are related to socioeconomic status.

Positive Use Cases

That sort of interaction between genes and social environment is at the heart of Conley's book but also raises questions about how we might use education PGI alongside traditional measures of student potential. Education PGI could provide a new way to identify students who would benefit from gifted education and other accelerated programs, for example.

Experts in the field of gifted education have long worried that schools are overlooking many students who would benefit from advanced education, especially students from low-income backgrounds, who are substantially underrepresented in gifted programs. While "universal screening" via test scores helps—it's certainly better than relying on teacher recommendations, which can suffer from racial and socioeconomic bias—it's still not perfect. Test scores for disadvantaged students are often suppressed by environmental factors, especially in the early grades. Imagine if a student's sky-high education PGI would reveal their likelihood of having impressive innate talents that could be developed with the right nurturing.

We can talk about whether someone has achieved their full genetic potential in terms of height. Might we also someday be able to determine whether someone has achieved their full genetic potential in terms of education attainment?

But that will only work if education PGI isn't itself biased against kids from low socioeconomic backgrounds, which it very well may be. That needs to be fixed.

Perhaps education PGI could be used in research studies someday as well. Imagine you want to know which states were doing a good job helping the children growing up there achieve their full academic potential. We understand that's not just about the effectiveness of schools, but also about everything that kids experience in their environment—from the moment of conception through high school graduation. Are states doing everything they can to help nurture students' talents? Via social service programs, support for parents, great schools, and everything else?

If you had a big enough sample of students' DNA and could follow their trajectories through their early to mid-20s, you could check to see whether college completion rates were in line with what their education PGIs predicted. That could identify states where students are thriving beyond what we might have expected, as well as how much academic potential other states were leaving on the table.

Making sure education PGI will be used for good and not for ill will be challenging, and even Conley acknowledges that some sort of regulation will be necessary before any of this is ready for primetime. But like AI and other transformative technologies, it's coming, sooner or later, so now is the time to prepare.

Michael J. Petrilli is president of the Thomas B. Fordham Institute, visiting fellow at Stanford University's Hoover Institution, and an executive editor of Education Next.