Four Facts about Human Capital
Faculty Research Working Paper Series

David Deming
Harvard Kennedy School

June 2022
RWP22-007

Visit the HKS Faculty Research Working Paper Series at:
https://www.hks.harvard.edu/research-insights/publications?f%5B0%5D=publication_types%3A121

The views expressed in the HKS Faculty Research Working Paper Series are those of the author(s) and do not necessarily reflect those of the John F. Kennedy School of Government or of Harvard University. Faculty Research Working Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate on important public policy challenges. Copyright belongs to the author(s). Papers may be downloaded for personal use only.
Four Facts about Human Capital
David J. Deming
NBER Working Paper No. 30149
June 2022
JEL No. I25,I26,J24

ABSTRACT

This paper synthesizes what economists have learned about human capital since Becker (1962) into four stylized facts. First, human capital explains at least one-third of the variation in labor earnings within countries and at least half of the variation across countries. Second, human capital investments have high economic returns throughout childhood and young adulthood. Third, we know how to build foundational skills such as literacy and numeracy, and resources are often the main constraint. Fourth, higher-order skills such as problem-solving and teamwork are increasingly valuable, and the technology for producing these skills is not well-understood. We know that investment in education works and that skills matter for earnings, but we do not always know why.

David J. Deming
Harvard Kennedy School
Malcolm Wiener Center for Social Policy
79 JFK St
Cambridge, MA 02138
and Harvard Kennedy School
and also NBER
david_deming@harvard.edu

For acknowledgments and financial disclosure information, see:
https://www.nber.org/papers/w30149
Human capital theory is the now widely accepted idea that education, training and other forms of learning are investments that pay off in the future. Like any capital investment, the costs of schooling are paid up front and the benefits are earned later. To be sure, schooling has benefits far beyond its monetary value, but the relationship between schooling quantity (and quality) and future earnings is one of the most robust findings in social science.

The term “human capital” was initially controversial among the pioneers of human capital theory, who wanted to explicitly reject the implication that people should be treated as property, or that workers are assets who in any sense “belong” to the owners of capital (Goldin and Katz 2020). Despite initial discomfort over terminology, the study of human capital has blossomed. This is in part because people all around the world spend much more money and time on education than they did a half-century ago. Between 1950 and 2010, the share of the world adult population with at least some secondary school education increased from 13 percent to 51 percent, and the share with some tertiary education increased nearly sevenfold, from 2.2 percent to 14.6 percent (Lee and Lee 2016). In the United States, education spending increased from 3.1 percent of GDP in 1950 to 7.1 percent in 2018, with most of the increase coming from the public sector (Digest of Education Statistics 2019, Table 106.10). This pattern generally holds for other countries around the world, with faster increases in public spending on education in developing countries (Roser and Ortiz-Ospina 2016).

Research interest in human capital within the economics profession has grown explosively in the last few decades. I conducted a text and title search on EconLit for the phrases “human capital,” “education,” and “skill.” At the time of the seminal work of Becker (1962), about 2.5 percent of articles included at least one of those phrases. This share didn’t change much until about 1990 but then it started rising, reaching about 15 percent of all articles since 2015.

This paper synthesizes what we have learned about human capital since Becker (1962) into four stylized facts. First, human capital explains a substantial share of the variation in labor earnings within and across countries. Second, human capital investments have high economic returns throughout childhood and young adulthood. Third, the technology for producing foundational skills such as numeracy and literacy is well-understood, and resources are the main constraint. Fourth, higher-order skills such as problem-solving and teamwork are increasingly economically valuable, and the technology for producing them is not well understood.

We have made substantial progress toward validating the empirical predictions of human capital theory. We know how to improve foundational skills like numeracy and literacy, and we know that investment in these skills pays off in adulthood. However, we have made much less progress on understanding the human capital production function itself. While we know that higher-order skills “matter” and are an important element of human capital, we do not know why.

1 The structure of the article models Nicholas Kaldor’s (1961) six “stylized” facts about economic growth as well as the six “new Kaldor facts” discussed in Jones and Romer (2010). Both articles summarized the state of knowledge about economic growth and successfully framed the research agenda going forward. That is also my goal for research on human capital.
Fact 1: Human capital explains a substantial share of the variation in labor earnings within and across countries

The “Mincer equation,” as it is colloquially known, is an important building block of human capital theory. Mincer (1974) starts with a formal model where identical agents make forward-looking investments in human capital to maximize the present value of future earnings and derives this relationship:

\[ \ln y_i = \alpha_i + \beta S_i + \gamma X_i + \delta X_i^2 + \varepsilon_i. \]

The Mincer equation models log annual earnings (or sometimes hourly wages) \( y_i \) as an additive function that is linear in years of schooling \( S_i \) and quadratic in years of experience \( X_i \). Although subsequent work has proposed adding higher order terms in experience and non-linearities in education, the Mincer equation has mostly withstood the test of time (Lemieux 2006).

The Mincer model’s simple functional form spawned a large literature of different approaches to estimating \( \beta \), the economic return to an additional year of schooling. Across many different countries and settings, estimates of \( \beta \) yield a coefficient of around 0.1, which implies that another year of schooling increases earnings by 10 percent (Gunderson and Oreopoulos 2020; Patrinos and Psacharopoulos 2020).

One immediately apparent issue is the potential endogeneity of schooling. Rational agents will invest more in schooling when they expect to receive higher returns, and thus a naïve comparison of earnings between individuals with different amounts of completed education will suffer from “ability bias” as noted by Griliches (1977), Card (1999), and many others.

One solution is to find an instrumental variable that affects schooling but is unrelated to ability or other determinants of earnings. The search for such instrumental variables has been a central focus for labor economists over the past few decades. Possibilities include distance to the nearest college (Card 1995), compulsory schooling laws that vary across countries and states and over time (Angrist and Krueger 1991), the timing of school construction (Duflo 2001), and the expansion of funding for primary schools (Khanna 2021).

An alternative “regression discontinuity” approach uses discontinuous changes in the probability of admission around grade or test thresholds to estimate returns to education. Zimmerman (2014) finds that students with a grade point average just high enough to be admitted to Florida International University have 22 percent higher earnings a decade after they apply. This translates into an 11 percent return for a year of education if there is no return to community college, or an 18 percent return if the plausible alternative for many of these students—a year of community college—is worth the same as a year at Florida International University.\(^2\) Several other studies find positive earnings impacts of

---

\(^2\) The admissions standards at Florida International University were more generous than any other four-year public university in Florida, and so students who did not meet that admissions threshold mostly attended community colleges or did not go to college at all. Barrow and Malamud (2015) calculate that the return to a year of college would be 18 percent if the earnings difference around the threshold in Zimmerman (2014) was due entirely to the difference in average years enrolled.
admission to high schools and colleges, with some emphasizing the quantity and others the quality of education.  

The bottom line is that naïve cross-sectional comparisons and studies with strong quasi-experimental research designs yield very similar estimates of the economic return to education. Overall, studies that identify returns to education using instrumental variables, regression discontinuity, and other quasi-experimental approaches yield estimates of an additional year of education ranging between 6 and 18 percent, with a median in the 10-12 percent range. This is slightly higher than the 10 percent return from a “naïve” Mincer model, most likely because of some combination of measurement error and higher returns for marginal students (Card 1999). Across all OECD countries, the median earnings premium for a four-year college / tertiary education is 52 percent, or roughly 13 percent per year of education (based on OECD.Stat data at https://stats.oecd.org/Index.aspx?DataSetCode=EAG_EARNINGS).

Card (1999) finds that a standard Mincer model with a linear schooling term explains between 20 and 35 percent of the variation in labor earnings using the Current Population Survey (CPS), a cross-sectional survey of U.S. workers. However, it is not possible with CPS data to follow workers over the life course or to account for possible “ability bias” in the return to schooling.

Table 1 uses data from the 1979 National Longitudinal Survey of Youth (NLSY79), which tracks a cohort of youth ages 14 to 22 in 1979 as they progress through the labor market. To estimate returns to education over the life-course, I compute the average inflation-adjusted hourly wage for individuals between the ages of 25 and 54 over multiple observations, and then regress log average hourly wages on years of education, race and gender indicators, and cognitive ability as measured by adolescent scores on the Armed Forces Qualifying Test (AFQT).

Column 1 shows that the average return to a year of education over an individual’s prime working years is 10.9 percent. The R-squared of this regression is 30 percent. Controlling for AFQT to account for “ability bias” reduces the coefficient on years of education to 7.2 percent and increases the R-squared of the regression to 35 percent. Column 3 shows the average return for different levels of educational attainment, with less than high school as the left-out category. High school graduates and four-year college graduates earn an average of 13 percent and 48 percent higher wages than those with less than a high school education, respectively. These results are similar in magnitude to the quasi-experimental studies discussed above and to naïve cross-sectional estimates from other data sources. Basic measures of human capital such as education and cognitive ability can explain at least one-third of the variation in wages in a recent cohort of US workers.  

However, one-third is probably a lower bound for the impact of human capital on earnings, for three reasons. First, the calculation here does not include variation in education quality between workers with the same level of attainment. Quality adjustment is particularly important, because nearly all expansion

---

3 For example, Hoekstra (2009) finds that white men have 20 percent higher earnings when they barely meet an admissions threshold at a state flagship university. Canaan and Mouganie (2018) find that students who marginally pass a French high school exit exam enroll in higher-quality colleges and earn 12.5 percent more, despite no increase in the quantity of education.

4 Hoffman, Lee, and Lemieux (2020) estimate that education is responsible for more than half of the growth in earnings inequality in the U.S. since the 1970s, and nearly 75 percent of the growth in inequality since the late 1980s.
of US postsecondary education over the last few decades has occurred within less-selective institutions. Carneiro and Lee (2011) estimate that the college premium would have grown an additional 30 percent between 1960 and 2000 if pre-college education quality were held constant for the marginal college graduate.

Second, several studies find a larger role for human capital when it is measured in a way that includes education but also other attributes. For example, Smith et al. (2019) study the impact of owner death or retirement on private pass-through businesses and find that 75 percent of profits are attributable to the owner’s human capital, rather than physical or financial assets. Card et al. (2018) and Song et al. (2019) decompose the variance of earnings in matched employer-employee data and find that “worker effects” account for 40 percent of the variance in earnings in West Germany and 50 percent in the US respectively. Because worker effects are invariant to firm pay premia and occupational shifts by construction, we can reasonably consider them an estimate of workers’ human capital.

Third, there is the possibility of human capital externalities, where one person’s education increases the earnings of others around them. The literature on human capital externalities is not settled, with some studies finding little or no evidence and others finding relatively large agglomeration effects of working in geographic areas or firms with higher levels of human capital (Acemoglu and Angrist 2001, Moretti 2004, Ciccone and Peri 2006, Gennaioli et al 2012). Externalities, if they exist, would only increase the importance of human capital for explaining variation in labor earnings.

Some authors argue that schooling simply reflects higher human capital, rather than causing it (Caplan 2019). The signaling model of Spence (1974) suggests that individuals invest in education because of the information value it sends to employers about their productivity. It is difficult to disentangle human capital and signaling empirically, and both explanations surely contribute somewhat to explaining returns to education. However, I think the contribution of signaling is probably small, for two reasons.

First, many studies find positive returns to education even when no degree or credential is earned. This is important because signaling theory requires employers to observe the signal, and most people don’t report years of education on a resume. For example, studies of compulsory schooling compare groups of students who all seek to drop out as soon as they can, but some are required to stay in school longer based on when they were born during a calendar year. Many of the youth staying in school for an extra year do not end up obtaining a high school degree at all—they drop out in 11th grade rather than 10th grade. Nonetheless, such studies show that additional education leads to gains in earnings. A school construction program in Indonesia studied by Duflo (2001) mostly worked by increasing primary school enrollment, not receipt of degrees—but still led to later gains in wages. Aryal, Bhuller and Lange (2022) cleverly exploit the differential observability of compulsory schooling laws across regions in Norway to separate returns to human capital from signaling, and find that human capital accounts for 70 percent of the private return to secondary school education.

Various studies find large labor market returns to increased coursework requirements and specific skills and knowledge learned in high school or college, even if they do not lead to more degrees being earned (for example, Arteaga 2018). Moreover, some fields such as engineering, law and medicine impart concrete skills and specialized knowledge differences that self-evidently reflect human capital accumulation. No one was born knowing how to be a heart surgeon.
Second, empirical support for signaling theory is scant. Clark and Martorell (2014) find no difference in earnings between high school students who barely pass or fail an exit exam, implying that there is no signaling value of a high school diploma. Some studies do find that the return to education decreases over time as employers learn workers’ true ability, which is a testable implication of the signaling model (Altonji and Pierret 2001; Lange 2007). Yet a similar test of the employer learning model in a more recent cohort finds that the return to education does not diminish with experience (Castex and Dechter 2014).

The evidence described above suggests that human capital explains at least one third of the variation in labor earnings within the US. How much of the cross-country variation in earnings can be explained by human capital?

Following Solow (1956) and Hall and Jones (1999), a standard approach here is to look at an aggregate production function for the economy, where total output is expressed in terms of inputs of quality-adjusted human capital, physical capital, and technology.\(^5\) While data on output, education, physical capital, and the labor force are widely available, data on technology is not, and so the measurement of technology – often called total factor productivity (TFP) – shows up in cross-country studies as the “Solow residual” that is not explained by other measured variables. Mankiw, Romer, and Weil (1992) show that countries with higher rates of human capital grow faster, and that human capital is positively related to GDP growth over a 25-year period.

However, just as in the Mincer model, cross-country differences in schooling are probably endogenous: that is, countries with better technology will benefit more from investments in human capital and thus tend to make such investments more often, and so causality cannot be inferred from the basic relationships.\(^6\) The solution in the individual case involves seeking out methodologies or experiments that change the level of schooling, holding other factors constant. For cross-country differences, the ideal experiment would vary a country’s human capital stock or its total factor productivity, holding the other factors constant.

Hendricks and Schoellman (2018) approximate this experiment by studying the wage gains from migration. If skills travel with individuals when they migrate, relative wages across countries with different technologies and institutions can inform us about the contribution of human capital to cross-country income differences. Hendricks and Schoellman (2018) measure pre- and post-migration wages of U.S. migrants using the New Immigrant Survey (NIS). They find that migrants to the United States

\(^5\) Hall and Jones (1999) consider an aggregate production function for the economy written in terms of log output per worker:

\[
\ln \left( \frac{Y}{L} \right) = \frac{\alpha}{1 - \alpha} \ln \left( \frac{K}{Y} \right) + \ln \left( \frac{H}{L} \right) + \ln \left( \frac{A}{L} \right)
\]

where \(Y\) represents total output in country \(c\), \(K\) is capital and \(\alpha\) is the capital share, \(H\) is quality-adjusted labor, and \(A\) is a term representing the state of technology, often called total factor productivity (TFP). This equation can be estimated using cross-country data on incomes and factor shares, with human capital per worker \(\frac{H}{L}\) measured using years of schooling or other data on educational attainment.

\(^6\) Development accounting estimates of the importance of human capital for economic growth depend greatly on measurement and on the assumed structure of the aggregate production function – Rossi (2018) and Hendricks and Schoellman (2021) are excellent reviews of the literature.
from low-income countries experience wage gains equal to 38 percent of the total GDP-per-worker gap in each source country. Intuitively, these migrants are experiencing a change in TFP and institutions while their human capital is held constant. If the wage gains from this change are equal to 38 percent of the cross-country difference in GDP-per-worker, the remaining 62 percent is explained by human capital.

Their approach has two potential sources of bias. First, human capital may not fully transfer across countries. However, when they apply their method to immigrants who come to the United States on employment visas, have job offers in hand, and work in the same occupation, they show that human capital still accounts for at least 50 percent of cross-country income differences in these cases. Second, immigrants may be self-selected in the sense that those with an expectation of larger earnings gains may be more likely to migrate. However, selection on gains would bias cross-country earnings differences upward, leading them to understate the importance of human capital. In a follow-up paper, Hendricks, Herrington, and Schoellman (2021) use the wage gains from migration to calibrate models of development accounting under different assumptions, and estimate that human capital explains between 50 and 75 percent of cross-country income differences.

Overall, the best evidence suggests that human capital accounts for at least one third of the variation in labor earnings within countries, and at least one half of the variation in earnings per worker across countries.

Fact 2: Human capital investments have high economic returns throughout childhood and young adulthood.

The last few decades have seen increased public support for early childhood investment in the United States and around the world. In the United States, the share of four-year olds enrolled in state-run preschool increased from 15 percent in 2003 to 34 percent in 2019 (National Institute for Early Education Research 2021). Between 1973 and 2014, the number of children in the world enrolled in pre-primary education increased from 43.6 million to 155.1 million (Roser and Ortiz-Ospina 2017). Some of the motivation for this increase is the belief that the payoff to early-life investment is especially high. However, while a body of empirical evidence confirms the high returns to early-life investment, evidence also confirms similarly high returns to an array of young adulthood investments.

In a series of papers, James Heckman and his co-authors have argued that the economic return to human capital investment diminishes as children age. Figure 1 reproduces the Heckman Curve, a key illustration of the concept of diminishing returns on skill investment (Heckman 2006). The figure shows a declining rate of return, with a horizontal “break even” line for public investment in human capital formation that intersects somewhere during school-age.

Cunha and Heckman (2007) formalize these ideas with a model of life-cycle skill formation. Agents are born with human capital (which could reflect genes, parental education, income, and other fixed factors) and an initial endowment of skills that can expand over time. They consider a general “technology of skill formation” where early investments can matter more than late investments, and where it is not always possible to fully remediate early skill deficits. A key idea from their model is “self-productivity,” which is captured by the memorable phrase “skills beget skills.” As an intuitive example, self-productivity matters for cumulative learning processes such as mathematics, where concepts build upon
one another. More broadly, early childhood investments can raise the level of human capital in a way that increases the productivity of later childhood investments.

Another key idea in the Cunha-Heckman model is “dynamic complementarity”. Imagine that there is a fixed budget of skill investment dollars available to be spent on each child. Dynamic complementarity suggests that a balanced investment portfolio yields higher returns than spending lots of money later on and very little early in life, for example. The combination of self-productivity and dynamic complementarity implies that later investments are not very productive and that they cannot easily remediate early skill deficits. This rationalizes the “Heckman Curve” in theory.

There is strong evidence supporting value of skill investments in early childhood. Perhaps best-known are two randomized evaluations of preschool interventions from the 1960s: the High/Scope Perry Preschool Project and the Carolina Abecedarian Project. These studies are from decades ago, involving small-scale and intensive interventions for highly disadvantaged families, and thus their results may not generalize to larger and more recent programs. However, several studies of more recent preschool interventions also find substantial impacts —although the evidence also provides a puzzle. Pre-K programs often provide only a short-term boost to test scores that fades out in a few years. Yet they have longer-run impacts on important life outcomes such as high school graduation and college attendance, as well as non-educational metrics like reductions in crime and teen pregnancy and improved health later in life (for example, Ludwig and Miller 2007, Deming 2009, Gray-Lobe, Pathak and Walters 2021).

In addition, there are many other methods of early childhood investment: pre-natal care, early child health care, food and nutrition support, home visits to encourage practices like breast-feeding and smoking cessation, and others. In recent essays in this journal, Aizer, Hoynes, and Lleras-Muney (2022) describe the evidence of long-term benefits from policy interventions affecting low-income children like Medicaid and food stamps, while Wust (2022) presents the evidence from the Nordic countries about the benefits of universal provision of early childhood investments in pre-natal care, health care at time of birth, and early childhood health care. Again, any benefit-cost analysis of these programs needs to take a long-term view, because many of these benefits only become apparent later in life.

The Heckman curve has practical implications for policymaking. Heckman (2006) writes that “early interventions targeted toward disadvantaged children have much higher returns than later interventions such as reduced pupil-teacher ratios, public job training, convict rehabilitation programs, tuition subsidies, or expenditures on police.”

But of course, evidence of high returns from early childhood interventions does not imply lower returns for other interventions. Indeed, a body of evidence suggests that human capital investments have high returns through childhood and young adulthood. Hendren and Sprung-Keyser (2020) summarize the findings of 133 experimental and quasi-experimental policy interventions in the United States—interventions affecting a wide variety of age groups—using a unified welfare analysis framework. They

---

7 Cunha and Heckman (2007) propose a two-period constant elasticity of substitution (CES) production function for adult skills $A = h[yI_1^\phi + (1 - y)I_2^\phi]^{1/\phi}$ where $0 \leq y \leq 1$ is a share parameter and $-1/(1-\phi)$ is the elasticity of substitution with $\phi \leq 1$. The importance of “self-productivity” is increasing in $y$ because of the higher relative weight on early life investments. “Dynamic complementarity” is decreasing in $\phi$, for example as $\phi \to -\infty$ the production function converges to the perfect complements case, $h[min(I_1, I_2)]$. 

calculate the “marginal value of public funds” for each of these studies as the ratio of recipients’ willingness to pay by the net cost to the government. A marginal value of public funds greater than 1 translates into a social welfare improvement over a non-distortionary cash transfer, and an infinite marginal value of public funds means that the program is a Pareto improvement that “pays for itself” due to the positive fiscal externality created when earnings increases are large enough to pay back program costs through increased tax revenue. For an introduction to the “marginal value of public funds” framework in this journal, see Finkelstein and Hendren (2020).

Figure 2 – which is reproduced from Hendren and Sprung-Keyser (2020) – plots marginal value of public funds estimates from all 133 policy interventions, sorted by the age of beneficiaries. Perry Preschool and Carolina Abecedarian are included here and have very high marginal value of public funds. But so do other policies aimed at other age groups, such as increased K-12 school spending in the 1970s and 1980s, financial aid for low-income college students, and sectoral employment programs for young adults. While individual estimates are noisy, they pool studies by category and find that child education, child health, and college policies all “pay for themselves” on average through increased future tax receipts. This pattern of results by age group is not an artifact of their welfare analysis framework, because many of the studies they cite have high returns when measured with more conventional approaches such as benefit-cost ratios or internal rates of return (Rea and Burton 2020).

In sum, the evidence suggests that human capital investments are, at least in rough terms, equally productive between the ages of 0 and 25. The key distinction is not age per se, but rather a focus on human capital. Skill investments improve outcomes for adult recipients, including higher income but also improvements in health and other benefits. However, higher income later in life benefits society as well as participants themselves, because the resulting increase in tax revenue lowers the long-run fiscal cost of a program. In contrast, programs for adults such as housing vouchers or disability insurance tend to reduce labor earnings, which pushes the marginal value of public funds below 1. To be clear, this does not mean that the policies are a bad idea, just that supporting such policies requires placing a higher welfare weight, in a given year, on beneficiaries than on the average taxpayer.

What are some reasons why skill investments would have similar returns throughout the life course? First, perhaps even young children are already on the flat of the Heckman Curve. Exposure to disease, pollution, and other adverse events have temporary impacts on adults and young children but long-lasting and permanent impacts on fetuses (for example, Almond, Currie, and Duque 2018). Perhaps sensitivity to early investments might be most important before a child is born, with the difference between early and late childhood being less important. Second, human capital investments during school-age may be more productive on the margin because schools are an efficient delivery mode for interventions.\(^8\) High fixed costs require interventions to be very intensive, whereas the marginal dollar can be spent more on increasing dosage when fixed costs are low. School spending and financial aid exploit the fact that the fixed costs have already been paid.

Third, while the Heckman Curve arises from an economic theory about the technological possibilities for human capital investment, the real world is much messier. Due to lack of opportunity and various market failures such as credit constraints and imperfect information, few people reach their full

\(^8\) I thank Todd Rogers for a helpful conversation that crystallized this point.
potential. If almost everyone is inside their own skill frontier, the Heckman Curve may not apply in practice even if it exists in principle.

**Fact 3: The technology for producing foundational skills such as numeracy and literacy is well-understood, and resources are the main constraint.**

Human capital investments can be productive at many stages of life. Yet not all interventions increase human capital, and what works in one setting may not work in others. *How should we invest in human capital?*

The available options can be divided into three main categories. First, we can allow schools to improve input quality by relaxing their financial constraints. With extra money, schools might buy smaller classes, higher-quality teachers, additional tutors, or other inputs. Second, we can change the investment decisions of individuals, families or schools by lowering the relative price (perhaps to zero) of specific inputs like tutoring or technology, or by increasing incentives for specific inputs like coming to class or reading books. Third, we can encourage students, teachers and other actors to expend more effort toward human capital investment through performance incentives.

There has been an ongoing controversy over which of these approaches holds greatest promise. In a highly influential review called “The Failure of Input-Based Schooling Policies,” Hanushek (2003) argues that investments such as lowering class size or increasing teacher pay do not work because schools do not use resources efficiently. He advocates instead for greater performance incentives to increase teacher and school effort. However, a wave of recent research reviewed in Jackson (2020) has concluded that additional resources do improve education outcomes. The discrepancy between these two findings arises partly from timing, but also from a debate about research design and methodology. Hanushek’s argument was based on time series and cross-sectional differences in education spending, and how it was difficult to find positive correlations in this data between spending and educational outcomes. The newer research is based on quasi-experimental evidence from school finance reforms.

As one example, Jackson, Johnson, and Persico (2016) use court-ordered changes in state K-12 funding formulas to predict a local district’s reform-induced expected change in per-pupil spending on class size, instructional time, and teacher quality, and then study the impact of these funding changes on student outcomes. They find that a 10 percent increase in per-pupil spending over 12 years of public schooling increases educational attainment by 0.3 years and adult wages by 7 percent. Jackson and Mackevicius (2021) conduct a meta-analysis of 31 quasi-experimental studies relating US public K-12 school spending to student outcomes. They estimate that a $1,000 increase in per-pupil spending over four years increases test scores by 0.035 standard deviations and increases college-going by 2.6 percentage points.

Additional school spending also boosts human capital in developing countries. Duflo (2001) finds that a large school construction program in Indonesia increased educational attainment and earnings. Khanna (2021) shows that school districts in India that received extra resources because of a funding formula discontinuity built more schools, hired more teachers, and improved existing facilities. Students in these districts completed about 0.7 more years of schooling and earned between 11 and 14 percent more as adults. However, other studies found no effect of increased school spending: De Ree et al. (2018) find no impact of a large increase in teacher salaries in India, and Mbiti et al (2019) find no impact of unconditional grants to schools in Indonesia.
One partial explanation of these varying results is that school finance reforms operate as “helicopter drops” of additional resources. While there is some heterogeneity in how the money is targeted or the characteristics of the student population, schools mostly seem to use extra resources to do more of what they were already doing. In the United States, at least, “more of the same” can be a good investment on the margin. Hendren and Sprung-Keyser (2020) calculate an infinite marginal value of public spending for the Jackson, Johnson and Persico (2016) results, suggesting that increased school spending caused by court decisions in the 1970s and 1980s “paid for itself” through increased future tax revenues. However, unconditional resource increases have often worked less well in developing countries: for example, the World Bank (2018) began its World Development Report with the sentence: “Schooling is not the same as learning.” The report discusses a common pattern across many countries, where resources devoted to getting children into classrooms have not been followed by a commensurate increase in academic achievement.

Which specific input investments reliably increase human capital? In the context of developed economies, Fryer (2017) reviews nearly 200 randomized educational interventions and finds wide and sometimes unpredictable variation in “what works.” Experiments that lower poverty, change parenting practices or alter the home environment have no average impact on academic outcomes. Early childhood and school-based interventions are effective on average, but with substantial heterogeneity. High-dosage tutoring sometimes increases math and reading achievement, but low-dosage tutoring does not (for example, Banerjee at al. 2007; Fryer and Howard-Noveck 2020). Teacher training and professional development sometimes increases achievement, and sometimes does not (for example, Borman et al. 2007, Loyalka et al. 2019). Providing computers at home or in school generally has no impact on measured human capital (e.g. Malamud and Pop-Eleches 2011, Cristia et al. 2017). Computer-assisted learning software has mixed impacts on achievement, with some evidence that technology-assisted personalization increases student achievement (Bulman and Fairlie 2016, Muralidharan, Singh and Ganimian 2019).

The overall picture is that some specific input investments work very well, but many do not, and it is often hard to predict ahead of time. An Institution of Education Sciences (2013) summary of randomized evidence on targeted school input interventions found that 11 out of 90 interventions produced positive and statistically significant impacts on achievement. Moreover, interventions that are effective in one context may not scale up or generalize to other settings. Kerwin and Thornton (2021) show that while the full-service version of a literacy program in Uganda boosted reading and writing skills, a lower-cost version implemented in the same context resulted in lower test scores. Beg et al. (2019) find that an educational technology intervention in a group of middle schools in Pakistan worked when teachers were trained ahead of time on how to use it in the classroom, but harmed learning when it was delivered directly to students.

Finally, the evidence on incentives is mixed. In experiments that included 250 urban schools in five US cities, Allan and Fryer (2011) show that paying students for performance directly rarely works. The largest randomized studies of teacher incentives in the United States find no impact on student achievement or other outcomes (Springer et al. 2012, Fryer 2013). Teacher incentives sometimes have positive impacts in developing countries, with larger impacts on tested vs. non-tested subjects (for example, Muralidharan and Sundararaman 2011; Filmer et al. 2020).
A large body of research studies school incentives in the form of test-based accountability, which threatens low-performing schools with sanctions such as failing grades, dismissal of teachers and principals, and school closure. This form of accountability pressure leads to large gains on high-stakes tests, modest and inconsistent gains on low-stakes tests, and a variety of harmful strategic responses (for example, Jacob and Levitt 2003, Neal and Schanzenbach 2010, Dee and Jacob 2011). In terms of long-run impacts, Deming et al (2016) find that accountability pressure in Texas increased college attendance and earnings for students in the lowest-performing schools but had negative long-run impacts for low-scoring students in higher-rated schools due to strategic classification of students as eligible for special education.

When incentives work, they often do so by diverting effort toward the narrow goal of meeting a performance target in ways that can create harmful side effects. This “multi-tasking” problem is well-known in the economics literature (Holmstrom and Milgrom 1991). Incentives increase pressure on students, teachers and schools to meet short-run targets, when the actual goal is long-run human capital development. In some form, this tradeoff may be unavoidable. Even though education is a long-run process, incentives usually work better for immediate and easy-to-verify metrics like attendance, enrollment, reading books and completing quizzes.

Some promising evidence suggests that resources and incentives can work well together. In a large-scale experiment in Tanzania, Mbiti et al. (2019) find that teacher incentives and unconditional cash grants to schools have little impact individually, but large impacts on achievement when implemented together. Andrabi et al. (2020) find that unconditional cash grants to private schools in Pakistan increased test scores, but only in villages where the grants were given to all schools rather than only one.

The combination of resources and incentives also works for some high-performing U.S. charter schools that follow a “no excuses” approach with an emphasis on rules of comportment, longer school days, and extra instructional time. These schools are publicly funded but receive significant additional private funding. Charter schools face higher levels of external accountability because they can be shut down more easily.

I interpret the evidence as follows. First, at least in the United States, increased school spending is productive on the margin. Increasing school spending from current levels would produce substantially more human capital and may even pay for itself in the long run. The technology for producing basic math and literacy skills in school-aged children is fairly well-understood. Smaller class sizes, better school facilities, and more instructional time all have reliable impacts on the development of foundational academic skills. The inputs with the best track record of effectiveness – high-dosage tutoring, extra instructional time, personalization and teaching to the right level – mostly deliver to students “more of the same”, rather than reinventing the learning process.

---

9 Another way to increase schools’ effort is through competition between traditional public schools and charter and/or private schools. Figlio, Hart and Karbownik (2020) find that the introduction of private school vouchers in Florida modestly increased achievement among students attending nearby public schools due to competitive pressure. However, in a large market-level experiment in India, Muralidharan and Sundararaman (2015) find no impact of private school competition on the achievement of public school students.
Second, while sharpening incentives works in some contexts, achievement gains are often short-lived and there is not much evidence of long-run benefits. An important caveat is that resources plus incentives appear to be more effective than resources alone.

Third, simply giving schools money - and allowing them to spend it flexibly - may be a more reliable way to increase human capital than pinning our collective hopes on any particular “silver bullet” approach that all schools would be required to follow. This makes education experts queasy, and rightfully so. In a perfect world, increases in resources are combined with transparency and accountability for results. Yet the evidence suggests that “helicopter drops” of money are spent well enough to be worth the investment, at least in developed countries and in schools with strong internal accountability.

However, just because school spending is economically productive on the margin does not mean that the money is spent optimally. It can be simultaneously true that school spending is productive and that much of it is wasted. We can probably always do better, and so innovation and experimentation are critical for increasing the productivity of human capital investments.

**Fact 4:** Higher-order skills such as problem-solving and teamwork are increasingly economically valuable, and the technology for producing them is not well understood.

Schools have a long and successful track record of teaching children how to read, write, and do arithmetic. But a good school does much more. The long-run impacts of educational interventions are often much larger than what would be predicted by achievement gains alone. A growing body of work emphasizes the importance of “non-cognitive” or “soft” skills like patience, self-control, conscientiousness, teamwork, and critical thinking. While such skills are clearly important, the very terms “soft” and “non-cognitive” reveal our lack of understanding about what these skills are and how to measure or develop them.

In my view, the appropriate term for capacities like problem-solving, critical thinking and teamwork is higher-order skills, following Bloom’s (1956) taxonomy of educational objectives. Bloom’s taxonomy establishes a hierarchy with factual knowledge as the base of the pyramid, followed by pattern recognition and classification, up to higher-order objectives such as application to new situations, experimentation and making connection to new ideas, evaluation and decision-making, and design and creation of new concepts. Tests like the SAT or the Armed Forces Qualification Test (AFQT) focus on the bottom two layers of the pyramid: recalling, explaining, and understanding ideas and concepts. As discussed above, we know a great deal about how to build these foundational skills: indeed, as the pyramid structure of the taxonomy implies, they are a precondition for developing higher-order skills such as applying conceptual knowledge to solve new problems and evaluating evidence from multiple sources to improve decision-making.

Despite our lack of understanding of how higher-order skills are measured and developed, a variety of studies have found ways to use the existing evidence do demonstrate their importance for life success. For example, Jackson et al. (2020) use survey evidence from ninth-grade students in Chicago public schools and find that schools with high “value-added” in promoting hard work and social well-being increase students’ high school graduation and college attendance, even after accounting for their impact on academic achievement. Using data from the population of Swedish military enlistees, Lindqvist and Vestman (2011) estimate high labor market returns to both cognitive and non-cognitive skills, where the latter is measured using scores from a personal interview administered by a trained psychologist.
Deming (2017) shows that the economic return to social skills in the United States more than doubled for a cohort of youth entering the labor market in the 2000s compared to the 1980s. In that study, discussed further below, I measure social skills by creating an index based on four factors: self-reported sociability; self-reported sociability at age six, as perceived by the adult respondent; number of clubs in which the respondent participated in high school; and participation in high school sports. Edin et al (2021) find similar returns to social skills in Sweden, using administrative data from the compulsory military draft that required men aged 18 or 19 to be tested on cognitive and non-cognitive skills. Attanasio et al. (2020) find growing inequality in socio-emotional skills across two British cohorts born 30 years apart, using survey tools filled out by mothers (or in some cases teachers) about behaviors of their children. Each of these studies measures “non-cognitive” skills using whatever measures are at hand, rather than relating them conceptually to particular higher-order skills.

Higher-order skills clearly seem important, yet measuring them well is a challenge. The typical approach uses self-reported questionnaires, which are often Likert scale items (1 to 5 or 1 to 7, ranging from “strongly disagree” to “strongly agree”) without any cardinal meaning. Their predictive power for different life outcomes varies widely depending on the exact measure, the outcomes used, and the social context.

One problem is that questionnaires can suffer from reference bias, where respondents make relative comparisons to those around them. West et al. (2016) find that students who win a lottery to attend “No Excuses” charter schools subsequently score lower on measures of conscientiousness and grit, despite having higher achievement and attending a school with longer hours and more homework, because they are now evaluating themselves in the context of a different set of institutional expectations. Some studies measure non-cognitive skills using behavior such as absences and school suspensions. Yet such behaviors capture not only skills but also social context, including racial discrimination, school context, and other unknown factors.

Conceptual clarity is a second challenge to our understanding of higher-order skills. In a standard human capital framework, more skills are always better. But certain skills may be effective in some contexts and counterproductive in others: as one example, conscientiousness positively predicts educational attainment and earnings, yet disruptive and aggressive behavior (the opposite of conscientiousness) sometimes predicts earnings and entrepreneurial success (Levine and Rubinstein 2017; Papageorge, Ronda, and Zheng 2019).

We need a systematic research program that seeks to understand the economic importance of higher-order skills. This research program would combine careful measurement and development of theory with experimentation and impact analysis using strong research designs. In the rest of this section, I illustrate the value of this approach by attempting to synthesize what we know (and what we do not know) about interpersonal and intrapersonal skills.

**Interpersonal Skills and the Science of Teamwork**

A survey of workers, employers and experts administered by the US Department of Labor found that teamwork is a “very” or “extremely” important job feature in 78 percent of all jobs (O*NET, 2022). A long literature in economics treats teamwork as a tradeoff between the benefits of increased productivity through specialization and the costs of coordination (Becker and Murphy 1992; Garicano and Rossi-Hansberg 2006). In this context, the rise of team production is a response to the increasing
complexity of work, and a well-functioning team can exploit comparative advantage between team members to increase productive efficiency.

Deming (2017) shows how the rise of teamwork has increased the value of social skills in the labor market. Between 1980 and 2012, jobs requiring high levels of social interaction grew by nearly 12 percentage points as a share of the US labor force, and the labor market return to social skills more than doubled. Deming explains these empirical results with a model of team production where social skills reduce the coordination costs of “trading tasks” between workers on a team, allowing them to exploit comparative advantage more fully. Several other recent papers show evidence of the economic value of social skills. For example, Hansen et al. (2021) use data on job descriptions for top executives to “show an increasing relevance of social skills in top managerial occupations.” Hoffman and Tadelis (2021) look at employee surveys within a single large firm to show that manager with better “people management skills” reduce attrition among those working for them and are better-compensated by the firm.

We know that social skills are rewarded in the labor market, and we know that teamwork is increasingly important. But can we draw a direct connection between social skills, teamwork, and increased productivity?

Weidmann and Deming (2021) develop a novel experimental method for identifying individual contributions to group performance. We first measure individuals’ productivity on a series of problem-solving tasks, and then randomly assign the same individuals to multiple teams, which perform group analogs of the same tasks. We use the individual scores to generate a performance prediction for each team, and then estimate a “team player” effect by combining the residual from the prediction across multiple random assignments of individuals to groups. People who consistently cause their teams to outperform its prediction are team players.

Weidmann and Deming (2021) find that individuals have persistent impacts on group performance – in other words, that team player effects exist. In addition, these effects are positively correlated with a commonly used and psychometrically validated measure of social intelligence called the Reading the Mind in the Eyes Test (Baron-Cohen et al 2001). The test presents participants with photos of faces, cropped so that only the eyes are visible. For each set of eyes, participants are asked to choose which of four emotions best describes the person in the image. This test measures the ability of participants to recognize emotions in others and, more broadly, the ability to reason about the mental state of others. Relative to other measures of social intelligence, the main value of the Reading the Mind in the Eyes is that it has right and wrong answers, has relatively high test-retest reliability, and can be quickly and reliably administered (Pinkham et al. 2014). Lab participants were also assessed on a standard measure of IQ and on three dimensions of the well-known “Big 5” personality inventory that are positively associated with group performance in other studies: Conscientiousness, Extraversion, and Agreeableness. None of these measures were correlated with the team player effect – only the Reading the Mind in the Eyes Test. In short, this experiment uses a lab setting to develop a clean test of the underlying economic mechanism relating social skills to team productivity.

---

10 The “memory test” involving words, imaged, and stories. In the “optimization test,” participants made a series of guesses between 0 and 300, observed how these guesses were translated by a complex unobserved function into final values, and then estimated the highest value for the function. In the “shapes test,” participants observed a series of shapes and the predicted what element was missing in the next shape of the series.
The importance of teamwork skills is also corroborated in a variety of field settings. Arcidiacono et al (2017) use data from US professional basketball to show how individual performance depends on peer effects. Devereux (2021) look at data on co-authorship of academic papers for economists within the University of California system; they find that the importance of an author’s value-added as a co-author is more closely tied to salary than the author’s own value-added. Bonhomme (2021) develops an econometric framework to estimate the impact of individual workers on team performance that allows for variation in teamwork skills, worker sorting, and complementarity, and estimates the framework using the research output of economists, and contributions of inventors to patent quality. However, much more work is needed to understand how social skills matter and under what conditions.

Another largely unexplored frontier is the development of social skills. There are a few studies, often in developing economies, looking at how a specific social skills program improved outcomes. With female workers in the garment industry in India, Adhvaryu, Kala and Nyshadham (2021) find that on-the-job “soft skills”, with a focus on communication, time management, financial literacy, successful task execution, and problem-solving, increases employee productivity by 13.5 percent, with larger impacts when work is more team-intensive and evidence of spillovers to untreated teammates. In a business training program in Togo, social skills training programs improve entrepreneurs’ ability to form business connections (Dimitriadis and Koning 2020). In a three-week business skills program for high school students in Uganda, Chioda et al (2021) find that teaching soft skills like self-efficacy, persuasion, and negotiation led to greater gains in earnings than a focus on hard skills. In an educational setting in Zambia, Ashraf et al (2020) find that a training program in negotiation skills for adolescent girls improved educational outcomes. We need more research and varying programs and contexts to build an economic theory of how and why teamwork skills matter.

Intrapersonal Skills and Economic Decision-Making

Good decision-making requires counterfactual reasoning, meaning a consideration of alternative actions and their likely consequences. In terms of Bloom’s (1956) taxonomy, this process requires the combination of several higher-order skills such acquiring information, applying information to new situations, testing and evaluating evidence, and making and justifying decisions.

A wide literature has considered various and overlapping aspects of decision-making: patience, self-control, grit to persevere through difficult tasks, habits of acquiring additional information or considering alternative strategies, developing cognitive shortcuts that will make sense in a variety of settings, and others. The economics literature on time and risk preference has mostly focused on broad cultural and familial influences for these decisions, or on behavioral biases and attributes of the choice environment. Yet in my view, decision-making errors and biases can be reinterpreted as arising from deficits in higher-order skills. This in turn suggests the need for more research on the “skill” of decision-making itself.

For example, patience and self-control are linked: patience is a willingness to think long-term, whereas self-control is the ability to overcome the temptations of the present (Fudenberg and Levine 2006). One approach in this literature asks the subject to comparing a tradeoff between now and the future to two equivalently spaced times in the future: for example, the offer of $100 today versus $110 tomorrow is accepted much more often than the offer of $100 seven days from now vs. $110 eight days from now. Waiting an extra day seems easier in the future than in the present. Self-control is positively related to academic achievement and other life outcomes: as one example, Duckworth et al (2019) survey a range
of evidence that measures of self-control are linked to academic achievement. Related to self-control is “grit”, or the ability to persevere through effortful, sometimes unpleasant tasks to achieve a desired long-run outcome.

There is promising evidence that self-control and grit are malleable and can be improved through low-cost, scalable investments. Alan and Ertac (2018) show that a school-based enrichment program for third and fourth-graders in Turkey that encourages forward-looking behavior and imagining the future increases behavior grades and patience up to three years later. Lührmann et al (2018) show that a financial literacy program for German high school students reduced time inconsistency and lowered discount rates.

What is the best way to improve grit? One view is that grit is developed through the adoption of a “growth mindset” (for example, Yeager and Dweck 2012), which refers to changes in beliefs about the returns to effort. An alternative view treats grit as cognitive endurance, the skill of maintaining focus over time (Brown et al 2021). Interventions focusing solely on growth mindset have shown mixed results, yet pairing mindset training with structured support for goal-setting and deliberate practice yields promising increases in academic performance in a nationally representative group of US high school students, as well as in 52 state-run elementary schools in Turkey (Yeager et al 2019, Alan, Boneva, and Ertac 2019). Brown et al (2021) randomly assign a group of 1,600 Indian primary school students to two types of effortful cognitive activity – one that is clearly academic, and one that is not. Both interventions increase the ability to concentrate and both lead to increased academic performance, suggesting that cognitive endurance improves with practice. This reframes “grit” as a skill that can be developed rather than as a mindset to be shifted.

Assessing alternative future states of the world is cognitively taxing, which may help explain the correlation between intelligence and patience (for example, Dohmen et al 2018). One route to better decision-making is to develop cognitive shortcuts and habits of mind that make long-run thinking easier. The success of cognitive behavioral therapy in reducing violence and other negative behaviors offers an intriguing example. Cognitive behavioral therapy focuses on decision-making directly by asking people to slow down and evaluate the consequences of their behavior patterns, and then reprogramming new behaviors through deliberate practice. Heller et al (2017) find large reductions in violent crime and increases in high school graduation from cognitive behavioral therapy interventions with high-risk young men in Chicago. Blattman, Jamison, and Sheridan (2017) find that a combination of cognitive behavioral therapy and cash grants greatly reduced crime and violence among criminally engaged men in Liberia. A decision-making intervention for young children in Switzerland called Promoting Alternative Thinking Strategies (PATHS) increased high school graduation and college attendance (Sorrenti et al 2020).

Finally, an emerging body of evidence suggests that strategic sophistication improves decision-making. Fe, Gill and Prowse (2020) worked with data from the Avon Longitudinal Study of Parents and Children (ALSPAC), which measured the theory-of-mind ability and cognitive ability at age eight of children from

---

11 Heller et al (2017) give the example of an exercise where students are paired up and one is given a ball. The other is given 30 seconds to try to get the ball from his partner. After 30 seconds of physical struggle, the group leader asks whether anyone decided instead to simply ask for the ball. When they say “no,” the leader then asks the ball holder whether they would have given it up, to which the typical answer is “I would have given it; it’s just a stupid ball.”
the Avon region in the Southwest of England, and found that “theory of mind” – the ability to attribute mental states to others – predicts strategic sophistication in children and is positively correlated with adult social skills and educational attainment. Gill and Prowse (2016) find that more intelligent US college students converge to Nash equilibrium faster and engage in more sophisticated level reasoning in a series of laboratory experiments.

Future studies should seek to develop theory and measurement paradigms that allow for a direct assessment of the skills and knowledge that are required to improve decision-making. One promising approach is to build on the “rational inattention” literature, which identifies the conditions under which decision mistakes are optimal given the costs of paying attention (for example, Sims 2003, Mackowiak, Matějka and Wiederholt 2021). Many of the biases and rules-of-thumb phenomena identified by behavioral economics can be rationalized by models of costly information acquisition. Viewed in this light, interventions that build the “skill” of lowering attention costs will manifest as a reduction in decision errors and an increase in patience, grit, and other higher order skills. When it comes to understanding the role of skills in improving economic decision-making, there are more questions than answers, which suggests many fruitful and exciting avenues for future work.
References


19


Table 1 - Returns to Education in the NLSY79

<table>
<thead>
<tr>
<th></th>
<th>Average Log Hourly Wages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Years of Education</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td>[0.002]</td>
</tr>
<tr>
<td>AFQT (standardized)</td>
<td>0.161</td>
</tr>
<tr>
<td></td>
<td>[0.006]</td>
</tr>
<tr>
<td>High School Graduate</td>
<td>0.127</td>
</tr>
<tr>
<td></td>
<td>[0.013]</td>
</tr>
<tr>
<td>Some College</td>
<td>0.247</td>
</tr>
<tr>
<td></td>
<td>[0.016]</td>
</tr>
<tr>
<td>Bachelor's Degree</td>
<td>0.479</td>
</tr>
<tr>
<td></td>
<td>[0.019]</td>
</tr>
<tr>
<td>Graduate Degree</td>
<td>0.535</td>
</tr>
<tr>
<td></td>
<td>[0.021]</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.296</td>
</tr>
<tr>
<td>Sample Size</td>
<td>10,876</td>
</tr>
</tbody>
</table>

Notes: Estimates are from a regression of inflation-adjusted average log hourly wages, measured between the ages of 25 and 54 using repeated observations of individuals in panel data from the 1979 National Longitudinal Survey of Youth (NLSY79). In columns 1 and 2, years of education is a continuous measure that is bounded below at 11 and above at 20. Column 3 shows results by level of attainment, where less that high school is the left-out category. The regression also includes indicators for race and gender. AFQT is the Armed Forces Qualifying Test, a measure of aptitude administered prior to labor market entry and standardized to have mean zero and standard deviation one. The average wage of respondents is $18.95 in 2016 dollars.
This figure is a reproduction of the “Heckman Curve”, from Figure 2 of Heckman (2006).
This figure is a reproduction of Figure IV.A in Hendren and Sprung-Keyser (2020), which computes 133 estimates of the Marginal Value of Public Funds (MVPF) for experimental and quasi-experimental interventions, grouped by the age of beneficiaries. An infinite MVPF indicates policies that “pay for themselves” through fiscal externalities such as increased future tax receipts.