INSTITUTE for RESEARCH on POVERTY
UNIVERSITY OF WISCONSIN–MADISON

Focus

Vol. 29, No. 2, Fall/Winter 2012–13 ISSN: 0195–5705

Hard evidence on soft skills 3

Year Up: Providing a pathway from poverty to a professional career for urban young adults 9

How does Year Up measure up? 13

Response from the author 18

Value-added measures of teachers: Research and policy 21

The long-term effects of teachers 21

Effects of value-added policies 23

Use of value added in teacher policy measures 25

I am very happy to introduce this issue of Focus*, which is centered on the theme of building human capital for youth from disadvantaged backgrounds, one of IRP's major research themes for the next two years (see http://www.irp.wisc.edu/research/humancapital.htm). Now almost four years after the end of the Great Recession, hiring and job quality are both far from a full recovery, especially for low-skill younger workers. Because we believe that the answer to poverty for non-disabled adults and their children lies in a good steady job with decent wages and long-term prospects for advancement, we are especially interested in how to improve job market outcomes for poor youth and adults.*

We lead off with a summary of the Robert J. Lampman Memorial Lecture given by James Heckman last spring. James discusses personality psychology, a new point of contact between the fields of economics and psychology. He argues that individual personality traits, sometimes called "soft skills," are predictors of success in many areas of economic and social life, including earnings outcomes. Individual variation in these skills is an important source of inequality. James details the ways that economists can both take from and contribute to personality psychology in order to better examine and address the sources of poverty.

Next is an article by Gerald Chertavian, which provides a real-life example of the importance of soft skills. The piece is based on the annual IRP New Perspectives in Social Policy Seminar delivered by Gerald last fall, with Carolyn Heinrich serving as the respondent. His is the first in a set of articles about Year Up, a one-year intensive training program that he founded. Year Up provides low-income young adults with a combination of technical and professional (or soft) skill development, college credits, and corporate internships. He describes the success of the program to date, and strategies to improve, expand, and replicate its results across the nation.

The Year Up article is followed by a summary of the seminar response by Carolyn Heinrich, co-organizer of IRP's Building Human Capital and Economic Potential project. Carolyn applauds the program's success, but also cautions about the difficulties of going to scale and replicating results in different locations. Gerald in return makes a constructive reply to Carolyn's remarks.

We close the issue with a summary of a wonderful panel on the value added of good teaching organized by IRP Affiliate Robert Moffitt as part of the IRP 2012 Summer Research Workshop. In their presentations, Raj Chetty, Eric Hanushek, and Jesse Rothstein reviewed research, policy, and practice in using value-added measures to improve educational outcomes, and perhaps labor market outcomes, especially for disadvantaged children.

Another resource for those interested in reducing the effects of poverty on the next generation is IRP Associate Director Katherine Magnuson's recent webinar on early childhood interventions for low-income children, in which she summarizes the state of research and practice in this crucial policy arena. Rather than

Do we have your e-mail address?

This is the last print edition of *Focus*. We are switching to electronic distribution.

Please see the notice on page 27.

including a summary of her webinar in this issue of Focus*, we have posted her presentation on our website so that you can access it directly (visit http://www.irp.wisc.edu/publications/media/webinars.htm or go to our home page and click on the IRP Webinars link in the left-hand column).*

As always, IRP is on the lookout for good opportunities to stimulate research, train young researchers, analyze policy, and improve program performance and practice to better the lot of the poor. We welcome any ideas you may have that will help us to achieve these goals.

—Timothy M. Smeeding, IRP Director

Digital Media Resources at IRP

Webinar Series
IRP has established a new series of interactive online seminars designed to more broadly disseminate current research on poverty and social inequality in the United States. http://www.irp.wisc.edu/publications/media/webinars.htm

IRP Podcasts
Also new, IRP offers brief discussions of poverty issues for nonspecialist audiences, posted monthly. http://www.irp.wisc.edu/publications/media/podcasts.htm

Video Archives
Videos of past IRP events are now archived online. http://www.irp.wisc.edu/newsevents/videos.htm

Social Media
IRP is on Facebook (https://www.facebook.com/irpwisc) and on Twitter (https://twitter.com/irp_uw).

Poverty Dispatch
IRP also compiles and emails weekly poverty-related news links. http://www.irp.wisc.edu/dispatch/

FOCUS is a Newsletter put out twice a year by the

Institute for Research on Poverty
1180 Observatory Drive
3412 Social Science Building
University of Wisconsin
Madison, Wisconsin 53706
(608) 262-6358
Fax (608) 265-3119

The purpose of *Focus* is to provide coverage of poverty-related research, events, and issues, and to acquaint a large audience with the work of the Institute by means of short essays on selected pieces of research. Full texts of Discussion Papers and Special Reports are available on the IRP Web site.

Focus is free of charge, although contributions to the UW Foundation–IRP Fund sent to the above address in support of *Focus* are encouraged.

Edited by Emma Caspar

This publication was supported by Grant Number AE00102 from the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation (ASPE), and awarded by the Substance Abuse and Mental Health Services Administration (SAMHSA). Its contents are solely the responsibility of the author(s) and do not necessarily represent the official views of ASPE or SAMHSA.

Hard evidence on soft skills

James J. Heckman

James J. Heckman is Henry Schultz Distinguished Service Professor of Economics at the University of Chicago. He delivered the annual Robert J. Lampman Memorial Lecture at Madison in May 2012. This article is adapted from his lecture.

The ascent of behavioral economics produced a union of economics with cognitive psychology and neuroscience. This article discusses a new point of contact between economics and psychology: personality psychology.[1] Personality psychology gives very rich descriptions of individual differences in traits and outcomes. Personality traits are important predictors of success in many areas of economic and social life. Individual variation in these personality traits—sometimes called "soft skills" or "character skills"—is an important source of inequality. Personality traits can be changed by intervention, and interventions that target personality are promising.

What can economists take from and contribute to personality psychology?

Measures of personality are informative descriptions of human differences that supplement those offered by the standard preference parameters used in economics, such as risk aversion or preference for leisure. Economics can contribute to the field of personality psychology in defining traits, in distinguishing measurements of traits from other traits, and in determining the causal effect of traits on outcomes. And in turn, lessons from personality psychology can help economists identify hitherto unidentified sources of inequality.

Psychological measurement systems

I begin with a brief description of psychological measurement systems. Measurement of cognition and educational attainment has been refined during the last century. Personality psychologists have constructed measures of personality traits and have shown that these traits predict many important life outcomes.

Cognitive ability

Modern intelligence tests have been used since 1904, when the first IQ test was created.[2] The standardized achievement test was created in the wake of IQ tests as an objective and cost-effective measure of acquired skills. In contrast to IQ tests, thought to measure a fixed trait, standardized achievement tests were designed to measure skills that could be acquired in school and through life experience and that were widely applicable beyond the classroom.[3]

Personality traits

IQ and achievement tests measure skills that have payoffs in labor markets. They do not measure personality traits that are also valued in the labor market, in school, and elsewhere in life. Personality psychologists have generally agreed on a taxonomy of traits called the "Big Five," which are Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism—OCEAN for short. Table 1 defines these traits and their facets.

How are psychological measurements validated?

The measurements of traits offered by psychologists should be taken with a grain of salt. Many of the validation studies in psychology have very limited objectives. For example, the validity of the SAT standardized assessment test is often based on how well it correlates with first-year college grades. There is a circular quality to many of these validation studies. For example, the validities of IQ tests are often based on correlations with other intelligence tests or with grades or scores on achievement tests. The validity standards for personality measurements are a bit better, as personality tests are often constructed to predict a wide array of behaviors, and thus may be validated by observing those behaviors.

The predictive power of personality

There are several difficulties involved in synthesizing the evidence of the effect of personality on outcomes. First, measures of personality and cognition differ among studies. Even the Big Five set of traits identified above, while having fairly wide acceptance, is not universally agreed upon. Second, different studies use different notions of the predictive power of the measures. Third, very few studies address the question of causality. That is, does the measured trait cause, rather than simply predict, the outcome of interest?

The existing studies on the power of personality can be summarized briefly as follows: Conscientiousness—the tendency to be hardworking and organized—is the most predictive Big Five trait across a variety of outcomes. For example, a study of correlations of the Big Five and intelligence with college course grades found that Conscientiousness was just as important a predictor as intelligence, as measured by an IQ test.[4] My colleagues and I have consistently found that cognitive and personality traits are equally predictive across a great variety of labor market and social outcomes.[5] The SAT provides an interesting example, as it is so widely used in college admissions decisions. Multiple studies have shown that in comparing the relative predictive power of Conscientiousness and SAT scores for college grades (the very thing that SAT scores were designed to predict), Conscientiousness is just as predictive as the SAT, though neither is very predictive.[6]

Table 1
The Big Five Domains and Their Facets

Big Five Personality Factor	American Psychology Association Dictionary Description	Facets (and Correlated Trait Adjective)	Related Traits	Analogous Childhood Temperament Traits
Openness to Experience	"The tendency to be open to new aesthetic, cultural, or intellectual experiences."	Fantasy (imaginative), Aesthetic (artistic), Feelings (excitable), Actions (wide interests), Ideas (curious), and Values (unconventional)		Sensory sensitivity, Pleasure in low-intensity activities, and Curiosity
Conscientiousness	"The tendency to be organized, responsible, and hardworking."	Competence (efficient), Order (organized), Dutifulness (not careless), Achievement striving (ambitious), Self-discipline (not lazy), and Deliberation (not impulsive)	Grit, Perseverance, Delay of gratification, Impulse control, Achievement striving, Ambition, and Work ethic	Attention/(lack of) distractibility, Effortful control, Impulse control/ delay of gratification, Persistence, and Activity[a]
Extraversion	"An orientation of one's interests and energies toward the outer world of people and things rather than the inner world of subjective experience; characterized by positive affect and sociability."	Warmth (friendly), Gregariousness (sociable), Assertiveness (self-confident), Activity (energetic), Excitement seeking (adventurous), and Positive emotions (enthusiastic)		Surgency, Social dominance, Social vitality, Sensation seeking, Shyness[a], Activity[a], Positive emotionality, and Sociability/affiliation
Agreeableness	"The tendency to act in a cooperative, unselfish manner."	Trust (forgiving), Straightforwardness (not demanding), Altruism (warm), Compliance (not stubborn), Modesty (not show-off), and Tender-mindedness (sympathetic)	Empathy, Perspective taking, Cooperation, and Competitiveness	Irritability[a], Aggressiveness, and Willfulness
Neuroticism / Emotional Stability	Emotional stability is "predictability and consistency in emotional reactions, with absence of rapid mood changes." Neuroticism is "a chronic level of emotional instability and proneness to psychological distress."	Anxiety (worrying), Hostility (irritable), Depression (not contented), Self-consciousness (shy), Impulsiveness (moody), Vulnerability to stress (not self-confident)	Internal vs. External, Locus of control, Core self-evaluation, Self-esteem, Self-efficacy, Optimism, and Axis I psychopathologies (mental disorders) including depression and anxiety disorders	Fearfulness / behavioral inhibition, Shyness[a], Irritability[a], Frustration, (Lack of) soothability, Sadness

Source: Adapted from O. P. John and S. Srivastava, "The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives," in *Handbook of Personality: Theory and Research*, Second Edition, eds. L. A. Pervin and O. P. John (New York: Guilford Press, 1999).

Notes: Trait adjectives in parentheses from the Adjective Check List (H. G. Gough and A. B. Heilbrun, Jr. *The Adjective Check List Manual* (Palo Alto, CA: Consulting Psychologists Press, 1983). [a]These temperament traits may be related to two Big Five factors.

An economic framework for defining and measuring traits

All measurement systems in psychology are based on performance on tasks, such as tests, or in observations of behavior by analysts. An IQ test measures how well the examinee performs on the task of taking the IQ test. A behavioral report on a student filed by a teacher measures how well the student performs on the tasks of being respectful of others and disciplined. Performances on tasks are frequently equated with traits.

It is important to distinguish traits from measurements of traits that are affected by multiple factors. Economics can assist in making this distinction. For example, productivity on tasks that the measurements capture can be modeled as a response function that depends on traits and effort. The behaviors that constitute the measurements of personality are patterns of actions in response to the constraints, endowments, and incentives that individuals face, given their goals and preferences.[7] If incentives and constraints are changed, then the measures will in turn change. These considerations complicate the interpretation of measured traits.

Distinguishing traits from measurements of traits

How can analysts recover traits from measurements of traits? Productivity can be observed in outcomes such as grades, test scores, and accomplishment of tasks. One challenge is to distinguish traits from effort, that is, to standardize for effort. Even if this is possible, it still leaves open the possibility that multiple traits may affect performance on any given task. It is extremely difficult to disentangle the separate roles of individual traits.

An important example illustrating that effort and incentives both affect performance is found in intelligence tests. Two studies in the 1970s offered incentives for performance. In both studies, those in the group that were offered incentives scored higher on intelligence tests than those in the group that was not offered incentives.[8] Thus, an "IQ gap" was created simply by offering incentives. Other studies have shown that levels of traits other than the traits sought to be measured matter. For example, people high in Conscientiousness are already highly motivated, and are much less likely to be influenced by incentives.[9] There is no "pure" measure of IQ. Even IQ scores need to be effort adjusted, and adjusted for personality traits that affect performance on IQ tests.

Causality

Another area where economics can contribute to personality psychology is in establishing a causal relationship between traits and outcomes. There are a number of difficulties in doing this, including the issues just discussed. Parsing the different factors that produce an outcome on a task is very difficult. Added to this problem is the fact that outcomes can be influenced by incentives. This problem is even more of a concern with personality measures than with cognition. Someone taking a personality test for a job is likely to give desirable answers even if they are not true.

These challenges are not insurmountable. I offer two examples of causal evidence of the effect of personality on outcomes—one from the GED testing program and one from a social experiment.

Evidence from the GED testing program

This first study of causality demonstrates the power of personality and the costs of neglecting it. The GED is a standardized achievement test that provides an alternative to a high school diploma. High school dropouts who pass the GED test are certified as high school graduate "equivalents." GED recipients have about the same cognitive ability as high school graduates, but differ in their personality traits. Figure 1 shows that the distributions of ability for GED recipients and high school graduates who do not go on to college are very similar, while the distribution for high school dropouts is very different and shifted to the left.

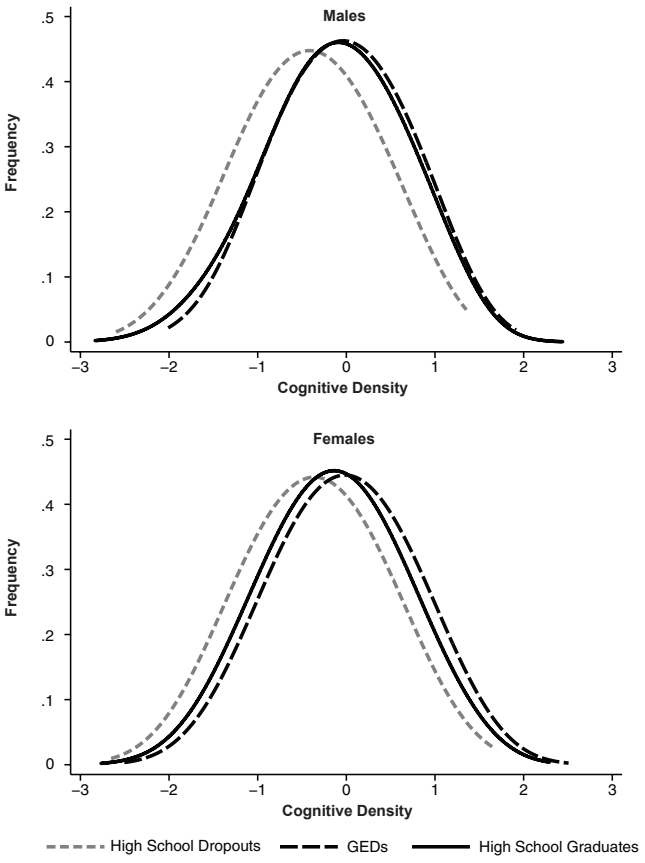

Figure 1. Cognitive ability by educational status.

Source: J. J. Heckman, J. E. Humphries, S. Urzua, and G. Veramendi, "The Effects of Educational Choices on Labor Market, Health, and Social Outcomes," *Journal of Political Economy*, (2013, under revision).

Note: Those who went on to college are not included in the figure.

Examining the distribution of noncognitive skills, on the other hand, tells a very different story. Figure 2 shows that the noncognitive density for GED recipients is very similar to that of high school dropouts, while that of high school graduates is shifted to the right.[10]

While many people with a GED go to college, their college graduation rates are much lower than those of high school graduates. Overall, those with a GED have similar earnings, employment, labor force participation, and hours worked as those who drop out of high school and do not GED certify. Without the ability to randomly assign noncognitive skills, it is difficult to definitively prove causality. However, in a forthcoming book, my coauthors—John Eric Humphries and Tim Kautz—and I estimate thousands of different empirical models on multiple datasets, and consistently find substantial differences in outcomes between those with a GED and those with a regular high school diploma.[11]

For males (the subject of the majority of studies on the effectiveness of the GED), there is no consistent evidence of any difference in outcomes between those with a GED and other high school dropouts. For females, however, there is some hint of a GED "effect." After accounting for differences in cognitive ability, female GED recipients have higher annual earnings than high school dropouts. Although there is no difference in their hourly wages, female GED recipients are more likely than their uncertified high school dropout counterparts to participate in the labor force. However, we cannot be sure of the source of this effect. The estimated GED effect appears to be a selection effect. As a group, GED females have better personality skills than their male GED counterparts. Women more motivated to work may take the GED regardless of any causal effect of the GED on their labor force participation.[12]

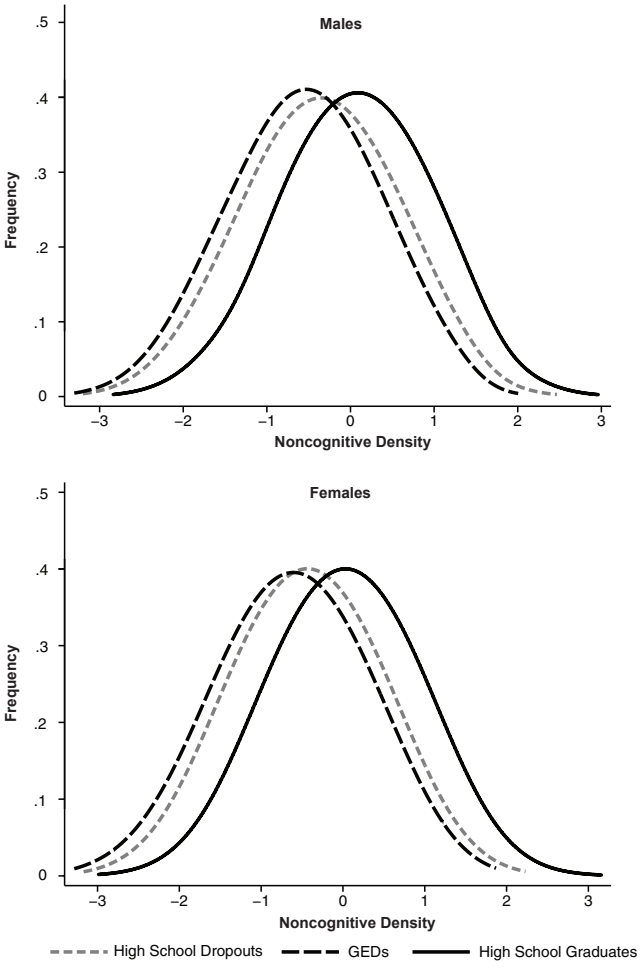

Figure 2. Noncognitive skills by educational status.

Source: J. J. Heckman, J. E. Humphries, S. Urzua, and G. Veramendi, "The Effects of Educational Choices on Labor Market, Health, and Social Outcomes," *Journal of Political Economy*, (2013, under revision).

Note: Those who went on to college are not included in the figure.

Are traits set in stone?

Some psychologists have argued that personality traits are not stable over situations or over time.[13] However, the stability of traits and behaviors before and after GED certification found for most GEDs argues against preference change. There is an evolution of traits as people age, although the mechanisms producing this change are still not well understood. There is, however, evidence that Conscientiousness, previously identified as particularly important among personality traits, increases as people age.[14] Other traits, such as Openness to Experience, appear to decrease with age. Interventions can change traits.

Evidence from a randomized intervention

My second study of causality and personality uses evidence from the Perry Preschool Program to show how personality traits can be changed in ways that produce beneficial lifetime outcomes. This enrichment program, carried out in the 1960s, was targeted to three- and four-year-old low-income black children who had IQs below 85 at age three. Participants were taught social skills, and home visits promoted parent-child interactions. A random-assignment evaluation of this program found no lasting program effects on IQ, leading some critics to dismiss the value of early education.[15] However, the evaluation did identify a variety of improved outcomes for participants with a statistically significant rate of return of 7 to 10 percent per year.[16] The program worked primarily through improving personality traits. Members in the treatment group have better measures of personal behavior, as well as "Externalizing Behavior," a psychological construct related to Agreeableness and Conscientiousness. For girls, the program also improved Openness to Experience. The program improved scores on the California Achievement Test, despite the finding of no program effect on IQ. This result is consistent with the notion that performance on achievement tests, generally regarded as measures of cognition, also reflects personality traits. Achievement tests capture acquired knowledge and higher levels of motivation that lead to greater learning.

Decomposing treatment effects into their experimentally determined sources produces an interesting pattern. Figure 3 shows the results of this analysis for a number of outcomes.

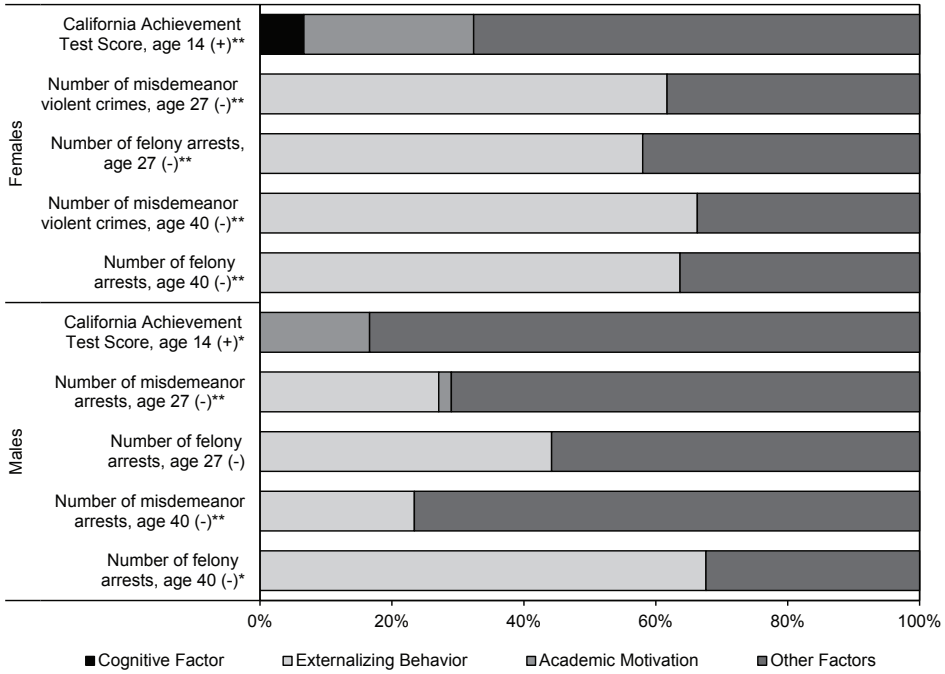

Figure 3. Decompositions of treatment effects on outcomes.

Source: J. J. Heckman, R. Pinto, and P. Savalyev, "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review* (2013, forthcoming).

Notes: Each bar represents the total treatment effect normalized to 100%. "(+)" and "(-)" indicate positive and negative total treatment effects. Statistical significance levels of decompositions are indicated as ** = 5 percent; * = 10 percent.

All reported treatment effects are statistically significant even after adjusting for the effect of multiple-hypothesis testing. The figure shows the proportion of the treatment effect displayed due to the indicated factors. Boosts in Externalizing Behavior play a major role in producing the Perry treatment effects. This evidence suggests that a policy that expands exposure to preschool can make a difference, not only in improving scores on achievement tests, but also in improving noncognitive outcomes—personality skills—where gains persist even into adulthood.

Conclusions

Economists can learn from and contribute to personality psychology. Measures of personality predict many behaviors, sometimes with the same strength as conventional measures of cognitive traits and sometimes even more strongly, as in the Perry Preschool effects of treatment-induced Externalizing Behavior on adult outcomes. Using personality traits augments our ability to predict behaviors.

Personality psychology considers a wider array of actions than are usually considered by economists. Drawing on the lessons of personality psychology enlarges the economist's ability to describe and model the world and understand the sources of poverty. Personality measures explain some of the variation in outcomes that produce inequality, though there is still much to be learned. Understanding personality helps us understand the nature of the tests that are used to monitor schools and societies. For example, tests intended to measure cognition, such as those used for No Child Left Behind, also partly measure motivation and effort.

Personality traits persist across situations. They are not set in stone, but change in stable ways over the life cycle. They are a possible avenue for effective interventions and wise public policy.

Economists can contribute to personality psychology by providing the precise models that personality psychologists lack. Economics provides a framework for recasting the field and collecting incentive-adjusted measures of personality, achievement, and IQ.

Economics is now playing a role in clarifying the concepts and empirical content of psychology. More precise models reveal basic identification problems that plague the measurements collected in psychology and warn economists not to use uncritically the measures developed by psychologists. The next wave of personality measures will incorporate and be improved by this research.

Economics can also be used to interpret the correlations reported in personality psychology. Many contemporaneously measured relationships between traits and behaviors are plagued by the problem of reverse causality. Economists can apply their tools to define and estimate causal mechanisms and to thus understand the causes of effects and not just the correlations among variables.

Future research

Future research in economics and psychology holds both challenges and research opportunities. Personality parameters and economic preference parameters do not correspond very closely.[17] However, the research required to control for the major confounding factors that determine psychological measurements has just started. It is possible that a tighter connection will emerge. More work needs to be done in developing rigorous methods for analyzing causal relationships in both fields. Since important policy decisions are being made based on findings reported in psychology, it is important to strive to establish which empirical relationships are causal.

It will also be necessary to develop a common language and framework to promote exchange between economics and personality psychology. Economists must be careful not to assume that basic questions of content and identification have been answered by psychologists at the level required for rigorous economic analysis. These questions should be reexamined using economic frameworks.

Economists should promote better systems of data collection that address the basic identification questions in the field. Personality measurements are being collected worldwide in a variety of contexts, and economists have the opportunity to contribute to and improve these measures. This offers a great opportunity to obtain a greater understanding of an important source of individual differences.■

---

[1]The work described in this article is detailed in J. J. Heckman and T. Kautz, "Hard Evidence on Soft Skills," *Labour Economics* 19, No. 4 (2012): 451–464.

[2]In 1904, La Société Libre pour l'Etude Psychologique de L'Enfant appointed a commission, led by Alfred Binet, to create a mechanism for identifying students in need of specialized education programs.

[3]See E. F. Lindquist, "Preliminary Considerations in Objective Test Construction," in *Educational Measurement*, ed. E. F. Lindquist, (Washington, D.C.: American Council on Education, 1951).

[4]A. E. Poropat, "A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance," *Psychological Bulletin* 135, No. 2 (2009): 322–338.

[5]J. J. Heckman, J. E. Humphries, S. Urzua, and G. Veramendi, "The Effects of Educational Choices on Labor Market, Health, and Social Outcomes," *Journal of Political Economy*, (2013, under revision).

[6]See for example, E. E. Noftle and R. W. Robins, "Personality Predictors of Academic Outcomes: Big Five Correlates of GPA and SAT scores," *Journal of Personality and Social Psychology* 93, No. 1 (2007): 116–130; and R. N. Wolfe and S. D. Johnson, "Personality as a Predictor of College Performance," *Educational and Psychological Measurement* 55, No. 2 (1995): 177–185.

[7]See M. Almlund, A. Duckworth, J. J. Heckman, and T. Kautz, "Personality, Psychology, and Economics," in *Handbook of the Economics of Education*, Vol. 4, 2011, pp. 1–181; and Heckman and Kautz, "Hard Evidence on Soft Skills."

[8]C. V. Edlund, "The Effect on the Behavior of Children, as Reflected in the IQ Scores, When Reinforced After Each Correct Response," *Journal of Applied Behavior Analysis* 5, No. 3 (1972): 317–319; and S. E. Breuning and W. F. Zella, "Effects of Individualized Incentives on Norm-Referenced IQ Test Performance of High School Students in Special Education Classes," *Journal of School Psychology* 16, No. 3 (1978): 220–226.

[9]L. Borghans, H. Meijers, and B. ter Weel, "The Role of Noncognitive Skills in Explaining Cognitive Test Scores," *Economic Inquiry* 46, No. 1 (2008): 2–12.

[10]Noncognitive skills can be measured by participation in risky behaviors such as crime, violence, sexual activity, smoking, and alcohol use during the adolescent years. The proportions of GED recipients and high school dropouts who exhibit these behaviors are similar, and much higher than those of high school graduates.

[11]J. J. Heckman, J. E. Humphries, and T. Kautz, eds., *The GED and the Problem of Character in American Life*, (Chicago: University of Chicago Press, forthcoming). See especially Chapter 5.

[12]For further discussion, see Heckman et al., *The GED and the Problem of Character*.

[13]W. Mischel, *Personality and Assessment* (New York: Wiley, 1968); and R. H. Thaler, "Short Course in Behavioral Economics," Edge Master Class, Sonoma, CA, July 25–27, 2008, at http://www.edge.org/3rd_culture/thaler_sendhil08/thaler_sendhil_index.html.

[14]See the evidence cited in L. Borghans, A. Duckworth, J. J. Heckman, and B. ter Weel, "The Economics and Psychology of Personality Traits," *Journal of Human Resources* 43, No. 4 (2008): 972–1059.

[15]For a summary of the literature see, for example, F. Cunha, J. J. Heckman, L. J. Lochner, and D. V. Masterov, "Interpreting the Evidence on Life Cycle Skill Formation," in *Handbook of the Economics of Education*, eds. E. A. Hanushek and F. Welch (Amsterdam: North-Holland, 2006), pp. 697–812; and J. J. Heckman and D. V. Masterov, "The Productivity Argument for Investing in Young Children," *Review of Agricultural Economics* 29, No. 3 (2007): 446–493.

[16]J. J. Heckman, S. H. Moon, R. Pinto, and A. Q. Yavitz, "The Rate of Return to the High Scope Perry Preschool Program," *Journal of Public Economics* 94, No. 1–2 (2010): 114–128.

[17]A. Becker, T. Deckers, T. Dohmen, A. Falk, and F. Kosse, "The Relationship between Economic Preferences and Psychological Personality Measures," *Annual Review of Economics* 4 (2012): 453–478.

Year Up: Providing a pathway from poverty to a professional career for urban young adults

Gerald Chertavian

Gerald Chertavian is founder and CEO of Year Up. He delivered the annual IRP New Perspectives in Social Policy Seminar on October 3, 2012.

This article is based on the book entitled A Year Up: How a Pioneering Program Teaches Young Adults Real Skills for Real Jobs with Real Success *and the 2012 to 2013 IRP New Perspectives in Social Policy Seminar. It is followed by a reaction by Carolyn Heinrich to the book, and a response to those comments by Gerald Chertavian.*

Lack of money gets less education
No jobs, and further degradation
It's getting devastatin'
And I walk both sides of the tracks
So who better to provide you the facts?

—Bakari Barrett, Year Up Graduate

There are 6.7 million young adults ages 16 to 24 nationwide who are not employed or in school and who do not have more than a high school diploma. Many of these young adults face multiple and significant challenges to entering the job market, even beyond national economic woes. They lack access to information about job openings as well as transportation and other resources that would allow them to work. Previous strategies to provide employment to this population have been largely unsuccessful; in the few cases where interventions resulted in a positive effect on employment and earnings, those gains generally disappeared over time.

There are many jobs requiring secondary education that go unfilled, while at the same time young adults with the potential to fill them lack access to the economic mainstream and to a way to obtain the needed skills. Year Up seeks to bridge this gap by providing a year of training to prepare low-income young adults for positions with good wages and career advancement opportunities in expanding fields. This is done using a high support, high expectation model that combines marketable job skills, stipends, internships, and college credits. We enhance students' professional and personal development in order to put these young adults on a viable path to economic self-sufficiency.

Skills gap

Figure 1 shows the percentage of 16- to 24 year-olds employed in the United States, from 1948 through 2011. After

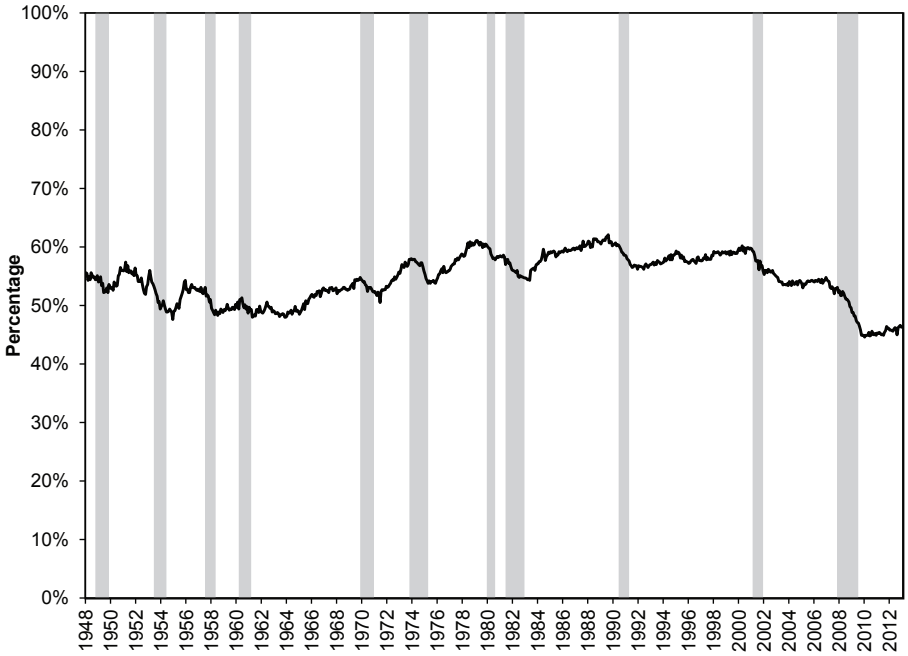

Figure 1. Percentage of 16- to 24 year-olds employed in the United States, 1948 to 2012.

Note: Grey bars reflect recessions.
Source: Seasonally adjusted employment-population ratio for 16- to 24-year-olds, Bureau of Labor Statistics. Recession periods are from the National Bureau of Economic Research.

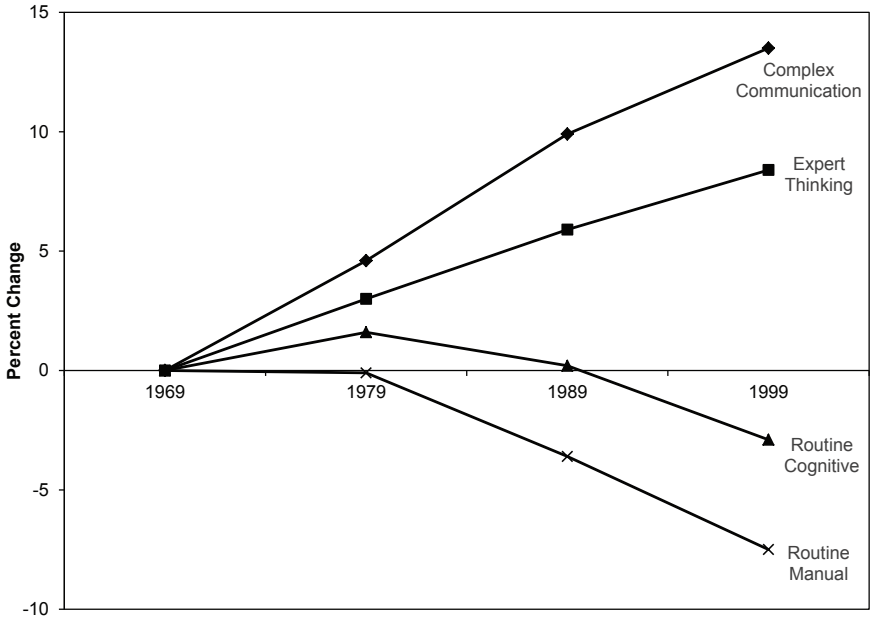

Figure 2. Changes in the tasks carried out by the American workforce, 1969 to 1999.

Source: Adapted from R. J. Murnane, "Preparing to Thrive in 21st Century America," presentation to the Mobile Area Education Foundation, February 29, 2008.

a steep decline over the last decade, the employment rate has now leveled out at about 45 percent. Disconnected young adults, or "Opportunity Youth" as we prefer to call the population, represent untapped potential for our nation. It is estimated that the immediate taxpayer burden of all disconnected young adults in lost revenues and increased social services is $93 billion, while the aggregate taxpayer burden of all disconnected young adults over their lives is $1.6 trillion.[1] Despite these challenges, a majority of these young adults are optimistic that they can achieve their goals, and they accept responsibility for their own futures.[2]

As the United States shifts to a knowledge economy, our demand for skilled workers is growing. Figure 2 shows a divergence of tasks carried out by the U.S. workforce from 1969 to 1999. The demand for routine manual and cognitive tasks has greatly decreased, while that for complex communication and expert thinking has increased.[3] This divergence will only continue, so in order to have a healthy economy, the United States must figure out how to provide workers with the needed skills.

Despite this increased demand for skilled workers, the supply is not keeping up. Even in a time of high unemployment, there are 3.7 million open job vacancies.[4] A recent study suggests that shortages of workers with some college-level skills could increase to more than 14 million by 2020.[5] Employers take longer to fill their job openings because applicants lack vital skills, such as communication, teamwork, professionalism, and critical thinking.

Year Up

Year Up was founded in 2001 in response to these challenges, in order to help close the gap between disconnected young adults and open job vacancies by providing urban young adults with the skills, experience, and support that will empower them to reach their potential through professional careers and higher education. The goal is, in one year, to take an individual from poverty to a professional career.

Program model

The Year Up program model is illustrated in Figure 3. A rigorous admissions process is followed by five months of skills training. Students learn marketable skills in areas such as information technology, financial operations, and quality assurance. Training is also provided in professional skills, everything from dressing and communicating professionally to managing personal finances. Year Up partners with colleges, allowing students to earn college credit for the satisfactory completion of classes. Students spend the next six months in full-time internships, applying their new skills with Year Up's corporate partners. Each site graduates two classes a year; when one class of students begins their internships, a new class begins the training phase.

All new students sign a performance contract agreeing to adhere to rigorous requirements including high attendance rates, punctuality, professional dress, and completion of assignments. A weekly stipend paid during both the classroom and internship phases of the program is tied to the performance contract; infractions of the contract requirements result in the loss of a portion of the week's stipend. Students who repeatedly fail to meet expectations effectively "fire" themselves from the program.

High expectations are accompanied by an extensive support system; students are assigned staff advisors who check in as often as needed. Mentors from the business community

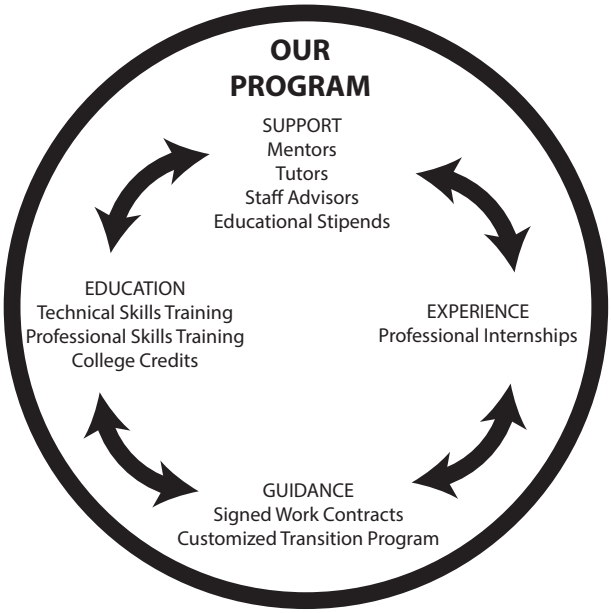

Figure 3. Year Up program model.

provide support, serve as role models, and provide networking opportunities. Mental health professionals are available on-site to help students cope with the persistent challenges they face in their daily lives. An essential component of this system is peer support; an orientation week is designed to create bonds that often last long after program completion, and students learn to rely on each other for support to complete the demanding program.

Year Up growth and results

In 2001, Year Up served 22 students in one site in Boston, Massachusetts. In 2013, Year Up will operate in 11 cities serving 1,900 students each year. The operating budget for 2012 was $48 million, and we have over 250 corporate partners. Since its founding, Year Up has served over 6,000 young adults, and currently has 3,464 alumni.

Within four months of graduation, 85 percent of our alumni are employed, in school full time, or both. Those with full-time employment earn an average of $15 per hour, or around $30,000 per year. Ninety-five percent of Year Up interns met or exceeded corporate partner expectations. An outside evaluation of the program (described in detail in the following article) found that Year Up participants earned an average of 30 percent more than a control group, and were more likely to be employed in the professional industries targeted by the program.

What have we learned about what works?

In looking at why the Year Up model has been successful, and what we have learned about providing services to young adults, there are three lessons that stand out: have high expectations and provide high support; emphasize soft skills; and align technical training with employers' needs.

High expectations and high support

The combination of high expectations and high support guides everything we do. Participants are expected to behave as they would in corporate America, and are treated as such; the program provides a set of expectations, and the student decides whether or not they want to abide by them. If the student chooses not to do so, they fire themselves from the program because they are accountable for their own actions. We often tell students that "the most respect we can pay you is to expect a lot from you."

Enforcing this is often extremely difficult to do. However, our staff understand the demands of the private sector and demonstrate to our students what will be expected when they enter the workforce. This professional support is augmented by social workers and clinical psychologists who provide essential assistance to students dealing with a wide range of issues including post-traumatic stress disorder, depression, self-medication, and sexual abuse. Without this support in place when a crisis inevitably occurs, few students would be able to complete the program.

Emphasis on soft skills

As noted in the article by James Heckman, professional or "soft" skills are extremely important for success in the workplace. Training in these professional skills—referred to in the Year Up vernacular as *pro skills*—is integrated into both the curriculum and the program culture. For example, an instructor might interrupt a lesson to point out a student's nonprofessional behavior during class, and use it as an opportunity to discuss what constitutes appropriate behavior. Staff members are trained in facilitation techniques, and know how to give feedback in a way that it will be heard and accepted by students. Students also receive direct instruction in a variety of professional skills including identifying and dealing with personality types and conflict behaviors, leadership, and team building. Students are taught business etiquette, including topics such as proper table manners for a business lunch, how to write thank-you notes, and appropriate body language.

Alignment with labor market

We are highly focused on providing the technical skill development that is required by businesses. We have a very close relationship with employers, surveying them twice during each internship period, and checking in with them every two weeks. We are responsive to feedback, continually adapting our curriculum to meet employer and industry needs. Ongoing communication with employers also makes us aware of specific skill gaps that corporations have identified.

There are several reasons why corporate partners agree to invest in Year Up. Foremost, we offer a "value proposition" that aligns well with the core business objectives of our corporate partners. Interns enter their workplace poised to develop the full range of skills that will ultimately make them valuable contributors and quality employees. This includes every detail of actually getting to their job on time and being prepared to work, including becoming familiar with the se-

curity procedure necessary to enter the building, and figuring out exactly how long it will take to commute to work from their home. In addition, corporate partners invest in our program because the internship is specifically structured to be low-risk and high-reward for employers; if the internship is unsuccessful, employers do not pay. Since Year Up depends on contributions from employers to operate, this provides a strong incentive to us to make sure we deliver.

As described above, a recent rigorous experimental evaluation has provided evidence that Year Up does deliver on our promises. Even James Heckman, who has argued that training programs aimed at older youth are often ineffective, and that limited resources should be invested in younger children, has noted that programs like Year Up that put an emphasis on soft skills have been finding success.[6]

How do we scale our effects?

While Year Up has certainly been effective at helping the young adults we have served find success in the workforce, we have to date been operating on a relatively small scale. The challenge and opportunity that we now face is to figure out how to expand our model to reach a greater number of people. The number of "disconnected youth" in the United States is large and growing, and our current reliance on private philanthropy restricts our growth.

In order to fully address the large and growing skills gap described above, direct service must be accompanied by systemic change. We believe that the current economic and political environments present a tremendous opportunity to effect such change, and that Year Up can make use of its increasing visibility and credibility on the national stage to assist in that effort.

Our strategy to expand and build on the success of Year Up has three parts: (1) to grow and strengthen our core model; (2) to develop a new "million person" model; and (3) to help create systems change.

Grow and strengthen the core

Although we recognize the need to develop a new model that can serve more people, we will of course continue to grow and strengthen our core model. We are adding programs in new cities, as well as expanding the program in current sites; by 2016, we expect to be serving 2,500 students each year. Growth in current sites is also helping us develop stronger ties to our local communities. We will continue to work on improving the program, and on ensuring the long-term success of our graduates. We are involved in the Innovative Strategies for Self-Sufficiency project (ISIS), a large-scale, rigorous evaluation of nine innovative career pathways programs across the country. ISIS is funded by the U.S. Department of Health and Human Services' Administration for Children and Families, and led by Abt Associates, a research and program implementation firm. Through this and other long-term evaluation tools, we will continue to prove and improve our program model.

Develop a million person model

At the same time, we are in the process of designing and piloting, in partnership with community colleges, alternative program models that can ultimately serve over one hundred thousand students each year. The colleges will provide the training, while we will provide needed academic, financial, social, and emotional support services to students, as well as internship placements. Through these efforts, we hope to increase graduation rates, and create more successful transitions into the labor market. By making use of the colleges' existing infrastructure, we are able to reduce program costs to a level where they can be covered by Pell Grants and internship fees, and thus require no philanthropy. These features make this type of model easier to scale up, so that much greater numbers of students may be served. We currently have community college partnerships in all of the cities that host Year Up sites, including transferable credits for Year Up course work.

Create systems change

We are also working to change the way people think about urban young adults and how we, as a nation, can better prepare them for the 21st century economy. Building on the credibility of our core program, our efforts focus on influencing the three key areas of perception, practices, and policies. Thus, we are working first to improve the perception of urban young adults, so that they are seen as economic assets rather than social liabilities. As an example of that effort, the nonprofit Ad Council has recently agreed to produce a series of public service announcements to promote this concept. Second is changing employer practices around finding and developing talent. This means working with employers to identify the skills they really require and the credentials that can provide them, and, for example, changing the common requirement of a four-year degree. Third and finally, we are supporting public policies that increase the number of effective pathways to work for young adults. We will continue these efforts at both the national level and in the various metropolitan areas where our sites operate. Our growing corps of alumni will be critical allies as we pursue this strategy to broaden access to meaningful career pathways to all young adults.■

---

[1]C. R. Belfield, H. M. Levin, and R. Rosen, *The Economic Value of Opportunity Youth*, (Washington, DC: Corporation for National and Community Service, 2012).

[2]J. M. Bridgeland and J. A. Milano, *Opportunity Road: The Promise and Challenge of America's Forgotten Youth*, (Washington, DC: Civic Enterprises, January 2012).

[3]R. J. Murnane, "Preparing to Thrive in 21st Century America," Presentation to the Mobile Area Education Foundation, February 29, 2008.

[4]Bureau of Labor Statistics, *Job Openings and Labor Turnover Summary*, October 2012.

[5]A. Carnevale and D. M. Desrochers, "Standards for What? The Economic Roots of K-16 Reform," Educational Testing Service, 2003.

[6]J. Heckman, Comments at Clinton Global Initiative Plenary Session, Chicago, IL, June 29, 2011.

How does Year Up measure up?

Carolyn Heinrich

Carolyn Heinrich is Sid Richardson Professor of Public Affairs, Professor of Economics, and Director of the Center for Health and Social Policy at the University of Texas at Austin, and an IRP affiliate.

The story of *Year Up's* founding and expansion is as poignant as it is instructive in an academic sense. In his book, *A Year Up: How a Pioneering Program Teaches Young Adults Real Skills for Real Jobs with Real Success*, Gerald Chertavian describes his approach to developing the Year Up program, which was simultaneously grass-roots and high-level in its search for guidance and evidence to build an innovative and viable model for helping urban young adults gain access to the job market. In fact, if I had not read this book and had instead come across the Year Up "High Expectations" program model on my own, I would have guessed that Year Up's founder sat down with all of the academic literature to design this program. In offering comments on *A Year Up* in this article, I bring both research and policy perspectives to bear in considering the key, inventive aspects of the program model, Year Up's implementation and reported program results, and its limitations as well as its potential for successful replication and extension of promising features to other interventions for urban youth.

Year Up program model

The Obama administration made clear from the beginning its intent to support only programs that are evidence-based and that can demonstrate success.[1] Year Up has a "leg up" in both respects, in that it has drawn on the best current evidence for what works in connecting urban young adults to the job market, and has opened its doors to evaluators. The program combines both professional and technical skills training, recognizing the importance of developing non-cognitive as well as cognitive skills to prepare youth for the job market, such as leadership skills, decision-making strategies, team building, and business etiquette. In addition, peer support is an integral component of the comprehensive support services that aid program participants in navigating their individual barriers and challenges to success. Year Up also spends considerable time and effort on curriculum development and training to ensure that they meet the labor needs of employers and that both program participants and prospective employers see their relevance. Through its close relationships with business partners and investors, Year Up is able to provide state-of-the-art training, teaching technology skills with cutting-edge software and equipment.

The key role that peers play in supporting the development of professional skills is one of the most important and distinctive innovations of Year Up. This support comes into play in a variety of ways, including helping with assignments, correcting each other's language, and making encouraging phone calls to keep peers engaged. Moreover, Year Up does not shy away from addressing some of the more difficult workplace issues through peer-led exercises such as "Turn Your Back," which is used for processing hurtful stereotyping and discrimination that are experienced on the job. These features of the Year Up culture simultaneously build lifelong friendships and professional networks, while contributing to the program's high completion rates.

Another important innovation of Year Up is the continued intensive support of young adults during their six-month internship. This support helps employers to see Year Up as a "hiring pipeline": students are trained and integrated into companies through the internships. In turn, the students can count on ongoing peer and staff support from Year Up, as well as financial support in the form of a weekly stipend that is tied to the performance contract. Students also earn college credit for training through Year Up partner institutions.

Finally, it is important to acknowledge the key role that the founder's executive connections have played in reaching the upper echelon of private sector firms and securing from them both donations and opportunities for the young adult participants. Philanthropy has also been essential to expansion of Year Up to other cities. As Chertavian explains: "You can't afford the people, the infrastructure, the benefits—the mission itself—without a constant, renewable source of philanthropic investment" (p. 272). However, this reliance on dedicated corporate partners could also pose a potential obstacle to implementing the program model on a much larger scale.

Year Up results

The results that Year Up reports are impressive: 70 percent of those entering the program complete it, and all *qualified* students are placed into internships. Ninety-five percent of Year Up interns meet or exceed their internship manager's expectations, and 85 percent of program graduates are either employed or attending college full time within four months of program completion. Employed Year Up graduates earn $15 an hour on average, or approximately $30,000 per year. However, what these figures do not tell us is how much of these results are attributable to the program, compared to what these highly motivated young adults would have achieved on their own.

An experimental evaluation of the Year Up program is being conducted by the Economic Mobility Corporation.[2] Eligible

Key Findings from an Evaluation of Year Up
Year Up's Initial Effects

Characteristics of study participants

Most Year Up participants are members of racial or ethnic groups that face discrimination in the labor market.

When they applied to Year Up, 81 percent of study participants lived with a parent or guardian, a higher percentage than for the overall population of young adults.

35 percent attended college at some point. Only one individual had obtained an associate's degree, and only six percent had attended college during the semester immediately prior to applying to Year Up.

In focus groups, participants who had attended college reported dropping out of college for financial reasons, or because required remedial courses prevented them from taking for-credit classes and progressing toward a degree.

Most study participants have some work experience, but fewer than half were employed at the time they applied to Year Up.

Most participants with work history had held low-wage jobs for short periods of time. The most common jobs were in food service and retail trade.

Table 1
Characteristics of Study Participants

Gender	
Male	57%
Female	43
Age	
18 to 21	72%
22 to 24	28
Race	
African American	50%
Latino	34
White	5
Asian	3
Other	7
Highest Degree	
GED	14%
High School Diploma	85
Associate's Degree	1
Work Experience	
Ever Worked for Pay	88%
Working at Time of Program Application	43%
Longest Job Less than One Year	58%
Median Hourly Wage in Longest Job	$8.25
Other Characteristics	
English not Primary Language	15%
Not a U.S. Citizen	8
Have Child(ren)	9
Convicted of a Crime	8
Live in Public Housing	18

Year Up Participants had greater earnings in the second year after random assignment

During the first year after random assignment, while treatment group members attended the program full-time, control group members had higher average earnings. However, during the second year after random assignment, following program completion, annual earnings for those in the treatment group were on average 30 percent higher than earnings for those in the control group ($15,082 compared to $11,621).

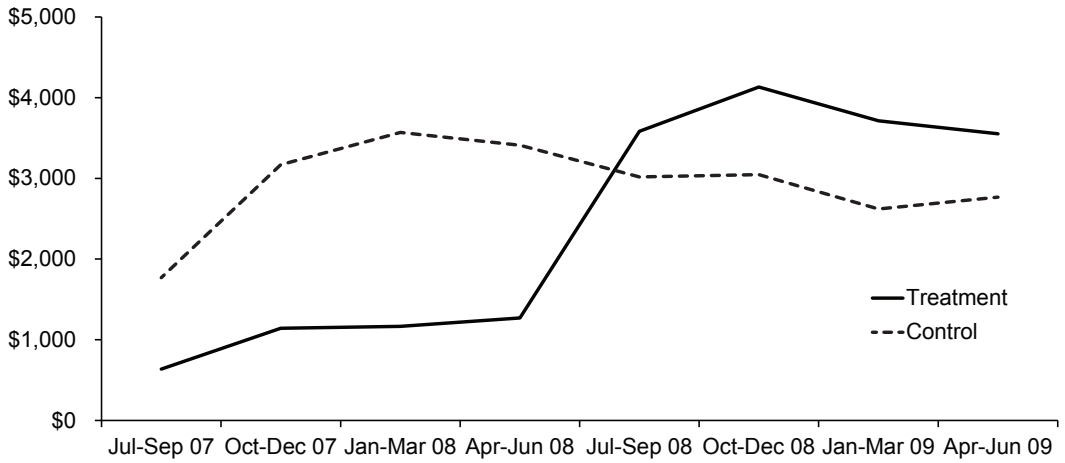

Figure 1. Total earnings during each quarter after random assignment.

Note: Differences are statistically significant at the $p < .05$ level in October to December 2008 and January to March 2008, and at the $p < .10$ level in April to June 2009.

Higher hourly wages drove the earnings difference

Both groups were equally likely to be employed during the second year, but Year Up participants tended to have higher paying jobs.

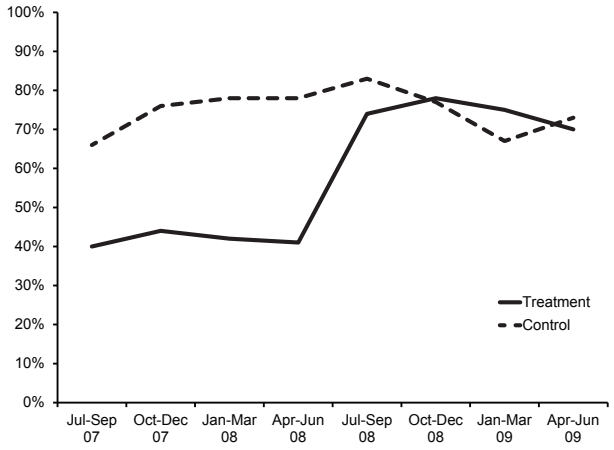

Figure 2. Percent employed at any time during each quarter after random assignment.

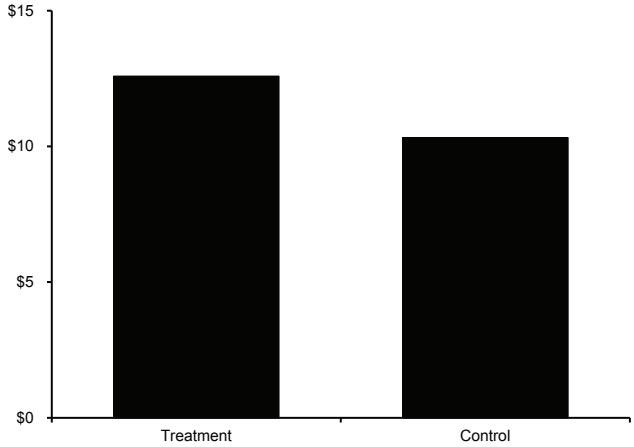

Figure 3. Average hourly wage at current or most recent job.

Year Up participants were significantly more likely than control group members to obtain jobs in the targeted fields, and were more likely to be working full time. The hourly wages of Year Up participants who worked in fields other than information technology and investment operations did not differ significantly from the wages of control group members.

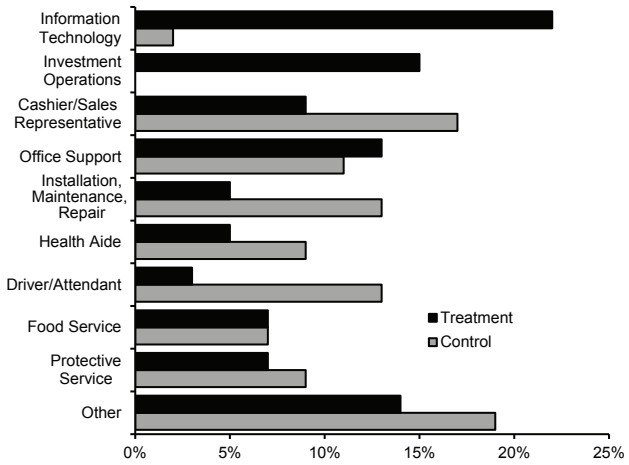

Figure 4. Type of Job held: Current or most recent job.

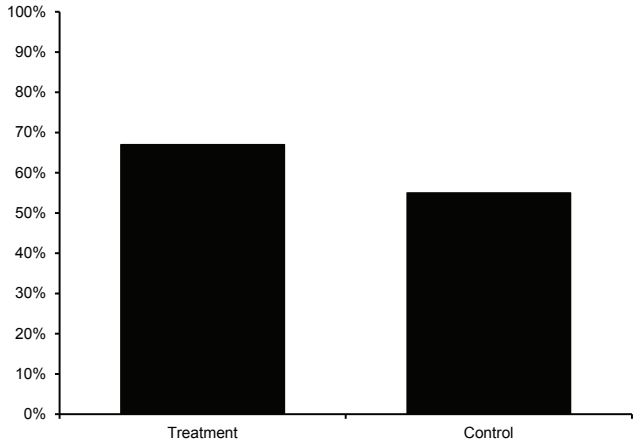

Figure 5. Current or most recent job is full-time.

Year Up participants were just as likely as those in the control group to attend college

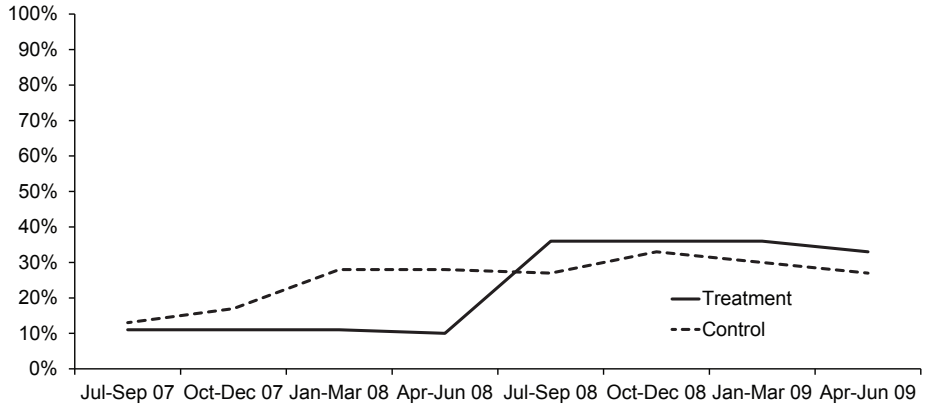

Figure 6. Percent attending college during each quarter after random assignment.

Note: Results from A. Roder and M. Elliott, *A Promising Start: Year Up's Initial Impacts on Low-Income Young Adults' Careers*, Economic Mobility Corporation, New York, NY, April 2011.

candidates in Boston, New York City, and Providence were randomly assigned to either a treatment group (who were able to participate in Year Up) or a control group (who had their names placed on a waiting list, and were told that they could reapply to Year Up after 10 months).[3] The sample was tracked for between 24 and 30 months following random assignment. The resulting sample was relatively small, 120 treatment and 44 control group members. The effects estimated in this evaluation represent the average effect of the intent to treat. That is, the analysis includes all members of the treatment group, regardless of whether or not they ever attended or graduated from the program.

Study results are summarized on pages 14–15. As one might expect, the evaluation found that during the course of the program, control group members earned more on average than Year Up participants. However, in the year following program participation, annual earnings for those in the treatment group were on average 30 percent higher than earnings for those in the control group ($15,082 compared to $11,621). Treatment and control group members were equally likely to be employed during the second year (86 percent for treatment group members, 83 percent for those in the control group), but treatment group members had higher hourly wages ($12.58 compared to $10.32), and were more likely to be working full time. Both groups were equally likely to be attending college during the second year after enrollment.

One somewhat surprising result of this study was the finding of no differences in the availability of employer-provided health benefits or tuition assistance for Year Up participants compared to those in the control group. Year Up cultivates relationships with many large, well-known employers, who might be expected to be more generous than average in their employee benefit offerings. However, these are only short-term results, and studies of other employment programs have often found that program impact estimates change over time. A future report from the Economic Mobility Corporation will address whether earnings gains persist over four years and will also look at program costs and cost-effectiveness.

Comparison to other programs

If I could strike a sentence from the book, it would be one that appears on page 11, in a discussion of the challenging environment that young job seekers face: "Add it all together and you get a workforce development system that really doesn't work." In fact, the pattern of impacts for Year Up looks very similar to the patterns that have been found in public training programs targeted toward young, disadvantaged adults, including those in a recent evaluation of the Workforce Investment Act (WIA) System that I conducted with colleagues.[4] We found an estimated average increase in earnings of approximately $2,400 per year, or 26 percent of average earnings, for disadvantaged women who participated in WIA; disadvantaged men had an average earnings increase of around $1,700, or 15 percent of average earnings.

The U.S. Job Corps program, which began in 1964 and is administered by the U.S. Department of Labor, is another useful comparison to make with Year Up, because both their missions and costs are very similar. Job Corps offers free education and vocational training to disadvantaged youth ages 16 to 24, primarily in a residential setting, with the ultimate goal of placing students in jobs that are well-matched to their newly-acquired skills. A national study conducted in the 1990s found that four years after program entry, average weekly earnings for treatment group members were $22 per week higher than that for control group members ($1,150 annually), a 12 percent earnings gain. The study found that earnings gains persisted beyond the fourth year, in years 5 through 10, but only for those who were 20 to 24 years old at program application.[5] Other findings of the Job Corps evaluation included increased receipt of GEDs and vocational certificates for those in the treatment group by more than 20 percentage points each, and significant reductions in criminal activity. The Job Corps evaluation raises the question of whether a broader range of potential impacts could also be measured in the Year Up evaluation.

The future of Year Up

Although Year Up's founder loathes the assertion that the program selects the "cream" among their applicant pool to get results, applicants are required to go through a rigorous admissions process that helps Year Up to identify those who appear to have the motivation and resilience necessary to succeed. Year Up does not work with students who do not have a high school diploma or a GED; this leaves out a large fraction of disconnected youth in the targeted age group. Just 52 percent of black males earned a high school diploma in 2010, up from 47 percent in 2008.[6] While Year Up clearly still serves a group of young people who are overcoming serious disadvantage, the evaluation results may only generalize to the more motivated group of young adults typically selected for the intervention, not to the larger population of disadvantaged youth.

This begs the question: Can we now take key innovations and insights from the successes of the Year Up program and extend them to other interventions or subgroups of youth? For example, would it be possible to intervene in similar ways earlier in the lives of those youth who drop out before completing high school? Recent research from the Harvard Center on the Developing Child suggests potential for enhancing the development of "executive function skills"—similar to Year Up "pro skills"—at younger ages.[7]

Year Up has likewise not been content in resting on its successes to date. It has revisited its own theory of change and posed the question: How can we now tackle the root causes that drive the need for Year Up? One approach that is being tried is the college-based pilot program, which attempts to move the Year Up model into community colleges. There may also be opportunity for Year Up to partner with and extend its innovative features to other organizations that serve

youth and young adults, including high school-based programs and organizations such as Jobs for Youth Chicago that share its mission and commitment to helping disadvantaged young people succeed in life.■

---

[1]R. Haskins and J. Baron, *Building the Connection Between Policy and Evidence* (London, UK: NESTA, 2011).

[2]A. Roder and M. Elliott, *A Promising Start: Year Up's Initial Impacts on Low-Income Young Adults' Careers*, Economic Mobility Corporation, New York, NY, April 2011.

[3]135 were placed in the treatment group, 60 in the control group. Individuals were most recently surveyed at between 24 and 30 months following random assignment. The follow-up survey response rates were 89 percent for the treatment group and 73 percent for the control group. Researchers found very few significant differences in attrition rates between the two groups and little evidence that those differences could have resulted in bias in the estimated program effects.

[4]C. J. Heinrich, P. R. Mueser, and K. R. Troske, "Workforce Investment Act Non-Experimental Net Impact Evaluation: Final Report," Report to U.S. Department of Labor, IMPAQ International, Columbia, MD, 2008.

[5]P. Z. Schochet, J. Burghardt, and S. McConnell, *National Job Corps Study and Longer-Term Follow-Up Study: Impact and Benefit-Cost Findings Using Survey and Summary Earnings Records Data,* Final Report, Mathematica Policy Research, Princeton, NJ, 2006.

[6]Schott Foundation, *The Urgency of Now: The Schott 50 State Report on Public Education and Black Males 2012*, The Schott Foundation for Public Education, Cambridge, MA, 2012.

[7]Center on the Developing Child at Harvard University, "Building the Brain's "Air Traffic Control" System: How Early Experiences Shape the Development of Executive Function," Working Paper No. 11, 2011. Retrieved from www.developingchild.harvard.edu.

Response from the author

Gerald Chertavian

I appreciate Carolyn Heinrich's thoughtful comments in response to *A Year Up* and welcome this opportunity to discuss the current nature and future of the program. I have broken down my responses into three categories, addressing issues of scaling, impact, and lessons learned.

Scaling

Heinrich raised a concern about whether the Year Up program model limits scalability. Our program does depend on a number of factors that limit where we are able to expand. For example, public transportation, innovative community college partners, and the depth of the local philanthropic base are factors we consider when looking to expand. However, we believe the basic components of our program model are scalable. These include high expectations and high support for our students, emphasis on both technical and professional skills, and education that is closely tied to work experience.

By 2016 we will serve 2,500 students annually in 12 cities. Even if we were to grow our program to 25 cities, we would still be serving less than one percent of "Opportunity Youth," 16- to 24-year-olds who are neither enrolled in school nor fully participating in the labor market.[1] Our direct service program is central to our enterprise and we are committed to building and strengthening Year Up as a vibrant community asset. However, in order to fully realize our mission of closing the Opportunity Divide, we are pursuing two complementary strategies.

First, we are working to eliminate the barriers that perpetuate the Opportunity Divide by changing national systems. We are engaging and partnering with corporate partners, academics, community-based organizations, and policymakers to build a climate where all young adults have access to meaningful careers and quality postsecondary education. Second, we are designing and piloting alternative program models that can grow rapidly to serve more than one hundred thousand young adults across the United States each year. These "Million Person Model" pilots are being built using the successful practices of our core program and in partnership with community colleges. Currently, we are piloting an alternative model in Baltimore with Baltimore City Community College and in Miami with Miami Dade College.

Impact

In reference to Heinrich's remarks about the experimental evaluation outcomes, it is important to remember that treatment group effects are generally only apparent following the completion of the training program, as was the case with Heinrich's own Workforce Investment Act (WIA) research. What is particularly notable about the Year Up results is the magnitude of the observed effects and the differences in the employment sectors. We are committed to helping our students' secure professional jobs in the sectors for which they were trained while attending Year Up. Thus, although rates of employment may be only nominally higher than the control group, the 30 percent boost in annual earnings is substantial and lasting. These results are even more significant when considering that they occurred in a program serving low-income young adults, a population that has traditionally been very difficult to serve. Furthermore, these results were attained without producing any negative impacts on the treatment group's college attendance. That is, even while getting into more demanding, higher-wage jobs, students were not forced to decide between education and work.

Year Up also has a strong focus on long-term outcomes. While we report postsecondary enrollment four months after the program, we continue to support our alumni in their work towards postsecondary completion. Part-time students in our target population only graduate at a rate ranging between 12 percent and 24 percent. While we are still collecting conclusive evidence, the early reports are very promising: a survey of our alumni shows that of those who enrolled in school, 84 percent have either completed or persisted in their studies. We are confident that our alumni, anchored by career pathway jobs and a base of college credit received while at Year Up, will complete postsecondary credentials at substantially higher rates than their peers.

The workforce development system

Heinrich took exception to the sentence from *A Year Up*, "Add it all together and you get a workforce development system that really doesn't work," noting that the pattern of effects for Year Up looks very similar to those observed for disadvantaged adults in the WIA System. While we believe strongly in the value of government-supported workforce development initiatives, I would counter that alternative training programs are not producing the outcomes needed. A 2010 Department of Labor study on federally financed workforce training programs found that "ultimate gains from participation are small or nonexistent," with graduates earning no more than nonparticipants three and four years later.[2] While nearly half of the respondents to a recent employer survey offered workforce readiness training for their new hires, most report the programs to be only "moderately" or "somewhat" successful.[3]

Year Up is providing a clear, workable solution to a persistent and significant problem. The result of failing to educate and train new workers is that businesses lack the talent they need

to thrive. Out of nearly 1,200 organizations polled in a study by the American Society for Training and Development, 79 percent state that a "skills gap" exists in their organization, defined as "a significant gap between an organization's current capabilities and the skills it needs to achieve its goals." The top reason given for that skills gap was that the skills of the current workforce do not match changes in company strategy, goals, markets, or business models.[4]

Measuring the program

Heinrich asked whether Year Up should measure a broader range of program effects. Given our relatively brief tenure as an organization, we currently assess outcomes that the program can reasonably influence in the short-term, such as employment, salary, and school enrollment. We are strongly committed to understanding the longer term impacts of our program as we continue to engage and support our growing alumni base. In our first Long-Term Graduate Success Survey, which surveyed over 1,200 (61 percent) of our alumni in 2012, we examined factors such as homeownership and family life to gain a broader and deeper knowledge of where our alumni are now. With this information, we are in a better position to develop strategies that will ensure their success in the future. We are also participating in the Innovative Strategies for Increasing Self-Sufficiency study, a rigorous longitudinal evaluation of a variety of "career pathways" organizations being conducted by the Administration for Children and Families. The study will evaluate Year Up's impact on a much wider range of factors, including healthy behaviors and psycho-social development.

Population served

As Heinrich noted, Year Up does not work with students who do not have a high school diploma or a GED. Through our partnership with local colleges, Year Up participants receive college credit for the classes they take while in the program. To earn these credits, our college partners require that our participants have a high school diploma or GED. Additionally, corporate partners usually require that employees have a high school diploma or GED. We do work with a number of educational and community-based organizations that assist young adults who do not have a high school diploma or GED. Many young adults express interest in our program, find out about the GED requirement, complete a GED, and then subsequently enroll in Year Up. For all enrollees through November 2012, roughly 1 in 5 did not graduate from high school and instead earned a GED.

Heinrich notes that Year Up is a selective program, and that the evaluation results can only be generalized to the more motivated group of young adults selected for the intervention. We believe that when all young adults are held to high expectations and given high support, they are able to achieve anything they set out to do. Our students, like millions of young people, do not have access to the tools and support they need to connect to a meaningful career path, but are motivated and determined. Many of our students lacked the resources, both financial and non-financial, to persist in postsecondary education. They typically attended underperforming high schools with overburdened college and career counselors. Often they did not have a support system or role models who could help them navigate the transition from school to a long-term career.

During the admissions process, our Student Services department conducts a comprehensive "Readiness Assessment" with every student to identify challenges that could hinder his or her success in the program. We have found that over 40 percent of students have significant financial risk such as receipt of public benefits, working significant hours outside of the program, or not having health insurance. One-third of students have a family risk; they have been in foster care or group homes, are the primary caregiver for a family member, or are dealing with a significant family crisis. Fourteen percent of our students are parents, and 9 percent are homeless or transient. Our students, like many young adults in this country, face tremendous challenges to success.

Lessons learned

Heinrich wondered whether lessons from Year Up could be used to intervene earlier in the lives of at-risk youth. We strongly believe parts of our model are transferable and beneficial for younger audiences. As a member of the Massachusetts Board of Elementary and Secondary Education, I chaired a cross-sector Task Force on Integrating College and Career Readiness, which in part sought to harness the innovations and insights from a number of successful programs, including Year Up, to serve all Massachusetts students. For instance, Year Up's program is rooted in a "high expectations, high support" culture. We hold our students to high standards because we believe in their potential and refuse to accept anything less. We also want to prepare them for the rigor of the corporate world when they begin their internships and start their careers. To make sure that our students are set up for success, the program also provides the high level of support they need to navigate a challenging environment—these supports include a strong learning community of peers, staff advisors, volunteer mentors, tutors, and educational stipends.

In terms of extending the features of Year Up to other organizations that serve youth and young adults, we believe this squarely falls within our goal of closing the Opportunity Divide. To accomplish this, we will need deep and wide-reaching changes in our nation to provide every young person with the opportunity to succeed. Year Up offers one solution out of many, but we know we cannot realize our vision for the future on our own; we do not have all the answers. We believe sharing knowledge and information is essential to ensuring that we are all working together effectively to increase opportunities for young people. We are willing to share what we have learned with others, and are eager to incorporate new insights from our allies in the field. Through our systems change strategy, we work with others to expand meaningful career pathways for young adults and

connect employers to this underutilized pool of talent. For instance, through coalitions and strategic partnerships with like-minded organizations, we advocate for government funding that rewards program outcomes rather than efforts. We also work to engage more employers in providing career opportunities for young people through mentoring, internships, or innovative hiring practices that target Opportunity Youth. These efforts often point to Year Up as an example of "what works" in developing effective career pathways, but our goal is not to elevate Year Up, but to promote systemic changes in perception, practice, and policy that expand opportunities for all young adults.■

---

[1]C. R. Belfield, H. M. Levin, and R. Rosen, *Economic Value of Opportunity Youth*, Civic Enterprises, Report to the Corporation for National and Community Service and the White House Council for Community Solutions, 2012, at www.civicenterprises.net/MediaLibrary/Docs/econ_value_opportunity_youth.pdf.

[2]P. S. Goodman, "After Training, Still Scrambling for Employment," *New York Times*, July 18, 2010.

[3]J. Casner-Lotto, E. Rosenblum, and M. Wright, *The Ill-Prepared U.S. Workforce: Exploring the Challenges of Employer-Provided Workforce Readiness Training*, The Conference Board, American Society for Training and Development, Society for Human Resource Management, and Corporate Voices for Working Families, Research Report, 2009.

[4]P. Galagan, *Bridging the Skills Gap: New Factors Compound the Growing Skills Shortage*, American Society for Training and Development, 2010.

Kids' Share Project

Kids' Share is a series of reports from the Urban Institute looking at trends in federal and state spending and tax expenditures on children.

Kids' Share 2012: Report on Federal Expenditures on Children Through 2011
Julia Isaacs [IRP affiliate], Katherine Toran, Heather Hahn, Karina Fortuny, C. Eugene Steuerle

The most recent annual report provides a comprehensive look at trends over the past 50 years in federal spending and tax expenditures on children. Key findings suggest that the size and composition of expenditures on children have changed considerably, and that children have not been a budget priority. http://www.urban.org/publications/412600.html

How Targeted Are Federal Expenditures on Children? A Kids' Share Analysis of Expenditures by Income in 2009
Tracy Vericker, Julia Isaacs [IRP affiliate], Heather Hahn, Katherine Toran, Stephanie Rennane

This report provides an analysis of how the allocation of public resources for children varies by family income. Key findings indicate that in 2009, 70 percent of all federal expenditures on children served the 42 percent of children who are living in families with incomes less than twice the federal poverty level. While low-income children received 84 percent of outlays on children, higher-income children received 82 percent of tax reductions benefiting children. http://www.urban.org/publications/412522.html

Kids' Share Website: http://www.urban.org/projects/kids_share.cfm

Value-added measures of teachers: Research and policy

Value-added models in education are used to attempt to measure the contributions to student achievement of individual teachers. Test scores for a particular teacher's students are compared to those of the same students in the previous year, as well as to those of students with other teachers in the same grade, in an effort to isolate the contribution of the given teacher. Advocates of these methods argue that these measures provide objective information that can be used to improve instruction, while critics counter that their validity as an indicator of teacher quality is still in question. School districts from Washington D.C. to Los Angeles have started to use value-added measures, and some teachers' ratings have been made publicly available, including recently in the Los Angeles Times *and the* New York Times.

On June 27, 2012, as part of IRP's annual Summer Research Workshop, three researchers participated in a roundtable discussion of teacher value-added measures. Raj Chetty presented outcomes from a long-term study of the effects of teachers on students from elementary school through early adulthood. Jesse Rothstein explored the potential dangers of using value-added measures to make teacher personnel decisions. Finally, Eric Hanushek looked at policy implications from a different perspective, exploring why and how value-added measures can be used most effectively. This set of articles summarizes the three presentations.

The long-term effects of teachers

There is considerable debate about the best way to measure and improve teacher quality. One method is to rate teachers based on their students' test score gains, known as the "value-added" approach.[1] School districts have begun to use these measures to make personnel decisions about teachers. For example, District of Columbia Public Schools lay teachers off or offer them bonuses using a teacher-performance assessment system that puts 50 percent weight on value-added scores. This article describes an extensive study designed to estimate the effects of teachers on student outcomes through early adulthood.[2]

Potential issues with value-added models

The debate about using teacher value-added models stems primarily from three issues. First is concern about the potential for bias in value-added estimates; do differences in test-score gains across teachers capture causal effects, or are they instead driven by student sorting? Second is lack of evidence on the long-term effects of teachers; do teachers who raise test scores also improve students' long-term outcomes, or are they simply better at teaching to the test? The third issue is measurement error; are estimates based on only a few years of data accurate enough to be used for policy decisions?

Assessing teachers over the long-term

The study done by Raj Chetty, John Friedman, and Jonah Rockoff addresses all of the above issues by using data on a million children, from childhood through early adulthood.[3] The researchers developed new quasi-experimental tests to assess bias in value-added estimates. They look at whether those who had high value-added teachers as children have better outcomes in adulthood. Finally, they assess the monetary gains to be made by selecting teachers with higher estimated value-added scores, given observed measurement error.

The estimated teacher effects include both direct and indirect outcomes. For example, having a good teacher in the fourth grade can improve a student's labor market outcomes in adulthood directly, but it can also have indirect effects, if students with good teachers in the fourth grade receive better teachers in subsequent grades.[4]

Are value-added estimates unbiased?

Since students are not assigned to teachers randomly, it is necessary to adjust for the composition of students assigned to a classroom; the standard approach is to control for prior year variables. Recent studies have reached conflicting conclusions about whether this approach is sufficient for obtaining consistent estimates of teacher effects.[5]

In this study, the researchers evaluate whether or not value-added estimates are biased by (1) testing for selection on observable characteristics, and (2) using quasi-experimental methods that make use of natural teacher turnover. In order to test for selection on observable characteristics, they look at whether parent characteristics are correlated with teacher value-added scores, and find no relationship. For example, the children of wealthier parents are no more likely to get higher value-added teachers. In order to test for selection on unobservable characteristics, they looked at changes in students' scores in the year before and after a switch in teachers due to teacher turnover. Again, they find no evidence of selection. Therefore, they conclude that their value-added measures provide unbiased estimates of teachers' causal effects on student test scores.

This article summarizes the presentation given by Raj Chetty.

Raj Chetty is Professor of Economics at Harvard University.

Effects on outcomes in adulthood

Next, the researchers assessed whether teachers who raise test scores also improve their students' outcomes in adulthood. They analyze the effects of teachers on three sets of outcomes; college attendance, earnings, and other indicators such as teenage birth rates.

Being assigned to a higher value-added teacher in a single grade significantly raises a student's likelihood of attending college. A one standard deviation increase in the value added of a teacher appears to increase the probability of that student attending college by age 20 by 1.25 percent. Students with higher value-added teachers are also more likely to attend a better college, as measured by projected average earnings at age 30.

Having a higher value-added teacher has a clear statistically significant effect on earnings. An increase in teacher value added of one standard deviation increases annual earnings at age 28 by $182. The lifetime financial value of having a teacher one standard deviation higher is approximately $4,600 per grade.[6]

Having a teacher one standard deviation higher in value added in a single year from grades 4 through 8 reduces the probability of a teen birth for female students by 1.25 percent.[7] Students with higher value-added teachers are also more likely to live in higher socioeconomic status neighborhoods as adults.

Measurement error and policy relevance

Any evaluation of teachers based on value-added measures must rely on only a few years of classroom data. This limited amount of data adds uncertainty to value-added estimates, thus potentially reducing their utility for performance evaluation. In order to evaluate how much the utility is reduced, it is necessary to look at a policy example. Thus, the researchers analyze the effects of retaining or firing teachers on the basis of their value-added scores.

On average, replacing a teacher in the bottom 5 percent with an average teacher for one year raises a child's cumulative lifetime income by $50,000. For a class of average size (28 students), the cumulative lifetime income gains from a high value-added teacher exceed $1.4 million. This is equivalent to $267,000 in present value at age 12, discounting future earnings gains at a 5 percent interest rate. Of course, data limitations do not allow certainty about which teachers are in the bottom 5 percent. In estimating the gains of deselecting teachers based on their *estimated* value added, there is still a substantial potential lifetime earnings gain. The present value of earnings gain from deselecting teachers below the fifth percentile increases with the number of classes observed per teacher. While the gain with even ten observed classes is still below the $267,000 value achievable with perfect knowledge of teacher rank, with even three or four observed classes, the lifetime gain is still around $200,000.

Policy implications

While the Chetty and colleagues study supports the idea that existing value-added measures are useful in identifying long-term effects of teachers, this conclusion alone is not sufficient to assess value added as a policy tool, for at least two reasons. First, it is necessary to weigh any potential gains against the cost of firing teachers. The researchers' calculations suggest that the financial benefits of such a policy far outweigh the costs. A second and more serious concern not addressed in this study is potential negative behavioral responses to testing when the stakes are so high, such as teaching to the test or even cheating.[8] It is possible that such responses, if sufficiently large, could completely counter any policy gains.

Parents should be interested in knowing the value added of their child's teacher, whether or not that information is useful as a policy tool. This analysis shows that high value-added teachers improve students' achievement and long-term outcomes. The most important lesson of this study is that finding policies to raise the quality of teaching—whether through the use of value-added measures, or through other tools such as salary structure changes or teacher training—is likely to have substantial economic and social benefits in the long run.■

---

[1]See, for example, E. A. Hanushek, "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data," *American Economic Review Papers and Proceedings* 61, No. 2 (1971): 280–88; and R. J. Murnane, *The Impact of School Resources on the Learning of Inner City Children* (Cambridge, MA: Ballinger, 1975).

[2]The study summarized here is described in detail in R. Chetty, J. N. Friedman, and J. E. Rockoff, "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood," NBER Working Paper No. 17699, National Bureau of Economic Research, 2011. http://www.nber.org/papers/w17699.

[3]The data link two large databases: student records from a large school district, including teachers, class assignment, and test scores from 1991–2009 for students in grades 3–8; and U.S. income tax records including both student outcomes (such as earnings, college, and teenage birth) and parent characteristics (such as income, savings, home ownership, mother's age at childbirth, and marital status).

[4]See, for example, T. J. Kane and D. O. Staiger, "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," NBER Working Paper No. 14607, National Bureau of Economic Research, 2008.

[5]Kane and Staiger, "Estimating Teacher Impacts on Student Achievement;" and J. Rothstein, "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement," *Quarterly Journal of Economics* 125, No. 1 (2010): 175–214.

[6]A one standard deviation increase in teacher value added in a single grade results in increased earnings at age 28 of $182, which is 0.9 percent of mean earnings in the regression sample. The researchers assume that the percentage gain in earnings remains constant at 0.9 percent over the lifecycle, and that earnings are discounted at a 3 percent real rate (that is, a 5 percent discount rate with 2 percent wage growth) back to age 12, the mean age in the sample. Under these assumptions, the mean present value of lifetime earnings at age 12 in the U.S. population is approximately $522,000. Thus, the financial value of having a one standard deviation higher value added teacher is 0.9 percent of $522,000, or approximately $4,600 per grade.

[7]The "teenage birth" measure indicates whether a tax return was filed that included a dependent born while the mother was a teenager.

[8]G. Barlevy and D. Neal, "Pay for Percentile," *American Economic Review* 102, No. 5 (2012): 1805–1831.

Effects of value-added policies

Jesse Rothstein

Jesse Rothstein is Associate Professor of Public Policy and Economics at the University of California, Berkeley, and Research Associate at the National Bureau of Economic Research.

It is important to distinguish between two topics that have often been mixed together: (1) the properties of value-added models and (2) the effects of value-added-based policies. Most research to date has focused on the first, nearly always in low-stakes settings, and many researchers and others have drawn strong policy conclusions from that research. But at this point we know very little about the effects of policies that would use value-added scores to make decisions about teachers. That should be the focus going forward. What really matters is not the effect of individual teachers, which is what most research estimates, but the effect of a policy.[1]

What do we know about the properties of value-added models?

A considerable amount of research has been devoted to developing models to estimate the contributions of individual teachers to student achievement. It is important to note that the things we have learned about the properties of value-added models nearly always come from low-stakes settings; that is, the value-added calculations for individual teachers have not generally been used to make decisions about teacher retention or bonuses. While much has been learned, there are still many unanswered questions. I'll review here what I see as a few of the most important outstanding issues.

Value-added measures have been shown to have substantial measurement error, although averaging a few years of data does help. The measures are also sensitive to student assignments. We know that assignment of students to teachers is not random, but it remains an open question whether assignment practices introduce large biases in individual teachers' evaluations. In a paper a few years ago, I showed that the available data were consistent with substantial biases or with essentially no bias.[2] Important papers by Kane and Staiger and Chetty, Friedman, and Rockoff have narrowed the plausible range somewhat.[3] However, both the Kane-Staiger and the Chetty and colleagues estimates have had very wide confidence intervals, so we still do not know the importance of biases due to student assignments.

The Chetty and colleagues study revealed an important fact that has not been incorporated into most thinking about value-added models to date. Specifically, they found that teacher effectiveness changes over time: Some teachers are ineffective at first but improve as they age, while others start better and then burn out. Under a policy that uses value-added measures to fire poor teachers and reward good ones, some teachers fired early for poor student achievement would have improved over time, while some teachers who receive early raises will continue to receive them even if the quality of their teaching declines. Both modeling and policy calculations will need to change to accommodate this fact, which could have important implications for the kinds of cost-benefit analyses that have been done to date (including in the Chetty and colleagues study).

Another unresolved issue is the choice of value-added specifications. Each author tends to focus on his or her preferred value-added model, and it isn't clear how much it matters. An important aspect of this issue is the distinction between within- and between-school comparisons. Researchers typically focus on within-school comparisons, including fixed effects to absorb any between-school differences. There is good reason for this, as while it is barely possible that students are randomly assigned to teachers within schools, it is clearly not the case that students or teachers are randomly assigned to schools. Proposed policy applications of value added, however, will need to make both within- and between-school comparisons. We do not have a consensus about how to do this, nor much evidence about how much it matters.

Finally, Chetty and colleagues show that teacher value added is predictive of students' future wages. However, the strength of this correlation is unknown. If we could measure teachers' impacts on student wages, would we find that their test score impacts (as measured by value added) were good proxies for them? We don't know. We also know very little about the interactions across grades; if a student has a high value-added teacher two years in a row, how should the values be combined to calculate the joint effect? Researchers typically treat the effects as additive, but there's no evidence for this and a good deal of reason to think it is incorrect.

What do we know about the effects of value-added-based policies?

Much less is known about the effects of value-added-based policies than about how to measure the contributions of teachers to student achievement. It is difficult to find studies that show that offering significant bonuses to high value-added teachers in the United States produces significant effects, and some of the highest-quality studies of the issue find no evidence of such effects.[4] And there is essentially no evidence on the effects of policies that use value added for deselection (i.e., firing) or for professional feedback.

A study by Carrell and West provides a cautionary tale: adjunct Air Force academy professors, whose continued employment depends on their measured teaching performance,

outscored their regular faculty peers on value-added-type measures based on end-of-year tests, but their students performed poorly in follow-on classes.[5] These results suggest the potential for teacher responses that improve the teacher value-added measure without improving future student outcomes.

What would we expect to happen if teacher policy is based on value added?

In the absence of extensive evidence on the effects of value-added policies, we can still make an educated guess using a long-standing principle in the education field known as Campbell's Law: "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor."[6] Campbell also states that "achievement tests may well be valuable indicators of general school achievement under conditions of normal teaching aimed at general competence. But when test scores become the goal of the teaching process, they both lose their value as indicators of educational status and distort the educational process in undesirable ways."

Thus, if teachers are told that their jobs depend on having a high value added, we should expect that value added will be high, but also worry that that might come at the cost of teachers not doing things that we would really like them to do, but that are not directly related to value-added scores. For example, since teachers are evaluated based on math and reading scores, they might spend less time teaching subjects that are not covered in achievement tests, such as history. Even within a tested subject, teachers might spend more time on topics that are covered on the test such as analogies, and less on topics that are not such as composition. There is anecdotal evidence that some teachers are unwilling to teach students whom they believe will not improve their value-added score. Teachers might also focus more on short-term learning (such as drills on multiple-choice questions) that is likely to be reflected in test scores, rather than on long-term learning that will serve students better after the tests are done. The Air Force Academy results mentioned above appear to indicate that these kinds of responses can be important.

David Figlio has done a lot of work looking at the unintended effects of school accountability, ranging from suspension of students who are expected to do poorly, to changing the food offered in the cafeteria on test day.[7] There are a great deal of factors that may affect test scores without affecting learning, and this may not be how we want our school resources to be used. We do not currently have a sense of how large these distortions would be, and thus how much they would undermine a policy that was based on value-added measures, but it does appear possible that they could completely negate the effects of a teacher policy based on value added.

Personnel economists have spent years studying incentive compensation, and there are lessons from that field that clearly apply to education. When a task is multidimensional, as teaching certainly is, and when a performance measure is subject to influence, as I believe value added is, it is important to ensure that the stakes are low for a particular measure; that multiple measures be used; that human discretion be part of the process; and finally that the process for helping employees improve be separate from the process through which personnel decisions are made. I believe that describes a viable teacher personnel policy, albeit one that looks quite different from what many districts are implementing. What would it take to implement this kind of policy? First there must be lots of administrators, all highly trained and carefully selected. It seems unreasonable for a single principal to be solely responsible for 40 teachers, accompanying staff, and all other aspects of a given school. While the consulting-world standard of one manager for every five workers is not likely to occur in the world of education, perhaps one administrator for every ten teachers is achievable? It is important that the administrator be capable—there is no reason to think that principal quality is any less important than is teacher quality. We should also be thinking at least as much about the best ways to develop and improve staff, rather than firing them. Finally, there should be an incentive pay component, but stakes need to be relatively low so as not to cause too much distortion of outcomes.■

---

[1]This point was made by D. B. Rubin, E. A. Stuart, and E. L. Zanutto, "A Potential Outcomes View of Value-Added Assessment in Education," *Journal of Educational and Behavioral Statistics*, 29, No. 1 (Spring 2004): 103–116.

[2]See J. Rothstein, "Student Sorting and Bias in Value Added Estimation: Selection on Observables and Unobservables." *Education Finance and Policy*, 4, No. 4 (Fall 2009): 537–571.

[3]See: T. J. Kane and D. O. Staiger, "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation." Working Paper No. 14607, National Bureau of Economic Research, 2008; and R. Chetty, J. N. Friedman, and J. E. Rockoff, "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood," Working Paper No. 17699, National Bureau of Economic Research, 2011. In addition, the Gates Foundation's Measures of Effective Teaching (MET) Project recently released results of a large-scale experiment along the lines of that carried out earlier, on a smaller scale, by Kane and Staiger. Unfortunately, the experiment was plagued by high rates of noncompliance, which limited its ability to answer the question at hand. See: T. J. Kane, D. F. McCaffrey, T. Miller, and D. O. Staiger, *Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment*, MET Project Research Paper, Bill and Melinda Gates Foundation, Seattle, WA, January 2013; and J. Rothstein and W. J. Mathis, Review of *Have We Identified Effective Teachers?* and *A Composite Estimator of Effective Teaching*: "Culminating Findings from the Measures of Effective Teaching Project," National Education Policy Center, Boulder, CO, January 31, 2013.

[4]See M. G. Springer, D. Ballou, L. Hamilton, V. Le, J. R. Lockwood, D. F. McCaffrey, M. Pepper, and B. M. Stecher, *Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching*, National Center on Performance Incentives at Vanderbilt University, Nashville, TN, 2010.

[5]S. E. Carrell and J. E. West, "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," *Journal of Political Economy*, 118, No. 3 (2010): 409–432.

[6]D. T. Campbell, "Assessing the Impact of Planned Social Change," *Evaluation and Program Planning*, 2, No. 1 (1979): 67–90.

[7]See, for example, D. N. Figlio and S. Loeb, "School Accountability," in *Handbook of Economics of Education*, Volume 3, eds. E. Hanushek, S. Machin, and L. Woessmann (The Netherlands: North-Holland, 2001).

Use of value added in teacher policy measures

Eric A. Hanushek

Eric A. Hanushek is Paul and Jean Hanna Senior Fellow at the Hoover Institution of Stanford University, and an IRP affiliate.

I would like to offer a different take on the policy issues related to value-added estimates than that provided in Jesse Rothstein's article. I believe that the primary value of these estimates is in illustrating how much difference there is between teachers. When the estimates are made in low-stakes situations where there is little incentive to teach to the test, estimates of the variance in teacher quality are very precise. In this article, I discuss the implications of the results of these types of studies, and then explore the implications for teacher policy. I believe that where Rothstein's argument falters is that there are not currently any school systems that make teacher personnel decisions solely on the basis of value-added estimates, nor am I aware of any current proposals for such a system. For example, in regard to the District of Columbia policy described by Raj Chetty, only 18 percent of teachers in the system have value-added scores available, so this information is clearly only a relatively small part of what goes into making firing and bonus decisions.

Magnitudes of effects

Estimates of the average standard deviation in gains in student achievement over one year attributable to higher value-added teachers within a given school range from 0.13 to 0.17. Any between-school differences in teacher effectiveness would need to be added on top of this. Although Chetty has already discussed some of the implications of these differences, I will very briefly offer my own calculations.[1] Estimates of the effect of test scores on earnings indicate that a standard-deviation increase in scores translates into a 13 to 20 percent annual increase in earnings. Figure 1 illustrates the effect on student lifetime income by class size and teacher effectiveness, allowing for some depreciation in scores over time. This figure shows the estimated marginal effect, compared to an average teacher, of having a teacher in various percentiles. Calculations for individual students are multiplied by class size. So, for example, the present value at the beginning of high school for a 75th percentile teacher with a class of 30 students is $430,000, while that for a 25th percentile teacher with the same class size is $425,000. These numbers appear large enough to suggest that, although there may be some error in particular teacher personnel policies, having no personnel policy at all cannot be the correct answer.

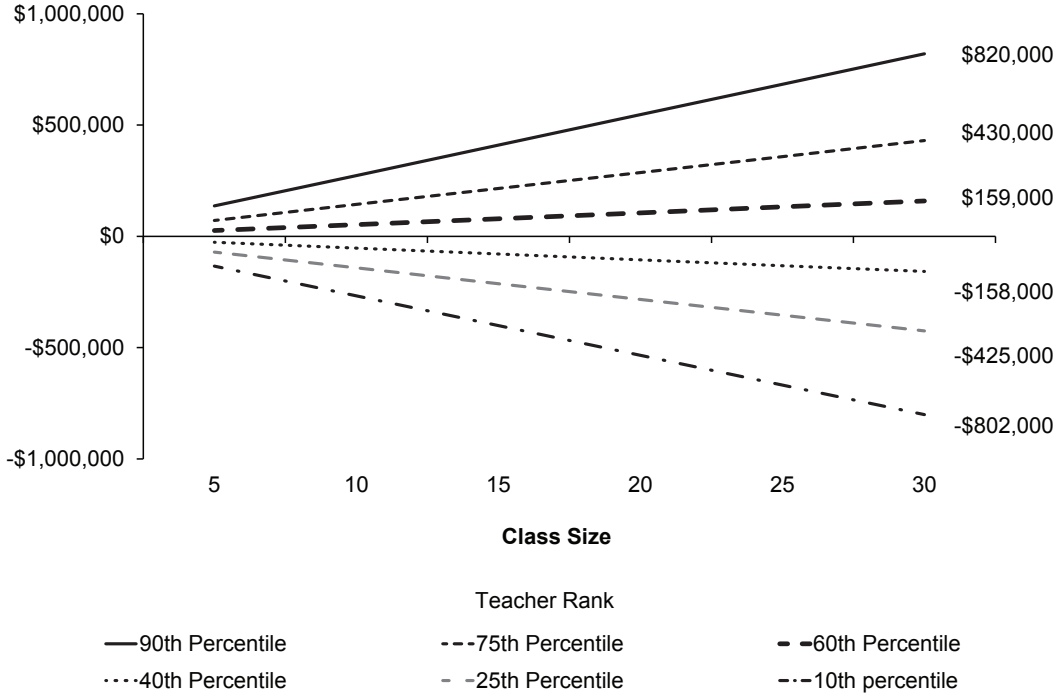

Figure 1. Effect on student lifetime incomes by class size and teacher effectiveness (compared to average teacher).

Source: Calculations by author relying on estimates of teacher quality using 0.2 standard deviations, and reflecting between-school calculations.

School districts have needed to lay off teachers in substantial numbers only quite recently, as a result of the recent recession. The standard policy for determining layoffs is to use teacher seniority. A recent simulation comparing this policy to one that used a measure of effectiveness found some differences between the two approaches.[2] Since seniority-based layoffs generally mean that those with lower salaries are more likely to lose their jobs, more layoffs are required to achieve a given budget reduction. In this simulation, a system based on value-added results in about 25 percent fewer layoffs than one based on seniority. In addition, the typical teacher laid off using a value-added system is less effective than the typical seniority-based layoff, by 26 percent of a standard deviation.

Another mental exercise is to imagine ranking all teachers in the United States based on effectiveness, and look at the performance gains that would result from deselecting some percentage of the lowest-ranked teachers, and replacing them with an average teacher. In this case, unlike the one-year effects that Rothstein estimated, I am looking at lifetime effects. I find that, depending on whether a high or low estimate of teacher effectiveness is used, a deselection rate of between 5 and 8 percent would result in achievement levels similar to that of Canada, a country that currently ranks 0.42 standard deviations above the United States. According to calculations I have made along with Ludger Woessman, such an increase in achievement is worth $72 trillion in GDP.[3] Larger estimates of the variation in teacher effectiveness result in even higher estimates. Although the precise value can certainly be argued, it is clear to me that the value of having policies based on teacher effectiveness is enormously higher than having no policy at all, and that policies based on teacher effectiveness in fact represent the future of the U.S. economy.

Use of value-added measures in teacher personnel policy

There has been a great deal of discussion about errors in estimating value added, and whether it is acceptable to, for example, have a 5 percent error rate in determining which teachers contribute the most to student achievement. I believe that the current state of having no policy translates to a 100 percent error rate, and that we should be striving not for perfection, but for a policy that improves teacher effectiveness overall.

Rothstein discussed some of the implications of making teacher-retention decisions based on imperfect value-added scores. If the rate of dismissal and replacement is somewhere between 5 and 8 percent, that translates to 2 to 3 individuals in a school of 30 teachers. I have found in all of my dealings with teachers, administrators, parents, and staff in numerous schools, that there is very little uncertainty about who the 2 to 3 least-effective teachers in any given school are. I believe that an evaluation process that allowed decisions based on this type of common knowledge would not necessarily need to depend on value-added data that might not be available in a timely manner, and that the evidence suggests that such a policy would likely result in substantial gains in student achievement.

As has been mentioned, both the *Los Angeles Times* and the *New York Times* have recently published teacher value-added scores for their respective school districts. This was extremely controversial, and the aftershocks are still being felt. I was one of the few researchers to support the idea of publishing value-added scores, not because I think that personnel policy should be done through newspapers, but because within a week of these publications, unions and school officials were meeting to discuss teacher-evaluation policy. This is an issue that had been on the agenda forever with no progress. It seems that providing a strict value-added ranking as one (extreme) option prompts people to develop better personnel systems that incorporate other teacher-evaluation tools, and this is exactly what is needed.

Issues and areas for further study

One could ask whether the currently available achievement tests are really up to the task of providing reliable value-added scores. I would say certainly not, and that value-added measures should never be the sole basis for personnel decisions. Rothstein also raised the possibility that value-added measures can become less reliable when used for consequential purposes. While this and the accompanying loss in reliability and validity is certainly possible, I believe such problems can be dealt with in feasible ways.

On the question of whether value-added measure can be used to rate principals, I agree with Rothstein that a parallel system is required. There are some indications that reliable value-added measures can be constructed. Preliminary estimates from work that I have been involved in suggest that principal quality is extremely important and that a one standard deviation increase in principal quality results in an increase of approximately 0.05 standard deviations in average student growth.[4] While this effect is much smaller than that seen for teachers within a given school, principals affect all students in a school, so an increase in principal quality will have effects much greater than a similar increase in the quality of a single teacher.■

---

[1]E. A. Hanushek, "The Economic Value of Higher Teacher Quality," *Economics of Education Review*, 30, No. 3 (June 2011): 466-479.

[2]D. Boyd, H. Lankford, S. Loeb, and J. Wyckoff, "Teacher Layoffs: An Empirical Illustration of Seniority versus Measures of Effectiveness," *Education Finance and Policy*, 6, No. 3 (Summer 2011): 439–454. The simulation was conducted using fourth- and fifth-grade math and language arts achievement scores for students in New York City public schools.

[3]E. A. Hanushek and L. Woessmann, *The high cost of low educational performance: The long-run economic impact of improving PISA outcomes* (Paris: Organisation for Economic Cooperation and Development, 2010).

[4]G. F. Branch, E. A. Hanushek, and S. G. Rivkin, "Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals," NBER Working Paper No. 17803, National Bureau of Economic Research, 2012.

This is the last time we will be printing and mailing *Focus*

To reduce the environmental impact and production costs of *Focus*, we are discontinuing printed copies, and distributing it solely as an electronic publication. **This will be the last printed issue of *Focus*.** If you would like to receive electronic notification of future *Focus* issues, we need to have your e-mail address.*

If you do not already receive "IRPFocusAlert" e-mail notifications (you should have received one about this issue), please send a message to irppubs@ssc.wisc.edu with the phrase "FOCUS ELECTRONIC" in the subject line. Please include the following in the body of the message:

Name

Mailing address

E-mail address

In addition to receiving an e-mail and link when a new issue of *Focus* is available on our website, IRPFocusAlert subscribers also receive notification of each new issue of *Fast Focus*, our electronic-only brief on recent poverty research.

We welcome your input. If you have questions or comments, please e-mail *Focus* editor Emma Caspar at ecaspar@ssc.wisc.edu.

*Rest assured that we will not share your e-mail address with anyone.

Focus

Institute for Research on Poverty
University of Wisconsin–Madison
1180 Observatory Drive
3412 Social Science Building
Madison, Wisconsin 53706

Read *Focus*
on the Web

http://www.irp.wisc.edu/