



Hard evidence on soft skills[☆]

James J. Heckman^{a,b,c,*}, Tim Kautz^a

^a Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637, USA

^b American Bar Foundation, 750 North Lake Shore Drive, Chicago, IL 60611, USA

^c University College Dublin, Ireland

ARTICLE INFO

Article history:

Received 10 May 2012

Accepted 13 May 2012

Available online 25 May 2012

JEL classification:

I20

D01

Keywords:

Personality

Achievement tests

IQ

Cognition

ABSTRACT

This paper summarizes recent evidence on what achievement tests measure; how achievement tests relate to other measures of “cognitive ability” like IQ and grades; the important skills that achievement tests miss or mismeasure, and how much these skills matter in life.

Achievement tests miss, or perhaps more accurately, do not adequately capture, *soft skills*—personality traits, goals, motivations, and preferences that are valued in the labor market, in school, and in many other domains. The larger message of this paper is that soft skills predict success in life, that they causally produce that success, and that programs that enhance soft skills have an important place in an effective portfolio of public policies.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Contemporary society places great value on standardized achievement tests to sift and sort people, to evaluate schools, and to assess the performance of nations. Admissions committees use tests like the SAT, the ACT, and the GRE (Graduate Record Examinations) to screen applicants. In the United States, the No Child Left Behind (NCLB) Act stipulates that government-run schools must administer standardized achievement tests in order to be eligible for federal funding.¹ The Programme for International Student Assessment (PISA) evaluates student

performance in math, science, and reading across countries. The results attract media attention and influence policy. The year 2000 PISA test results caused Germany to re-evaluate its educational system and introduce a variety of educational reforms (Grek, 2009).

Despite the widespread use of standardized achievement tests, the traits that they measure are not well-understood. This paper summarizes recent evidence on what achievement tests capture; how achievement tests relate to other measures of “cognitive ability” like IQ and grades; the important skills that achievement tests miss or mismeasure, and how much these other skills matter in life.

Achievement tests miss, or more accurately, do not adequately capture, *soft skills*—personality traits, goals, motivations, and preferences that are valued in the labor market, in school, and in many other domains. The larger message of this paper is that soft skills predict success in life, that they produce that success, and that programs that enhance soft skills have an important place in an effective portfolio of public policies.²

Measurement of cognition and educational attainment has been refined during the past century. Psychometricians have shown that cognitive ability has multiple facets.³ This progress is not widely appreciated. Many social scientists—even many psychologists—continue to use IQ tests, standardized achievement tests, and grades interchangeably to proxy “cognitive ability.”⁴ Even though scores on IQ tests, standardized

[☆] This paper was presented as the Adam Smith Lecture at the Annual Meeting of the European Association of Labour Economists held in Cyprus, September 2011. This research was supported in part by the University of Chicago, A New Science of Virtues: A Project of the University of Chicago, the American Bar Foundation, a conference series from the Spencer Foundation, the JB & MK Pritzker Family Foundation, Susan Thompson Buffett Foundation, the Geary Institute, University College Dublin, Ireland, NICHD R37 HD065072 and R01 HD054702, a European Research Council grant hosted by University College Dublin, DEVHEALTH 269874, a grant to the Becker Friedman Institute for Research and Economics from the Institute for New Economic Thinking (INET), and an anonymous funder. The views expressed in this paper are those of the authors and not necessarily those of the funders or commentators mentioned here. We thank Linor Kiknadze for research assistance. Paul Devereux made helpful comments that are incorporated in this paper.

* Corresponding author at: Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637, USA. Tel.: +1 773 702 0634; fax: +1 773 702 8490.

E-mail addresses: jjh@uchicago.edu (J.J. Heckman), tkautz@uchicago.edu (T. Kautz).

¹ Sales of achievement tests have increased by nearly 400% between 1959 and 2005 (Digest of Education Statistics, various years; The Bowker Annual: Library and Book Trade Almanac, various years).

² This paper draws on and supplements Borghans et al. (2008a), Almlund et al. (2011), and Heckman et al. (2012a).

³ See Carroll (1993) and Ackerman and Heggestad (1997) for a discussion.

⁴ Many call this “IQ”, e.g., Flynn (2007), Nisbett (2009), and Nisbett et al. (2012).

achievement tests, and grades are positively correlated with each other, the recent literature shows that they measure different skills and depend on different facets of cognitive ability. Recent research also shows that all three measures are associated with personality, but to different degrees across various cognitive measures.

Standardized achievement tests were designed to capture “general knowledge” produced in schools and through life experiences. Such knowledge is thought to be relevant to success inside and outside of the classroom. However, achievement tests are often validated using other standardized achievement tests or other measures of cognitive ability—surely a circular practice.

A more relevant validity criterion is how well these tests predict meaningful outcomes, such as educational attainment, labor market success, crime, and health. No single measure of cognitive ability predicts much of the variance in these outcomes, and measurement error does not account for most of the remaining variance, leaving much room for other determinants of success.⁵

Success in life depends on personality traits that are not well captured by measures of cognition. Conscientiousness, perseverance, sociability, and curiosity matter. While economists have largely ignored these traits, personality psychologists have studied them over the last century.⁶ They have constructed measures of them and provide evidence that these traits predict meaningful life outcomes.

Many scholars—inside and outside of psychology—have questioned the existence of stable personality traits, arguing that constraints and incentives in situations almost entirely determine behavior. These scholars claim that people are like chameleons—they adapt to any situation.⁷ A substantial body of evidence shows that stable traits exist. People tend to behave in the same fashion across a wide range of situations.⁸ Evidence from genetics and neuroscience provides a biological basis for the existence of such traits, suggesting that something tied to the person, not the just the situation, affects behavior.⁹

Throughout this paper we use the term “personality traits” to describe the personal attributes not thought to be captured by measures of abstract reasoning power. These attributes go by many names in the literature, including soft skills, personality traits, noncognitive skills, noncognitive abilities, character, and socioemotional skills. These different names connote different properties. The term “traits” suggests a sense of permanence and possibly also of heritability. The terms “skills” and “character” suggest that they can be learned. In reality, the extent to which these personal attributes can change lies on a spectrum. Both cognitive and personality traits can change and be changed over the life cycle but through different mechanisms and to different degrees at different ages. To avoid confusion, throughout this paper we use the term “trait” to capture the set of personal attributes we study.¹⁰

Psychological traits are not directly observed. There is no ruler for perseverance, no caliper for intelligence. All cognitive and personality traits are measured using performance on “tasks,” broadly defined. Different tasks require different traits in different combinations. Some distinguish between measurements of traits and measurements of

outcomes, but this distinction is misleading. Both traits and outcomes are measured using performance on some task or set of tasks.

Psychologists sometimes claim to circumvent this measurement issue by creating taxonomies of traits and by applying intuitive names to responses on questionnaires. These questionnaires are not windows to the soul. They are still rooted in task performance or behavior. Responding to a questionnaire is itself a task. Additionally, many of the questionnaires inquire directly about behavior, e.g., a measure of Agreeableness used in the German Socioeconomic Panel asks the extent to which a respondent “is sometimes somewhat rude to others.”¹¹ How else can one answer that question but reflect on one’s behavior? IQ tests and standardized achievement tests also measure performance on different “cognitively loaded” tasks.

Performance on most tasks depends on effort, personality traits, cognitive ability, and incentives, although the importance of each differs by task. This dependence creates a fundamental problem in measuring traits. Most studies in psychology devise a set of measures to capture a trait but do not standardize for incentives in the situation in which the trait is being measured or for other traits. Measured cognitive ability and measured personality depend on a constellation of factors. The identification problem arising from the multiple determinants of performance on tasks is empirically important, even for measures of cognitive ability. Incentives can affect performance on IQ tests. Multiple traits affect performance on cognitive tasks. For example, personality traits affect achievement test scores and grades.¹² Caution is required in taking the measures developed by psychologists too literally.

Nonetheless, measures of personality traits predict meaningful life outcomes. Conscientiousness—the tendency to be organized, responsible, and hardworking—is the most widely predictive of the commonly used personality measures. It predicts educational attainment, health, and labor market outcomes as strongly as measures of cognitive ability.¹³

Most studies in psychology only report correlations between measured traits and outcomes without addressing whether the traits *cause* the outcomes and without controlling for the other traits and incentives that determine performance on the tasks used to measure the traits. While traits are relatively stable across situations, they are not set in stone. They change over the life cycle. On average, Agreeableness and Conscientiousness tend to grow with age. Different facets of cognitive ability peak at different ages. Interventions, education, and parenting can affect traits in lasting ways.

This paper summarizes recent evidence that personality causally affects life outcomes. We review some of the literature from psychology and economics and then focus on two particularly compelling examples.¹⁴

First, we show how an achievement test, the General Educational Development (GED) test, fails to capture important traits that affect success in life. High school dropouts can take the GED to certify to employers and post-secondary institutions that their skills are equivalent to those of high school graduates who do not attend college. After accounting for differences in pre-existing cognitive ability, GED recipients perform much worse in the labor market than high school graduates and much more like other high school dropouts. GED recipients lack important personality traits (see Heckman et al., 2011a, 2012a).

Second, we show how an early childhood intervention, the Perry Preschool Program, improved the lives of disadvantaged children, even though the program did not permanently change the IQ of its participants. The program changed their personality traits in a lasting way (see Heckman et al., 2012b). Other interventions and observational

⁵ On the magnitudes of measurement error on a variety of economic measures, see Bound et al. (2001). These authors report that at most 15–30% of earnings variance is due to measurement error.

⁶ Some early studies in economics are Bowles and Gintis (1976), and Bowles et al. (2001). An important study in sociology is Jencks (1979). Work in psychology going back to Terman et al. (1925) shows that personality traits matter (Murray, 1938; Terman et al., 1947; and the discussion in Gensowski, 2012).

⁷ See Mischel (1968). Some behavioral economists share this view. See, e.g., Thaler (2008).

⁸ See Epstein (1979) for an early paper showing that personality traits are stable across multiple situations. See the special issue of *Journal of Research in Personality* (43), “Personality and Assessment at Age 40,” for a more recent discussion.

⁹ See Bouchard and Loehlin (2001) for estimates of the heritability of traits. See Canli (2006) and DeYoung et al. (2010) for evidence that regions of the brain are associated with different traits.

¹⁰ Drawing on the literature in psychology, Borghans et al. (2008a) present one definition of cognitive traits.

¹¹ Throughout this paper, we adopt the convention of capitalizing traits from the “Big Five” personality taxonomy. See Table 3 for a description of the Big Five.

¹² See Borghans et al. (2011a).

¹³ See the evidence collected in Almlund et al. (2011), Borghans et al. (2008a), and Roberts et al. (2007).

¹⁴ Borghans et al. (2008a) and Almlund et al. (2011) present extensive surveys of this literature.

Table 1
Cognitive ability validities.

Test	Validation domain	Estimate(s)	Source(s)
SAT (Achievement)	1st Year College GPA	0.35–0.53	Kobrin et al. (2008)
ACT (Achievement)	Early College GPA	0.42	ACT, Inc. (2007)
GED (Achievement)	HS Senior GPA	0.33–0.49	GED Testing Service (2009)
DAT (Achievement)	College GPA	0.13–0.62 [†]	Omizo (1980)
AFQT (Achievement)	9th Grade GPA	0.54	Borghans et al. (2011a)
WAIS (IQ)	College GPA	0.38–0.43	Feingold (1982)
WAIS (IQ)	HS GPA	0.62	Feingold (1982)
Various IQ**	9th Grade GPA	0.42	Borghans et al. (2011a)
WISC (IQ)	WRAT (Achievement)	0.44–0.75 [‡]	Hartlage and Steele (1977)
WISC-R (IQ)	WRAT (Achievement)	0.35–0.76 [‡]	Hartlage and Steele (1977)
Various IQ**	AFQT (Achievement)	0.65	Borghans et al. (2011a)
Stanford Binet (IQ)	WISC-R (IQ)	0.77–0.87	Rothlisberg (1987), Greene et al. (1990)
Raven's (IQ)	WAIS-R (IQ)	0.74–0.84	O'Leary et al. (1991)
WIAT (Achievement)	CAT/2 (Achievement)	0.69–0.83*	Michalko and Saklofske (1996)

Definitions: WISC—Wechsler Intelligence Scale for Children, WISC-R—Wechsler Intelligence Scale for Children—Revised, WAIS - Wechsler Adult Intelligence Scale, Raven's IQ—Raven's Standard Progressive Matrices, GED—General Educational Development, DAT—Differential Aptitude Tests, WIAT—Wechsler Individual Achievement Test, CAT—California Achievement Test, WRAT—Wide Range Achievement Test, AFQT - Armed Forces Qualification Test.

[†] Large range is due to varying validity of eight subtests of DAT.

[‡] Ranges are given because correlations vary by academic subject.

* Ranges are given because correlations vary by grade level.

** IQ test scores in the NLSY79 are pooled across several IQ tests using IQ percentiles.

studies provide supporting evidence that early-childhood investments improve outcomes through their effects on personality.¹⁵

2. Defining and measuring personality traits

2.1. History and measurement of cognitive ability

Modern intelligence tests have been used for just over a century, beginning when a French minister of public instruction wished to identify retarded pupils in need of specialized education programs. In response, Alfred Binet created the first IQ test.¹⁶ IQ scores were interpreted as measuring a stable trait. The standardized achievement test was created in the wake of the perceived success of IQ tests as an objective and cost-effective measure of acquired skills. In contrast to IQ tests, standardized achievement tests were designed to measure “general knowledge” that could be acquired in schools and through life experiences and was widely applicable beyond the classroom to workplace and social functioning.¹⁷

Achievement tests are typically validated on other achievement tests, IQ tests, and grades, rather than on tasks or outcomes in the labor market and in social functioning. Table 1 shows correlations among scores on standardized achievement tests, IQ tests, and grades. Standardized achievement tests are correlated with IQ tests, but the correlation depends on the subject area of the standardized achievement test. Hartlage and Steele (1977) find that the arithmetic portions of standardized achievement tests are the most highly correlated with IQ. Grades and scores on IQ tests and standardized achievement tests are far from perfectly correlated, suggesting that they measure different aspects of “cognitive functioning.”¹⁸

Psychologists distinguish between fluid intelligence (the rate at which people learn) and crystallized intelligence (acquired knowledge).¹⁹ Achievement tests are heavily weighted towards crystallized intelligence,²⁰ whereas IQ tests like Raven's progressive matrices (1962) are heavily weighted toward fluid intelligence.^{21,22} Many psychologists do not recognize the differences among these measures and interchangeably use IQ, achievement tests, and grades to measure “cognitive ability” or “intelligence,” and this practice is also widespread in economics.²³

2.2. Defining and measuring psychological traits

Validating one measure of cognitive ability using other measures of cognitive ability is circular. More relevant is how well these measures predict important life outcomes. Table 2 shows the extent to which IQ, standardized achievement tests, and grades explain the variance of outcomes at age 35 in the National Longitudinal Survey of Youth, 1979 (NLSY79) data. The three groups of columns under each category show results for different sub-samples based on the availability of the different cognitive measures. For each category, the first column shows the explained variance using only the designated measure of cognitive ability. Achievement tests and grades are more predictive than IQ. But none of these measures explains much of the variation of any outcome, leaving considerable room for other determinants. As noted in the introduction, it is unlikely that measurement error accounts for all of the remaining variance.

Personality is one missing ingredient. The second columns in each category preview our later discussion of the explanatory power of personality. They show the variance explained by measures of personality.²⁴

¹⁵ The “Tools of the Mind” intervention is designed to promote “executive functioning,” which has both cognitive and personality components. Barnett et al. (2006, 2008), Bierman et al. (2010), Bodrova and Leong (2001, 2007), Dee and West (2011), Diamond et al. (2007), Durlak et al. (2011), Lillard and Else-Quest (2006) report success of this intervention. For a contrary view, see the study by Farran et al. (2011).

¹⁶ In 1904, *La Société Libre pour l'Etude Psychologique de l'Enfant* appointed a commission to create a mechanism for identifying these pupils in need of alternative education led by Binet. See Hermstein and Murray (1994) for an overview of Binet's life and work.

¹⁷ See Lindquist (1951).

¹⁸ It is an irony of the testing literature that high school grades are more predictive of first year college performance than SAT scores (Bowen et al., 2009). The SAT and related tests are thought to be more objective measures of student quality than high school grades (Lemann, 1999).

¹⁹ See, e.g., Nisbett et al. (2012).

²⁰ See Roberts et al. (2000).

²¹ See Raven et al. (1988). The high correlation between intelligence and achievement tests is in part due to the fact that both require cognitive ability and knowledge. Common developmental factors may affect both of these traits. Fluid intelligence promotes the acquisition of crystallized intelligence.

²² Carroll (1993) and Ackerman and Heggstad (1997) discuss more disaggregated facets of cognitive ability.

²³ See Flynn (2007) and Nisbett et al. (2012). For examples in economics, see Benjamin et al. (2006).

²⁴ They include measures of adolescent risky behavior, self-esteem and locus of control (the extent to which people feel they have control over their lives). For precise definitions of the measures used, see the notes to Table 2.

Table 2
Predictive validities in outcomes that matter (adjusted R-squared).

	IQ Sample			AFQT Sample			GPA Sample		
	IQ	Personality	Both	AFQT	Personality	Both	GPA	Personality	Both
Males									
Earnings at age 35	0.07	0.05	0.09	0.17	0.07	0.18	0.09	0.06	0.12
Hourly wage at age 35	0.07	0.03	0.08	0.13	0.06	0.14	0.07	0.06	0.09
Hours worked at age 35	0.01	0.03	0.04	0.03	0.02	0.03	0.02	0.01	0.02
Jail by age 35	0.03	0.02	0.04	0.06	0.06	0.09	0.03	0.03	0.04
Welfare at age 35	0.01	0.00	0.01	0.03	0.01	0.03	0.01	0.00	0.01
Married at age 35	0.01	0.05	0.05	0.04	0.03	0.06	0.03	0.03	0.04
B.A. degree by age 35	0.12	0.08	0.16	0.19	0.10	0.22	0.14	0.10	0.18
Depression in 1992	0.01	0.05	0.05	0.04	0.04	0.06	0.02	0.04	0.04
Adj, R ² Cog, personality	0.07			0.17			0.11		
Females									
Earnings at age 35	0.01	0.03	0.03	0.09	0.05	0.11	0.05	0.04	0.07
Hourly wage at age 35	0.05	0.03	0.06	0.12	0.05	0.14	0.06	0.04	0.08
Hours worked at age 35	0.00	0.02	0.02	0.00	0.01	0.00	0.00	0.01	0.01
Jail by age 35	0.00	0.01	0.00	0.01	0.02	0.02	0.01	0.01	0.02
Welfare at age 35	0.02	0.04	0.05	0.10	0.05	0.12	0.05	0.05	0.07
Married at age 35	0.03	0.03	0.05	0.05	0.04	0.07	0.03	0.03	0.05
B.A. degree by age 35	0.10	0.08	0.14	0.17	0.09	0.20	0.10	0.08	0.13
Depression in 1992	0.02	0.05	0.05	0.04	0.05	0.07	0.02	0.05	0.05
Adj, R ² Cog, personality	0.10			0.15			0.10		

Source: National Longitudinal Survey of Youth 1979. Table Description: The table shows the adjusted R-squared from regressions of later-life outcomes on measures of personality and cognition. For each cognitive measure, the first column shows the explained variance using only the measures of cognitive ability, the second column shows the explained variance from using only the measure of personality (Personality), and the third column shows the explained variance from using both the measures of personality and cognition (Both). The last row shows the adjusted R-squared from a regression of each cognitive measure on the personality measures. Measures of Personality and Cognition: The measures of personality include minor illegal activity in 1979 (vandalism, shoplifting, petty theft, fraud and fencing), major illegal activity in 1979 (auto theft, breaking/entering private property, grand theft), participation in violent crime in 1979 (fighting, assault and aggravated assault), tried marijuana before age 15, daily smoking before age 15, regular drinking before age 15, and any intercourse before age 15. It also includes measures of self-esteem and locus of control. Self-esteem is measured using the ten-item Rosenberg scale administered in 1980. Locus of control is a measure of how much control an individual believes they have over their life and is measured using the 4-item Rotter scale. IQ and grades are from high school transcripts. IQ is pooled across several IQ tests using IQ percentiles. GPA is the individual's core-subject GPA from 9th grade. The Armed Forces Qualification Test (AFQT) was adjusted for schooling at the time of the test conditional on final schooling as described in Hansen et al. (2004). Outcomes: Due to the biennial nature of the survey after 1994, some respondents are not interviewed at age 35, for these individuals age 36 is used. Earnings include zero-earners and excludes observations over \$200,000 (2005 dollars). Hourly wage excludes observations less than \$3 or over \$200 (2005 dollars). Hours worked excludes observations less than 80 or more than 4000. Jail by age 35 indicates whether the respondent had listed residing in a jail or prison at some point before age 35. Welfare at age 35 indicates whether the respondent received any positive amount of welfare at age 35. Married at age 35 indicates whether the respondent was currently married. B.A. degree by age 35 indicates whether the respondent received a B.A. degree (or higher) by age 35. Depression in 1992 is based on the 7-item Center for Epidemiologic Studies Depression Scale (CES-D). Sample: The sample excludes the military oversample. The samples differ across the IQ, AFQT, and GPA due to missing measures across the samples.

In many cases, the variance explained by personality measures rivals that explained by measures of cognitive ability. The relative importance of personality depends on the outcome. The third column for each subsample shows the variance explained when both the cognitive and personality measures are used as predictors. In many cases, including the measures of personality in a regression with cognitive measures explains additional variance. The correlations between the set of measures of personality and the measures of cognition are positive, but not especially strong (see the bottom row of each table). Each set of traits has an independent influence on the outcomes in the table.

Even though economists have largely ignored personality traits, the pioneers of the original IQ tests recognized their importance.²⁵ Alfred Binet the creator of the first IQ test (the Stanford-Binet test), noted that:

"[Success in school] ...admits of other things than intelligence; to succeed in his studies, one must have qualities which depend on attention, will, and character; for example a certain docility, a regularity of habits, and especially continuity of effort. A child, even if intelligent, will learn little in class if he never listens, if he spends his time in playing tricks, in giggling, in playing truant."
-(Binet and Simon, 1916, p. 254)

Since the middle of the 19th century, personality psychologists have studied these traits. One leading personality psychologist defines personality traits in the following way:

"Personality traits are the relatively enduring patterns of thoughts, feelings, and behaviors that reflect the tendency to respond in certain ways under certain circumstances."-(Roberts, 2009, p. 140)

Personality traits are manifested through thoughts, feelings, and behaviors, and therefore, must be inferred from measures of performance on "tasks," broadly defined. Under this definition, performance on IQ tests is a personality trait because it is an enduring pattern of behavior (how one "behaves" or "performs" on an IQ test).²⁶

Personality psychologists primarily measure personality traits using self-reported surveys. They have arrived at a relatively well-accepted taxonomy of traits called the "Big Five," which includes Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Table 3 defines these traits and their multiple facets.²⁷ Some argue that the Big Five are the longitude and latitude of personality, by which all more narrowly defined traits may be categorized (see Costa and McCrae, 1992a). While the Big Five measures are now the most widely used, there are several other taxonomies, including the Big Three, the MPQ, and the Big Nine. They are conceptually and

²⁵ Lewis Terman, who created the Stanford-Binet test, even collected data on personality traits of a high-ability sample. In this sample, Conscientiousness is highly predictive of health and earnings (Savelyev, 2011; Gensowski, 2012).

²⁶ Studies of test-retest reliability of IQ tests show that scores are highly correlated across repeated testing occasions (see, e.g., Níolón, 2005).

²⁷ See, e.g., Borghans et al. (2008a).

Table 3
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
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, Activity*
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, Curiosity
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*, activity*, positive emotionality, and sociability/affiliation
Agreeableness	"the tendency to act in a cooperative, unselfish manner"	Trust (forgiving), Straight-forwardness (not demanding), Altruism (warm), Compliance (not stubborn), Modesty (not show-off), and Tender-mindedness (sympathetic)	Empathy, Perspective taking, Cooperation, and Competitiveness	Irritability*, 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*, irritability*, frustration (lack of) soothability, Sadness

Notes: Facets specified by the NEO PI-R personality inventory (Costa and McCrae, 1992b). Trait adjectives in parentheses from the Adjective Check List (Gough and Heilbrun, 1983).

*These temperament traits may be related to two Big Five factors.

Source: Table adapted from John and Srivastava (1999).

empirically related to the Big Five.²⁸ Other taxonomies, including psychopathology as measured by the DSM IV and measures of temperament, have also been related to the Big Five.²⁹ Almlund et al. (2011) show that economic preference parameters are not all that closely related to psychological traits and apparently represent different traits that, along with the psychological traits, govern behavior.

A deeper issue, as yet not systematically investigated in the literature in economics or psychology, is whether the traits captured by the alternative measurement systems are the manifestation of a deeper set of preferences or goals. Achieving certain goals requires certain traits, e.g., a surgeon has to be careful and intelligent; a salesman has to be outgoing and engaging and so forth, etc. Under this view, traits are developed through practice, investment, and habituation. The deeper traits may be the preference parameters that generate the manifest traits. The apparent stability of expressed traits across situations may be a consequence of the stability of the goals and incentives to achieve these goals.³⁰

2.3. Identification problems in measuring traits

Measuring traits is difficult, because, as suggested by Roberts' definition of personality, all psychological measurements are calibrated on measured behavior, and the behaviors used to measure one trait can be influenced by incentives and other traits. To infer traits from behaviors requires standardizing for all of the other contributing factors that produce the observed behavior. The inability to parse and localize behaviors that depend on a single trait or ability gives rise to a fundamental identification problem that is typically ignored in empirical research investigating how psychological traits affect outcomes.³¹

There are two primary issues. First, behavior depends on incentives created in situations. Different incentives elicit different amounts of effort on the tasks used to measure traits. Accurately measuring personality traits requires standardizing for the effort applied in any task. Second, behavior in one task can depend on multiple traits. Not standardizing for incentives and other traits can produce misleading estimates of any trait.

These identification problems are empirically important when measuring any given trait. For example, incentives partly determine scores on IQ tests, even though some have argued that performance on IQ tests reflects maximal effort.³² A series of studies conducted over the past 40 years shows that incentives, like money or candy, can increase IQ scores, particularly among low-IQ individuals. The Black-White gap in IQ can be completely eliminated by incentivizing students with M&M candies.³³ The incentives in one test do not affect performance on future tests.

The recent literature shows that personality traits are associated with standardized achievement test scores, which many analysts use interchangeably with IQ scores.³⁴ Figs. 1 and 2 show how the variance in the scores on two achievement tests, the Armed Forces Qualification Test (AFQT) and the closely related Differential Aptitudes Tests (DAT),³⁵ are decomposed into IQ and personality measures. Personality traits explain a substantial portion of the variances in both AFQT scores and DAT

²⁸ See Borghans et al. (2008a) and Almlund et al. (2011) for a comparison of these taxonomies.

²⁹ See, e.g., Cloninger et al. (1999).

³⁰ McAdams and Pals (2006) adds goals to the list of possible traits. Almlund et al. (2011) develop a model in which preferences and traits determine the effort applied to tasks.

³¹ See Borghans et al. (2011a) and Almlund et al. (2011).

³² A leading psychometrician, Carroll (1993), does not accept the notion that IQ captures maximal effort.

³³ See Ayllon and Kelly (1972), Borghans et al. (2008b), Breuning and Zella (1978), Clingman and Fowler (1976), Edlund (1972), Holt and Hobbs (1979), Larson et al. (1994), Segal (2008). This evidence is summarized in Borghans et al. (2008a) and Almlund et al. (2011).

³⁴ See, e.g., Nisbett (2009).

³⁵ The correlation between DAT and AFQT scores in the National Longitudinal Study of Youth 1979 (NLSY79) is 0.76 to 0.80 (Borghans et al., 2011b). Friedman and Streicher (1985) estimate correlations between 0.65 and 0.82 for in a sample of high school sophomores and juniors. Kettner (1976) estimates correlations between DAT and the AFQT subtests of 0.76 to 0.89 in a sample of juniors and seniors.

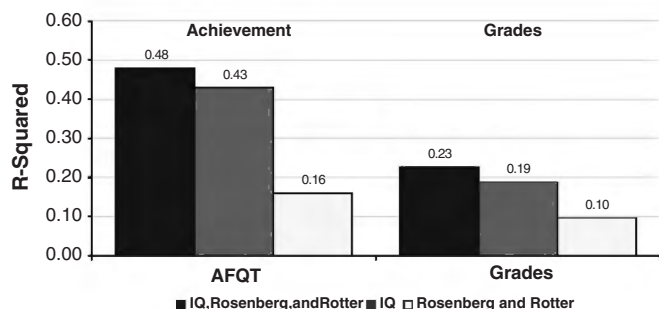


Fig. 1. Decomposing achievement tests and grades into IQ and personality [NLSY79]. Source: Borghans et al. (2011a), National Longitudinal Survey of Youth 1979 (NLSY79). Notes: Rotter is a measure of locus of control designed to measure the extent to which individuals believe that they have control over their lives through self-motivation or self-determination as opposed to the extent to which individuals believe that the environment controls their lives (Rotter, 1966). Rosenberg is a measure of self-esteem designed to measure the degree of approval or disapproval toward oneself (Rosenberg, 1965). The Armed Forces Qualification Test (AFQT) score is constructed from the Arithmetic Reasoning, Word Knowledge, Mathematical Knowledge, and Paragraph Comprehension Armed Services Vocational Aptitude Battery (ASVAB) subtests. Rotter was administered in 1979. The ASVAB and Rosenberg were administered in 1980. IQ and GPA are from high school transcript data. AFQT, Rosenberg, and Rotter have been adjusted for schooling at the time of the test conditional on final schooling, as described in Hansen et al. (2004). IQ is pooled across several IQ tests using IQ percentiles. GPA is the individual's core subject GPA from 9th grade. Sample excludes the military oversample.

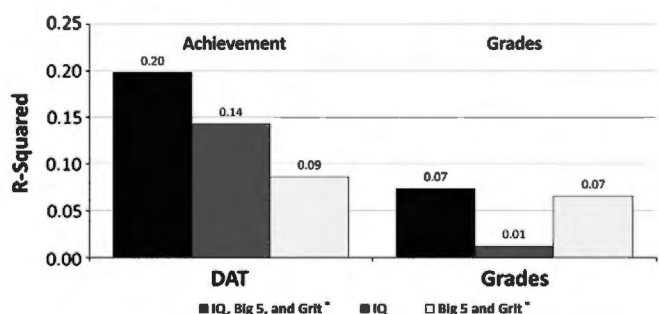


Fig. 2. Decomposing achievement tests and grades into IQ and personality [Stella Maris]. *Grit is a measure of persistence on tasks (Duckworth et al., 2007). Source: Borghans et al. (2011a).

scores.³⁶ The personality traits are incrementally valid in that they explain the variance above and beyond the variance that IQ explains in a regression. These findings caution the interpretation that standardized achievement tests only measure cognitive ability. They are bundled with personality traits. In data from the Stella Maris secondary school in Maastricht, Holland, Openness to Experience is strongly correlated with IQ.³⁷

Further complicating identification, not everyone responds to incentives in the same way. Borghans et al. (2008b) show that adults spend substantially more time answering questions on IQ tests when rewards are higher. Subjects high in Emotional Stability and Conscientiousness are less affected by rewards. Similarly, Segal (2008) shows that introducing cash incentives for performance on the coding speed test of the Armed Services Vocational Battery (ASVAB) increases performance substantially, particularly for men with lower levels of Conscientiousness.

2.4. Are there stable personality traits?

Many have questioned whether personality traits exist. The publication of Walter Mischel's (1968) book, *Personality and Assessment*,

³⁶ The lower explained variance in the sample with DAT is likely a consequence of restriction on range. The DAT data come from a single school, whereas the AFQT data come from a national sample.

³⁷ See Borghans et al. (2011b) for information on the Stella Maris secondary school and the analysis described in the text.

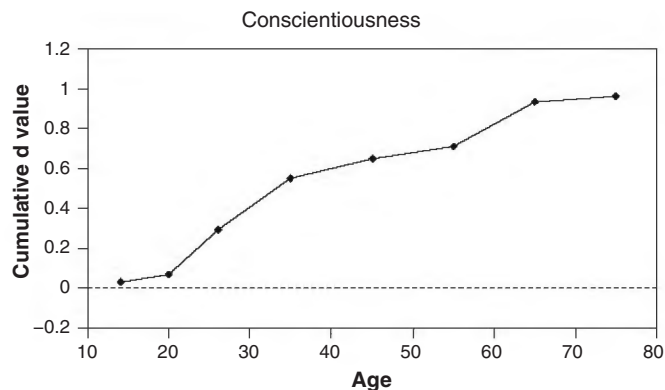


Fig. 3. Cumulative mean-level changes in personality across the life cycle. Note: Cumulative d values represent total lifetime change in units of standard deviations ("effect sizes"). Source: Figure taken from Roberts et al. (2006) and Roberts and Mroczek (2008). Reprinted with permission of the authors.

gave rise to a heated "personality-situation" debate within psychology, which pitted the social psychologists who favored situational factors as explaining behavior against those who considered stable personality traits as more consequential. Mischel argued that aspects of situations overshadow any effect of personality on behavior.³⁸

A large body of evidence reviewed in Almlund et al. (2011) shows that stable personality traits exist and are predictive of many behaviors.³⁹ An important paper by Epstein (1979) presents compelling evidence that, averaging over tasks and situations at a point in time, people act in a predictable fashion with a high level of reliability (R^2 of 0.6–0.8) of average behavior ("measured personality") across situations. The incentives in any situation also matter. Heritability studies show that measures of personality traits tend to be about 40%–60% heritable, suggesting that something tied to the person, rather than the situation, influences behavior (Bouchard and Loehlin, 2001).⁴⁰ Evidence in neuroscience suggests that expression of traits is related to regions of the brain (see Canli, 2006; DeYoung et al., 2010).

2.5. The evolution of personality traits over the life cycle

Even though personality traits are relatively stable across situations, they are not set in stone. They change over the life cycle. Fig. 3 shows that Conscientiousness tends to increase monotonically over the life cycle. Other traits change in different ways over the life cycle.⁴¹ Crystallized intelligence tends to increase monotonically for most of the life cycle, whereas fluid intelligence tends to peak in early adulthood and then decline.⁴²

This evidence does not address whether these changes occur naturally ("ontogenic change") or whether they are due to changes in the environments commonly experienced over the life cycle ("sociogenic change"). No evidence is available in the published literature on the distributions of these profiles over the life cycle. Almlund et al. (2011) review the evidence on how parental investment and interventions promote changes in personality.

³⁸ This theme has been picked up in behavioral economics. See Thaler (2008).

³⁹ See the special issue of *Journal of Research in Personality* (43), entitled "Personality and Assessment at Age 40" for a recent discussion.

⁴⁰ Devlin et al. (1997) suggest that traditional estimates of the heritability of IQ may be inflated because they fail to take into account the effect of the environment on conditions in the maternal womb. See also Rutter (2006) and an emerging literature on epigenetics.

⁴¹ See the evidence collected in Borghans et al. (2008a) and Almlund et al. (2011) for a variety of other traits.

⁴² See McArdle et al. (2000).

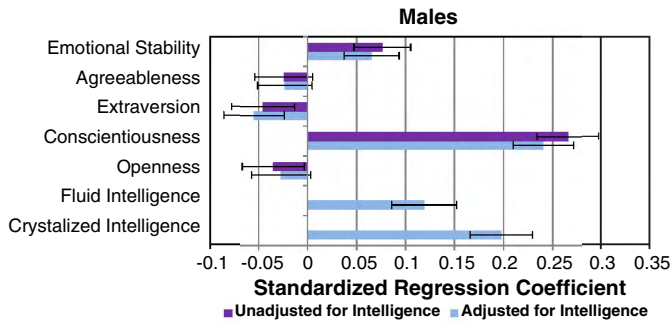


Fig. 4. Association of the Big Five and intelligence with years of completed schooling. Notes: The figure displays standardized regression coefficients from a multivariate regression of years of school attended on the Big Five and intelligence, controlling for age and age squared. The bars represent standard errors. The Big Five coefficients are corrected for attenuation bias. The Big Five were measured in 2005. Years of schooling were measured in 2008. Intelligence was measured in 2006. The measures of intelligence were based on components of the Wechsler Adult Intelligence Scale (WAIS). The data are a representative sample of German adults between the ages 21 and 94. Source: Almlund et al. (2011), German Socio-Economic Panel (GSOEP), waves 2004–2008.

2.6. The predictive power of personality

Table 2 shows that personality traits predict many later-life outcomes as strongly as measures of cognitive ability. **Conscientiousness—the tendency to be perseverant and hardworking—stands out as the most predictive of the Big Five traits across many outcomes.** Fig. 4 presents for males correlations between the Big Five and educational attainment, adjusting and not adjusting for fluid and crystallized intelligence. Conscientiousness predicts educational attainment more than either of the facets of intelligence.⁴³ Similar patterns appear for many other outcomes, including labor market performance, grades, and health.⁴⁴

A recurrent finding in the literature is that measured IQ is highly predictive of performance on complex tasks and jobs (Gottfredson, 1997). The importance of IQ increases with job complexity, defined as the information processing requirements of the job: cognitive skills are more important for professors, scientists, and senior managers than for semi-skilled or unskilled laborers (Schmidt and Hunter, 2004). In contrast, the importance of Conscientiousness does not vary much with job complexity (Barrick and Mount, 1991), suggesting that it pertains to a wider spectrum of jobs.

The literature in economics establishes that the same bundle of traits has different productivity in different tasks. People also differ in their endowments of traits. These two features lead to sorting in the tasks people pursue in life and are a manifestation of the general principle of comparative advantage in the labor market and in life (see Almlund et al., 2011; Borghans et al., 2008a; Cattani, 2012; Heckman et al., 2006a, 2011b).

Achievement test scores are crude, low-dimensional summaries of high-dimensional vectors of traits that operate in conjunction with effort. It is unlikely that these summaries capture the precise combinations of traits required for success in specific life tasks. The thrust of recent research in personality and economics is to isolate the traits that determine life outcomes and to understand how those diverse traits determine choices of tasks.

Most of the evidence in personality psychology is correlational. The reported correlations do not prove that personality traits *cause* higher educational attainment although it is consistent with it. For example, the reported pattern in Fig. 4 could arise if educational attainment increased Conscientiousness. We next present causal evidence.

Table 4
Validities of GED test.

Test	Correlation	Source(s)
Armed Forces Qualification Test (AFQT)	0.75–0.79 ^a	Means and Laurence (1984)
Iowa Test of Educational Development	0.88 ^a	Means and Laurence (1984)
ACT	0.80 ^a	Means and Laurence (1984)
Adult Performance Level (APL) Survey	0.81 ^a	Means and Laurence (1984)
New York's Degrees of Reading Power (DRP) Test	0.77 ^a	Means and Laurence (1984)
Test of Adult Basic Education (TABE)	0.66–0.68 ^a	Means and Laurence (1984)
General Aptitude Test Battery (GATB)	0.61–0.67 ^a	Means and Laurence (1984)
National Adult Literacy Survey (NALS) factor	0.78 ^b	Baldwin (1995)

^a Uses mean GED subtest scores.

^b Uses a general GED factor.

3. Causal evidence

3.1. Problems with establishing causality

Most studies in personality psychology do not address the question of causality, i.e., whether measured traits *cause* (rather than just predict) outcomes. Empirical associations are not a reliable basis for policy analysis. In this section, we discuss difficulties in establishing causality. We also summarize several studies that provide evidence that personality traits *cause* outcomes.

We introduce a simple framework to analyze the effect of traits on outcomes and how traits evolve over time.⁴⁵ Eq. (1) shows how an outcome at age a , T_a , which is the performance on a task, depends on cognition C_a , personality P_a , other acquired skills such as education and job training K_a , and the effort allocated to the task e_{T_a} :

$$T_a = \phi_a \left(\underbrace{C_a}_{\text{Cognition}}, \underbrace{P_a}_{\text{Personality}}, \underbrace{K_a}_{\text{Other acquired skills}}, \underbrace{e_{T_a}}_{\text{Effort devoted to task}} \right), a = 1, \dots, A. \quad (1)$$

Eq. (2) shows how the effort allocated to task T_a depends on cognition C_a , personality P_a , other acquired skills K_a , incentives R_{T_a} , and preferences γ_a :

$$e_{T_a} = \psi_{T_a} \left(\underbrace{C_a, P_a, K_a}_{\text{Incentives to perform on task}}, \underbrace{R_{T_a}}_{\text{Incentives to perform on task}}, \underbrace{\gamma_a}_{\text{Preferences}} \right). \quad (2)$$

The effort applied to a task is the outcome of a choice problem that depends on traits, preferences, and incentives, much like a supply equation in the standard theory of consumer choice. Preferences can be thought of as additional traits.⁴⁶ Some psychological theories posit that people have limited effort that they can divide among different tasks (see, e.g., Baumeister and Tierney, 2011).

Eqs. (1) and (2) formalize the difficulty in establishing a causal relationship between outcomes and traits. Multiple traits, effort, and acquired skills generate performance in a given task. Many studies in psychology and economics do not control for these inputs and equate measurement of a set of outcomes with the trait the analyst is trying to measure.⁴⁷ This practice can lead to a substantial bias in inference about any particular trait.

⁴⁵ This framework draws on Almlund et al. (2011).

⁴⁶ The empirical relationship between measured preference parameters and Big Five measures is weak (see Almlund et al., 2011).

⁴⁷ Selecting measures and verifying them is part of the sometimes mysterious and inherently subjective process of “construct validity” in psychology. For a discussion, see Borghans et al. (2008a).

⁴³ Results are similar for women (see Almlund et al., 2011).

⁴⁴ See Almlund et al. (2011), Borghans et al. (2008a), and Roberts et al. (2007) for comprehensive reviews of the evidence.

An additional point is that most studies assume a linear relationship between outcomes and traits. This practice is particularly problematic for measuring personality traits, where the effect of a trait on an outcome is not always linear or even monotonic. Too much of a good thing can be bad ($\frac{\partial \phi_a}{\partial P_a} < 0$ for $P_a > \bar{P}$ for threshold \bar{P}). For example, extreme levels of traits are associated with psychopathologies. High levels of Conscientiousness are associated with Obsessive Compulsive Disorder, which hinders task performance (Samuel and Widiger, 2008). Nonlinearities can also arise when traits and incentives interact, as in the analyses of Borghans et al. (2008b) and Segal (2008) who show that people with different personality traits respond differently to incentives on tests.⁴⁸

The traits and other acquired skills evolve over time through investment and habituation. Eq. (3) shows that traits at age $a+1$ are age-dependent functions of cognitive ability, personality traits, other acquired skills, and investment I_a at age a . In this way, previous levels of traits and acquired skill affect current levels of traits and acquired skill. Eq. (3) formalizes the notion that the traits and skills governing performance at a point in time are themselves the outcome of investment and habituation:

$$(C_{a+1}, P_{a+1}, K_{a+1}) = \eta_a \left(C_a, P_a, K_a, \underbrace{I_a}_{\text{Investment and experience}} \right), \quad a = 1, \dots, A. \quad (3)$$

In conjunction with resource constraints, a “deeper” set of preference parameters at age a may govern investment decisions and effort allocated to tasks.

3.2. Extreme Examples of Personality Change

Laboratory experiments and brain lesion studies provide some of the most compelling evidence that personality traits can change and that the change affects behaviors. The most famous example is that of Phineas Gage, a railway construction foreman whose head was impaled by a metal spike. Miraculously he retained his problem solving abilities, but he changed from being polite and dependable to being rude and unreliable. His personality change caused him to lose his job and alienate family members (Damasio et al., 2005). Laboratory experiments show that expressed traits can be manipulated temporarily. Magnetic disruption of the left lateral prefrontal cortex can increase experimentally elicited discount rates (Figner et al., 2010) and nasal sprays of oxytocin increase trust (Kosfeld et al., 2005).

3.3. Evidence from the GED testing program

The GED is a standardized achievement test that serves as an alternative to a high school diploma. High school dropouts can take the seven-and-a-half-hour GED exam to certify that they have the “general knowledge” of a high school graduate. The test is widely used. The GED testing program currently produces 12% of high school certificates each year in the United States. We draw on the analysis of Heckman et al. (2012a) and first present results for males. The GED program provides insight into the effects of personality traits on outcomes. GED recipients have the same cognitive ability as high school graduates, but differ in their personality traits.

Table 4 shows the correlations between GED test scores and other achievement test scores. GED test scores are strongly correlated with scores on other standardized achievement tests. The correlations range from 0.61 with the General Aptitude Test Battery (GATB) to 0.88 with the Iowa Test of Educational Development, the progenitor of the GED.

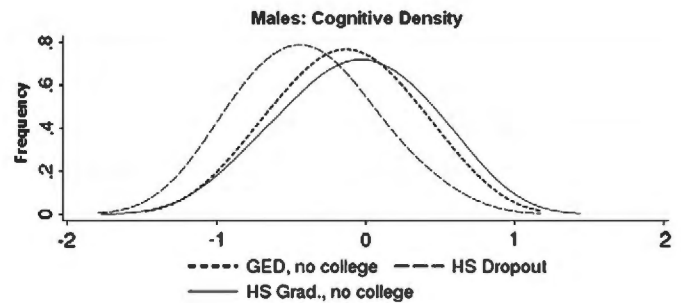


Fig. 5. Cognitive ability by educational status. Source: Reproduced from Heckman et al. (2011b), which uses data from the National Longitudinal Study of Youth 1979 (NLSY79). Notes: The distributions above represent cognitive ability factors estimated using a subset of the Armed Services Vocational Aptitude Battery (ASVAB) and educational attainment as laid out in Hansen et al. (2004). The sample is restricted to the cross-sectional subsample for both males and females. Distributions show only those with no post-secondary educational attainment. The cognitive ability factors are normalized by gender to be mean zero standard deviation one.

GED recipients are smarter than other dropouts. Fig. 5 shows the distributions of a factor extracted from the components of the ASVAB for male high school dropouts, GED recipients, and high school graduates.⁴⁹ The sample excludes people who attend post-secondary education. The distribution of the scores of GED recipients is much more like that of high school graduates than that of high school dropouts.

If they have the same cognitive ability as high school graduates, then why do they drop out of high school? Success in school requires other traits. On a variety of other dimensions, GED recipients behave much more like other dropouts. Fig. 6 shows measures of early adolescent drug use, crime, sex, and violence extracted from three data sources.⁵⁰ Male high school graduates perform better on all measures than high school dropouts or GED recipients. GED recipients are much more similar to dropouts, but in several cases are statistically significantly more likely to engage in risky behaviors than other dropouts. On no outcome measure in that figure are dropouts statistically significantly more likely to engage in risky behaviors compared to GED recipients. Fig. 7 summarizes these adolescent behaviors using a single factor and shows that unlike the cognitive summary measures, the distribution of the non-cognitive (personality) summary measure of GED recipients is much closer to that of dropouts than to that of high school graduates.

The traits that cause GED recipients to drop out of high school manifest themselves in many other life outcomes. One potential benefit of the GED certificate is that it opens doors to post-secondary education. Fig. 8 shows post-secondary educational attainment for GED recipients and high school graduates. About 40% of GED recipients enroll in a 2- or 4-year college. Nearly half drop out within the first year. Fewer than 5% earn a B.A. degree and fewer than 10% earn an A.A. degree (associate degree typically completed in two years).

GED recipients lack persistence in a variety of tasks in life. Fig. 9 shows the survival rates in employment, marriage, and in the condition of not having been incarcerated. GED recipients tend to exit employment, become divorced, and enter jail at rates similar to those of high school dropouts, while high school graduates are much more persistent.

Adjusting for their differences in cognitive ability, male GED recipients perform virtually the same as high school dropouts in the labor market. Fig. 10 shows the hourly wages and annual earnings of male GED recipients and high school graduates compared to high school dropouts for different age groups. The first set of bars shows the

⁴⁹ Similar results are found for females.

⁵⁰ The data sets are the National Longitudinal Survey of Youth 1979 (NLSY79), National Longitudinal Survey of Youth 1997 (NLSY97), and National Educational Longitudinal Survey (NELS). For discussion of these data sets, see Heckman et al. (2012a).

⁴⁸ Formally, this occurs when $\frac{\partial^2 \psi_{T_a}}{\partial P_a \partial K_{T_a}} \neq 0$.

outcomes after adjusting for age, race, year, and region of residence. The second set of bars shows the effects after additionally adjusting for AFQT scores. The third set of bars shows the effects after additionally adjusting for standard measures of family background. GED recipients and high school graduates outperform dropouts in regressions that only adjust for age, race, year, and region of residence. After adjusting for cognitive ability, GED recipients are indistinguishable from dropouts, whereas high school graduates earn more and have higher hourly

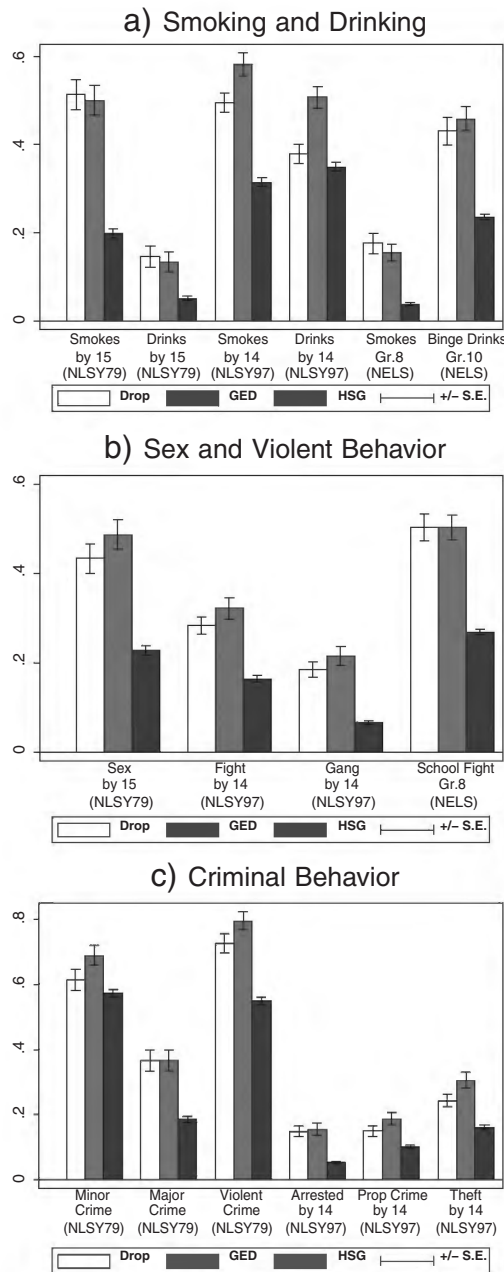


Fig. 6. Measures of adolescent behaviors for male dropouts, GED recipients, and high school graduates. Sources: Heckman et al. (2012a, Chapter 3). National Longitudinal Survey of Youth 1979 (NLSY79), National Longitudinal Survey of Youth 1997 (NLSY97), National Educational Longitudinal Survey (NELS). Notes: Minor crime includes vandalism, shoplifting, petty theft, fraud, holding or selling stolen goods. Major crime includes auto theft, breaking/entering private property, grand theft. Violent crime includes fighting, assault, aggravated assault. Tests of Significance: The estimates for GED recipients and high school graduates are statistically significantly different at the 5% level for all variables. The estimates for dropouts and high school graduates are statistically significantly different at the 5% level for all variables, except for "Minor Crime (NLSY79)" and "Drinks by 14 (NLSY97)." The estimates of "Smokes by 14 (NLSY97)," "Drinks by 14 (NLSY97)," and "Theft by 14 (NLSY97)" between GED recipients and dropouts are statistically significantly different at the 5% level.

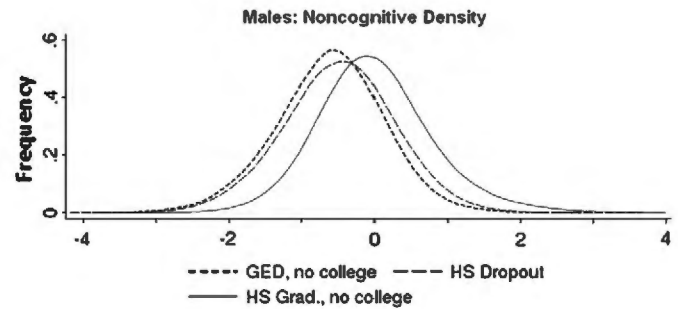


Fig. 7. Distribution of a summary measure of noncognitive ability by education group. Source: Reproduced from Heckman et al. (2011b), which uses data from the National Longitudinal Study of Youth 1979 (NLSY79). Notes: The distributions above represent noncognitive ability factors estimated using measures of early violent crime, minor crime, marijuana use, regular smoking, drinking, early sexual intercourse, and educational attainment as in Hansen et al. (2004). Sample restricted to the cross-sectional subsample for both males and females. Distributions show only those with no post-secondary educational attainment. The noncognitive ability factors normalized to be mean zero standard deviation one.

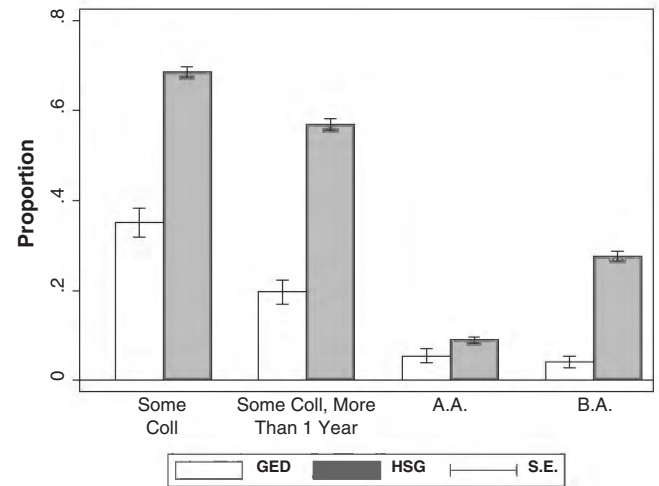


Fig. 8. Post-secondary educational attainment across education groups through age 40 males. Sources: Heckman et al. (2012a, Chapter 4). National Longitudinal Survey of Youth 1979 (NLSY79). Notes: The graph shows post-secondary educational attainment of GED recipients and high school graduates. Variable Definitions: "Some College" represents people who entered any post-secondary institution ever. "Some College, More Than a Year" represents people who completed at least a year of some post-secondary education ever. "A.A." represents people who obtained associate degrees ever. "B.A." represents people who obtained bachelor's degrees ever. "B.A." also includes people with higher education: M.A. Ph.D and professional degrees. Tests of Significance: The estimates for GED recipients and high school graduates are statistically significantly different at the 5% level for all but attainment of the A.A. degree.

wages. Controlling for family background characteristics does not change the story.

Most of the patterns found for women parallel those found for men. However, there are some important differences.⁵¹ While female GED recipients share similar cognitive and personality traits as male GED recipients, their outcomes differ. After accounting for differences in cognitive ability, female GED recipients do not earn higher hourly wages than other dropouts, but unlike men they have higher annual earnings because they are more likely to participate in the labor force.⁵²

⁵¹ See Heckman et al. (2012a).

⁵² The increased labor supply response is largely due to female GED recipients who attain some post-secondary education or who have dropped out of high school due to pregnancy. See Heckman et al. (2012a) for a full discussion of the evidence on the performance of GED recipients.

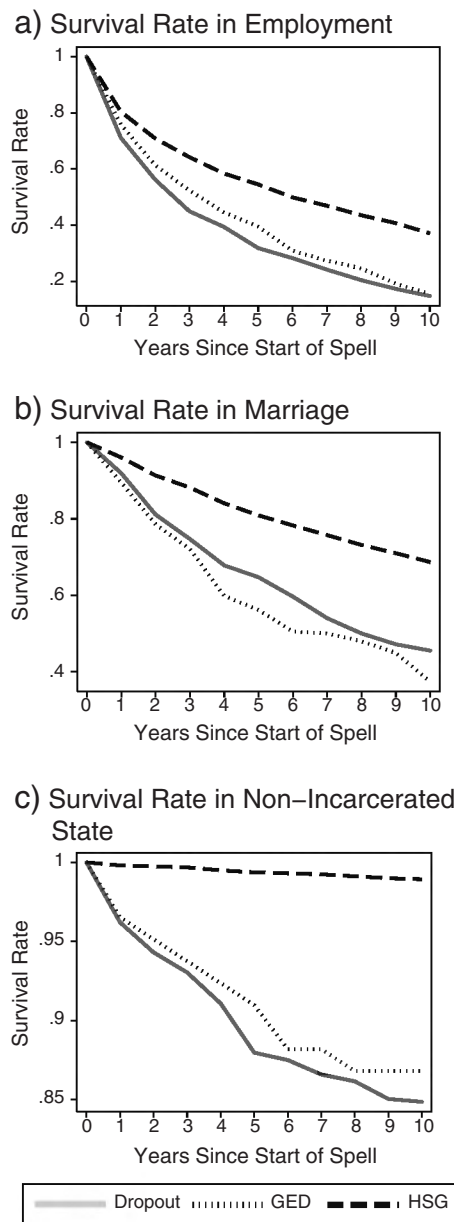


Fig. 9. Survival rates in various states for male dropouts, GED recipients, and high school graduates. Source: Heckman et al. (2012a, Chapter 4). National Longitudinal Survey of Youth 1979 (NLSY79), nationally representative cross sectional sample. Notes: The "Survival Rate in Marriage" is based on the first marriage spell that began after 1979. People who were already married or divorced in 1979 were excluded from the sample because the length of their spell is unknown. The "Survival Rate in Employment" is based on all employment spells that began after 1979 and after the respondent was 16. People are excluded from the employment estimates if they have been to jail. The "Survival Rate in Non-Incarcerated State" is based on the time until first incarceration starting at age 22 (the youngest age for which the jail status is available for all respondents). Respondents who were already in jail at age 22 were excluded from the sample. Tests of Significance: The estimates for GED recipients and high school graduates are statistically significantly different at the 5% level in all cases except for the first year of the "Survival Rate in Employment" and the first year of the "Survival Rate in Marriage." The estimates for dropouts and high school graduates are statistically significantly different at the 5% level in all cases. The estimates for dropouts and GED recipients are significantly different at the 5% level for years three through five of the "Survival Rate in Employment."

3.4. Evidence from the Perry Preschool Program and other interventions

Evidence from the Perry Preschool Program shows how personality traits can be changed in ways that produce beneficial lifetime outcomes.

The Perry preschool Program enriched the lives of 3- and 4-year-old low-income, Black children with initial IQs below 85 at age 3.⁵³

Participants were taught social skills in a "plan-do-review" sequence where students planned a task, executed it, and then reviewed it with teachers and fellow students. They learned to work with others when problems arose.⁵⁴ In addition, home visits promoted parent–child interactions. The program ended after 2 years of enrollment and both treatments and controls entered the same school. The program was evaluated by the method of random assignment.

The program did not improve IQ scores in a lasting way. Fig. 11 shows that, by age ten, treatment and control groups had the same average IQ scores. Many critics of early childhood programs seize on this finding and related evidence to dismiss the value of early intervention studies.

Nevertheless, the program improved outcomes for both boys and girls, resulting in a statistically significant rate of return of around 6–10% per annum for both boys and girls (see Heckman et al., 2010a,b). These returns are above the post-World War II, pre-2008 meltdown in stock market returns to equity estimated to be 5.8% per annum.⁵⁵

The Perry Preschool Program worked primarily through improving personality traits. Participants had better direct measures of personal behavior (a weighted average of "absences and truancies," "lying and cheating," "stealing," and "swears or uses obscene words" measured by teachers in the elementary school years). Participants of both genders improved their "externalizing behavior," a psychological construct related to Agreeableness and Conscientiousness. For girls, the program improved Openness to Experience (proxied by academic motivation). The program also improved scores on the California Achievement Test (CAT). This evidence is consistent with the evidence presented in the previous section that shows that performance on achievement tests depends on personality traits.

Other studies are broadly consistent with the evidence from the Perry Preschool study. Analyses of data from Project STAR, a program that randomly assigned kindergartners and teachers to classes of different sizes, yields results similar to the Perry Program. Using data from Project STAR, Dee and West (2011) find that assignment to a small class is associated with positive changes in personality. In a follow-up analysis, Chetty et al. (2011) examine the Project STAR program and find that students placed in higher quality kindergarten classes—as measured by their peer's average performance on a Stanford Achievement Test—had significantly higher earnings in early adulthood.

The curriculum of Promoting Alternative Thinking Strategies (PATHS) teaches self-control, emotional awareness, and social problem-solving skills and is aimed at elementary school children (see Bierman et al., 2010). A recent random-assignment, longitudinal study demonstrates that the PATHS curriculum reduces teacher and peer ratings of aggression, improves teacher and peer ratings of prosocial behavior, and improves teacher ratings of academic engagement.⁵⁶ PATHS is an exemplar of school-based social and emotional learning (SEL) programs. A recent meta-analysis shows that the program improved grades by 0.33 standard deviations and achievement test scores by 0.27 standard deviations (Durlak et al., 2011).⁵⁷

⁵³ We draw on the analysis of Heckman et al. (2012b).

⁵⁴ Sylva (1997) describes the Perry program as a Vygotskian program fostering personality traits. Vygotsky developed a psychology of child development in structured social settings that emphasized development of social and personality skills. The Vygotskian approach strongly influences the Tools of the Mind program (see Bodrova and Leong, 2001, 2007).

⁵⁵ See DeLong and Magin (2009).

⁵⁶ See Bierman et al. (2010).

⁵⁷ Note, however, that the largest federal study to date on character education programs, including PATHS, failed to find evidence for improvements in behavior or academic performance (see Social and Character Development Research Consortium, 2010).

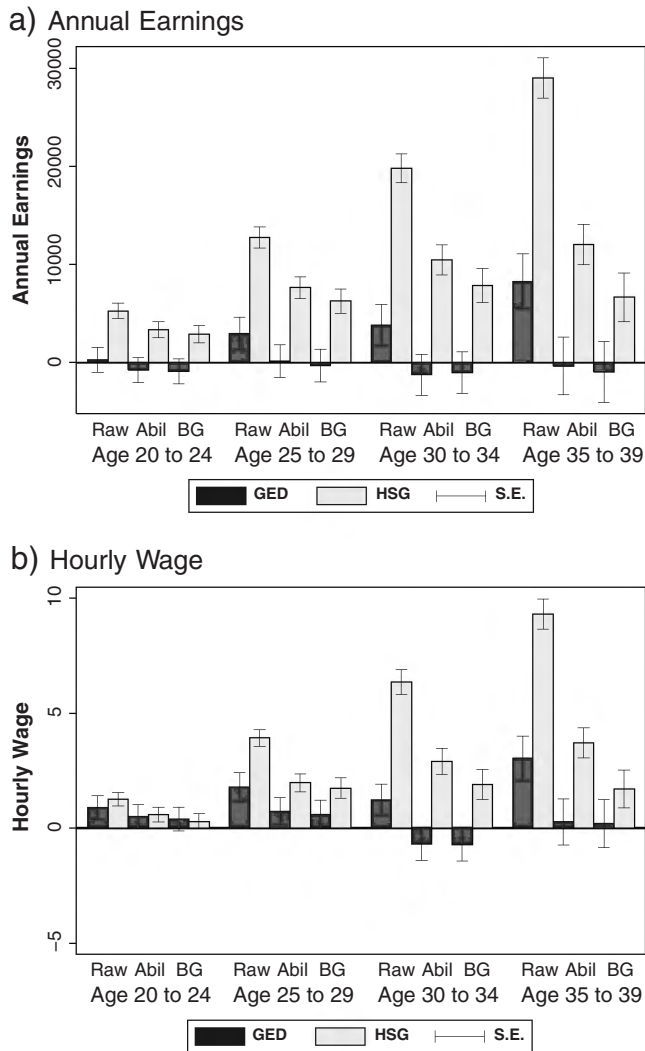


Fig. 10. Labor market outcomes differences. By age—NLSY79—males. Source: Heckman et al. (2012a, Chapter 3). National Longitudinal Survey of Youth 1979 (NLSY79). Controls: “Raw”—age, race, year, and region of residence; “Abil”—age, race, region of residence, and Armed Forces Qualification Test (AFQT) adjusted for schooling at time of test; “BG”—age, race, year, region of residence, mother’s highest grade completed, urban status at age 14, family income in 1978, broken home status at age 14, South at age 14, AFQT, and factors based on adolescent behavioral measures, crime and school performance. Regressions exclude those reporting earning more than \$300,000 or working more than 4000 hours. Notes: All regressions allow for heteroskedastic errors and when appropriate clustering at the individual level.

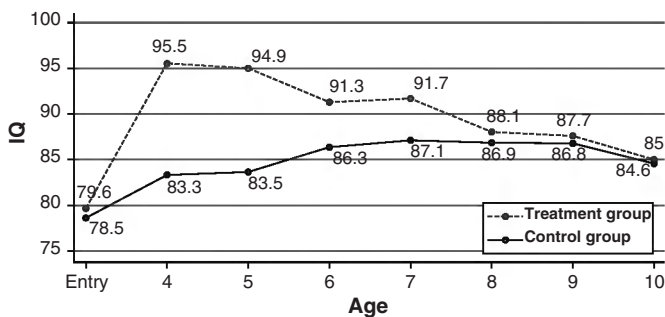


Fig. 11. Perry Preschool Program: IQ, by age and treatment group. Notes: IQ measured on the Stanford-Binet Intelligence Scale (Terman and Merrill, 1960). The test was administered at program entry and at each of the ages indicated. Source: Cunha et al. (2006) and Heckman and Masterov (2007) based on data provided by the High Scope Foundation.

Likewise, several random assignment evaluations of *Tools of the Mind*, a preschool and early primary school curriculum targeting development of self-control, show that it improves classroom behavior as well as executive function, defined as higher-level cognitive skills including inhibitory control, working memory, and cognitive flexibility (Barnett et al., 2006, 2008; Bodrova and Leong, 2001, 2007; Diamond et al., 2007; Lillard and Else-Quest, 2006).⁵⁸ Positive findings are reported for the Montessori preschool curriculum (Lillard and Else-Quest, 2006). Unlike the Perry study, these studies do not have long-term followups.

There is evidence that targeted intervention efforts can improve aspects of Conscientiousness. In contrast to the multi-faceted curricula described above, studies targeting improvement in this trait are designed to isolate a particular mechanism producing behavioral change. For instance, Rueda et al. (2005) designed a set of computer exercises to train attention in children between 4 and 6 years of age. Children in the intervention group improved in performance on computer tasks of attention relative to children who instead watched interactive videos for a comparable amount of time. Similarly, Stevens et al. (2008) designed a 6-week computerized intervention and showed that it can improve selective auditory attention (i.e., the ability to attend to a target auditory signal in the face of an irrelevant, distracting auditory signal). As is typical of much of the literature, all of these programs have only short-term follow-ups.

Several studies suggest that personality can be remediated in adolescence. Martins (2010) analyzes data from EPSIS, a program developed to improve student achievement of 13–15 year-olds in Portugal by increasing motivation, self-esteem, and study skills. The program consists of one-on-one meetings with a trained staff member or meetings in small groups. The intervention was tailored to each participant’s individual skill deficit. Overall, the program was successful and cost-effective, decreasing grade retention by 10 percentage points.

Other life experiences, like employment, can improve personality. Gottschalk (2005) analyzes evidence from a randomized control trial that working at a job can improve locus of control, a trait related to Neuroticism that measures the extent to which individuals believe that they have control over their lives through self-motivation or self-determination as opposed to the extent that the environment controls their lives (Rotter, 1966).⁵⁹ He uses data from the Self-Sufficiency Project (SSP) in which some welfare recipients were randomly offered substantial subsidies to work. The subsidy more than doubled the earnings of a minimum wage worker. People in the experimental group worked about 30% more hours than those in the control group. After 36 months, those who received the subsidy were more likely to have an improved locus of control.

3.5. Additional evidence

Studies that account for the endogeneity of investment and education provide further evidence of the causal effect of education and cognitive and personality traits on outcomes. Heckman et al. (2011b) estimate a sequential model of education to study the effects of education on a variety of outcomes. Correcting for selection into education, they find that early cognitive and personality traits affect schooling choices, labor market outcomes, adult health, and social outcomes and that increasing education promotes beneficial labor market, health, and social outcomes.

⁵⁸ However, a more recent large-scale study (Farran et al., 2011) does not find any effect of the program on self-regulation or literacy, language, and mathematics achievement.

⁵⁹ The relationship between locus of control and the Big Five trait of Neuroticism is discussed in Almlund et al. (2011).

Heckman et al. (2006b) estimate a version of Eq. (3) to analyze the effects of increases in education on measured cognition and personality measures.⁶⁰ Controlling for the problem of reverse causality that schooling may be caused by traits, they find that schooling improves both personality and cognitive traits and that these traits, in turn, boost outcomes.⁶¹

Cunha et al. (2010) estimate a model of the technology of skill formation using longitudinal data on the development of children with rich measures of parental investment and child traits. They control for the endogeneity of investment using shocks to family income along with other instruments. Their model is a version of Eq. (3). Traits are self-productive and exhibit dynamic complementarity—current values of traits affect the evolution of future traits through direct and cross effects. A leading example of a cross effect is that more motivated children are more likely to learn. They estimate parameters that summarize how past personality traits affect future cognitive traits.

They find that self-productivity becomes stronger as children become older, for both cognitive and personality traits. The elasticity of substitution for cognitive inputs is *smaller* later in life. This means that it is more difficult to compensate for the effects of adverse environments on cognitive endowments at later ages than it is at earlier ages. This finding is consistent with the high rank stability of cognition over ages past 10–12 reported in the literature. It also helps to explain the evidence on the ineffectiveness of cognitive remediation strategies for disadvantaged adolescents documented in Cunha et al. (2006), Knudsen et al. (2006) and Cunha and Heckman (2007).

Personality traits foster the development of cognition but not vice versa. It is equally easy at all stages of the child's life cycle to compensate for early disadvantage in endowments using personality traits. (Elasticities of substitution for these traits are essentially the same at different stages of the life cycle.) The most effective adolescent interventions target personality traits.⁶²

4. Summary

This paper reviews recent evidence on the importance of personality in economic and social life. It shows that success in life depends on many traits, not just those measured by IQ, grades, and standardized achievements tests. Personality traits predict and *cause* outcomes.

All psychological traits are measured by performance on tasks. Psychological traits have different productivities in different tasks. Performance on tasks depends on incentives and multiple traits, giving rise to a fundamental identification problem when measuring any single trait. This identification problem is empirically important even for measures of cognitive traits.

The importance of cognitive ability increases with the complexity of the task. Given their endowments of traits and the incentives they face, people sort into tasks in life in pursuit of their comparative advantage.

Traits are stable across situations, but their manifestation depends on incentives to apply effort in the situations where they are measured and also depends on other traits and skills. However, traits are not set in stone. They change over the life cycle and can be

enhanced by education, parenting, and environment to different degrees at different ages.

Scores on achievement tests capture both cognitive and personality traits. Children who are more academically motivated and more curious learn more and have higher test scores. More motivated children also try harder on achievement tests.

The evidence in this paper should give pause to analysts and policy makers who rely solely on achievement tests to monitor school performance and school systems. Standardized achievement tests do not adequately capture many skills that matter in life. GED recipients perform about as well as high school graduates on achievement tests but perform much worse in many aspects of life because they lack important personality traits. Categorizing GED recipients as high school graduates misrepresents national statistics on educational attainment.⁶³ The Perry Preschool Program improved the lives of its participants without increasing their IQ scores, demonstrating why it is problematic to focus curricula exclusively on improving cognitive test scores.

Monitoring school progress and creating programs to enhance skills requires a broader framework of measurement. Interventions that promote beneficial changes in personality have an important place in a portfolio of public policies to foster human development.

References

- Ackerman, P.L., Heggestad, E.D., 1997. Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin* 121, 219–245.
- ACT, Inc., 2007. The ACT Technical Manual. ACT, Inc., Iowa City, IA.
- Almlund, M., Duckworth, A., Heckman, J.J., Kautz, T., 2011. Personality psychology and economics. In: Hanushek, E.A., Machin, S., Woßmann, L. (Eds.), *Handbook of the Economics of Education*, vol. 4. Elsevier, Amsterdam, pp. 1–181.
- Ayllon, T., Kelly, K., 1972. Effects of reinforcement on standardized test performance. *Journal of Applied Behavior Analysis* 5 (4), 477–484.
- Baldwin, J., 1995. Who Took the GED? GED 1994 Statistical Report. American Council on Education, GED Testing Service, Center for Adult Learning, Washington, D.C.
- Barnett, W.S., Yarosz, D.J., Thomas, J., Hornbeck, A., 2006. Educational effectiveness of a Vygotskian approach to preschool education: A randomized trial. Technical report, National Institute for Early Education Research, Rutgers, The State University of New Jersey.
- Barnett, W.S., Jung, K., Yarosz, D.J., Thomas, J., Hornbeck, A., Stechuk, R., Burns, S., 2008. Educational effects of the Tools of the Mind curriculum: A randomized trial. *Early Childhood Research Quarterly* 23 (3), 299–313.
- Barrick, M.R., Mount, M.K., 1991. The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology* 44 (1), 1–26.
- Baumeister, R.F., Tierney, J.J.M., 2011. *Willpower: Rediscovering the Greatest Human Strength*. Penguin Press, New York.
- Benjamin, D.J., Brown, S.A., Shapiro, J.M., 2006, May. Who is 'behavioral'? Cognitive ability and anomalous preferences. Unpublished manuscript, Cornell University, Ithaca, NY.
- Bierman, K.L., Coie, J.D., Dodge, K.A., Greenberg, M.T., Lochman, J.E., McMahon, R.J., Pinderhughes, E., 2010. The effects of a multiyear universal social-emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology* 78 (2), 156–168.
- Binet, A., Simon, T., 1916. The development of intelligence in children (The Binet–Simon Scale). Psychological Science. Williams & Wilkins Co., Baltimore, MD.
- Bodrova, E., Leong, D.J., 2001. Tools of the Mind: A Case Study of Implementing the Vygotskian Approach in American Early Childhood and Primary Classrooms. International Bureau of Education, UNESCO, Geneva.
- Bodrova, E., Leong, D.J., 2007. Tools of the Mind: The Vygotskian Approach to Early Childhood Education. Pearson Education, Inc., Upper Saddle River.
- Borghans, L., Duckworth, A.L., Heckman, J.J., ter Weel, B., 2008a. The economics and psychology of personality traits. *Journal of Human Resources* 43 (4), 972–1059 (Fall).
- Borghans, L., Meijers, H., ter Weel, B., 2008b. The role of noncognitive skills in explaining cognitive test scores. *Economic Inquiry* 46 (1), 2–12 (January).
- Borghans, L., Golsteyn, B.H.H., Heckman, J.J., Humphries, J.E., 2011a. Identification problems in personality psychology. In: Ferguson, E., Heckman, J.J., Corr, P. (Eds.), *Personality and Individual Differences* 51, 315–320 (Special Issue on Personality and Economics).
- Borghans, L., Golsteyn, B.H.H., Heckman, J.J., Humphries, J.E., 2011b. Reinterpreting estimated effects of cognition on social outcomes. Unpublished manuscript, Department of Economics, University of Chicago.
- Bouchard, T.J., Loehlin, J.C., 2001. Genes, evolution and personality. *Behavior Genetics* 31 (3), 243–273 (May).
- Bound, J., Brown, C., Mathiowetz, N., 2001. Measurement error in survey data. In: Heckman, J.J., Leamer, E. (Eds.), *Handbook of Econometrics*. Volume 5 of *Handbooks in Economics*. Elsevier Science, Amsterdam, pp. 3705–3843.

⁶³ See Heckman and LaFontaine (2010).

⁶⁰ They estimate the effect of schooling on self-esteem and locus of control, personality traits related to Neuroticism. The Rosenberg Self-Esteem Scale attempts to assess the degree of approval or disapproval of oneself (Rosenberg, 1965). The relationship between these measures and the Big Five traits of Neuroticism is discussed in Almlund et al. (2011).

⁶¹ Both Heckman et al. (2011b) and Heckman et al. (2006b) use an identification strategy based on matching on proxies for unobserved traits that corrects for measurement error and the endogeneity of schooling.

⁶² Cunha et al. (2006) report that 16% of the variation in educational attainment is explained by adolescent cognitive traits, 12% is due to adolescent personality (socio-emotional traits), and 15% is due to measured parental investments.

- Bowen, W.G., Chingos, M.M., McPherson, M.S., 2009. Test scores and high school grades as predictors. Crossing the Finish Line: Completing College at America's Public Universities. Princeton University Press, Princeton, NJ, pp. 112–133.
- Bowles, S., Gintis, H., 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. Basic Books, New York.
- Bowles, S., Gintis, H., Osborne, M., 2001. The determinants of earnings: A behavioral approach. *Journal of Economic Literature* 39 (4), 1137–1176 (December).
- Breuning, S.E., Zella, W.F., 1978. Effects of individualized incentives on norm-referenced IQ test performance of high school students in special education classes. *Journal of School Psychology* 16 (3), 220.
- Canli, T., 2006. *Biology of Personality and Individual Differences*. Guilford Press, New York.
- Carroll, J.B., 1993. *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge University Press, New York.
- Cattan, S., 2012. Heterogeneity and selection in the labor market. PhD Thesis, Economics Department, University of Chicago.
- Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Diane, S.W., Yagan, D., 2011. How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Quarterly Journal of Economics* 126 (4), 1593–1660.
- Clingman, J.P., Fowler, R.L., 1976. The effects of primary reward on the IQ performance of grade-school children as a function of initial IQ level. *Journal of Applied Behavior Analysis* 9 (1), 19–23.
- Cloninger, C.R., Svrakic, D.M., Bayon, C., Przybeck, T.R., 1999. Measurement of psychopathology as variants of personality. In: Cloninger, C.R. (Ed.), *Personality and Psychopathology*. American Psychiatric Publishing, Inc., Arlington, VA.
- Costa, P.T., McCrae, R.R., 1992a. Four ways five factors are basic. *Personality and Individual Differences* 13 (6), 653–665.
- Costa, P.T., McCrae, R.R., 1992b. Revised NEO Personality Inventory (NEO PI-R) and the NEO Five-Factor Inventory (NEO-FFI) Professional Manual. Psychological Assessment Resources, Odessa, FL.
- Cunha, F., Heckman, J.J., 2007. The technology of skill formation. *American Economic Review* 97 (2), 31–47 (May).
- Cunha, F., Heckman, J.J., Lochner, L.J., Masterov, D.V., 2006. Interpreting the evidence on life cycle skill formation. In: Hanushek, E.A., Welch, F. (Eds.), *Handbook of the Economics of Education*. North-Holland, Amsterdam, pp. 697–812 (Chapter 12).
- Cunha, F., Heckman, J.J., Schennach, S.M., 2010. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78 (3), 883–931 (May).
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A.M., Damasio, A.R., 2005. The return of Phineas Gage: Clues about the brain from the skull of a famous patient. In: Cacioppo, J.T., Berntson, G.G. (Eds.), *Social Neuroscience: Key Readings*. Psychology Press, New York, NY, pp. 21–28.
- Dee, T.S., West, M.R., 2011. The non-cognitive returns to class size. *Educational Evaluation and Policy Analysis* 33 (1), 23–46.
- DeLong, J., Magin, K., 2009. The U.S. equity return premium: Past, present and future. *Journal of Economic Perspectives* 23 (1), 193–208 (Winter).
- Devlin, B., Daniels, M., Roeder, K., 1997. The heritability of IQ. *Nature* 388 (6641), 468–471.
- DeYoung, C.G., Hirsh, J.B., Shane, M.S., Papademetris, X., Rajeevan, N., Gray, J.R., 2010. Testing predictions from personality neuroscience: Brain structure and the Big Five. *Psychological Science* 21 (6), 820–828.
- Diamond, A., Barnett, S., Thomas, J., Munro, S., 2007. Preschool program improves cognitive control. *Science* 318 (5855), 1387–1388.
- Duckworth, A.L., Peterson, C., Matthews, M.D., Kelly, D.R., 2007. Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology* 92 (6), 1087–1101 (June).
- Durlak, J.A., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D., Schellinger, K.B., 2011. The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development* 82 (1), 405–432.
- Eidlund, C.V., 1972. 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 (3), 317–319.
- Epstein, S., 1979. The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology* 37 (7), 1097–1126.
- Farran, Dale C., Lipsey, M.W., Wilson, S., 2011. Experimental evaluation of the Tools of the Mind Pre-K curriculum. Technical report. Peabody Research Institute Report (August).
- Feingold, A., 1982. The validity of the information and vocabulary subtests of the WAIS. *Journal of Clinical Psychology* 38 (1), 169–174.
- Figuer, B., Knoch, D., Johnson, E.J., Krosch, A.R., Lisanby, S.H., Fehr, E., Weber, E.U., 2010. Lateral prefrontal cortex and self-control in intertemporal choice. *Nature Neuroscience* 13 (5), 538–539.
- Flynn, J.R., 2007. *What is Intelligence? Beyond the Flynn Effect*. Cambridge University Press, New York.
- Friedman, D., Streicher, A.H., 1985. Reliability of scores for fiscal year 1981 army applicants: Armed Services Vocational Aptitude Battery forms 8, 9, and 10. Technical report. Defense Technical Information Center.
- GED Testing Service, 2009. The technical manual: 2002 series GED tests. (March) Technical Manual. American Council on Education and GED Testing Service, Washington, DC. (http://www.acenet.edu/Content/NavigationMenu/ged/pubs/TechnicalManual_2002SeriesGEDTests.pdf).
- Gensowski, M., 2012. Personality, IQ, and lifetime earnings. Unpublished manuscript, University of Chicago, Department of Economics.
- Gottfredson, L.S., 1997. Why g matters: The complexity of everyday life. *Intelligence* 24 (1), 79–132 (January–February).
- Gottschalk, P., 2005. Can work alter welfare recipients' beliefs? *Journal of Policy Analysis and Management* 24 (3), 485–498.
- Gough, H.G., Heilbrun, A.B., 1983. *The Adjective Check List Manual*. Consulting Psychologists Press, Palo Alto, CA.
- Greene, A.C., Sapp, G.L., Chissom, B., 1990. Validation of the Stanford–Binet Intelligence Scale: Fourth edition with exceptional black male students. *Psychology in the Schools* 27 (1), 35–41.
- Grek, S., 2009. Governing by numbers: The PISA 'effect' in Europe. *Journal of Education Policy* 24 (1), 23–37.
- Hansen, K.T., Heckman, J.J., Mullen, K.J., 2004. The effect of schooling and ability on achievement test scores. *Journal of Econometrics* 121 (1–2), 39–98 (July–August).
- Hartlage, L.C., Steele, C.T., 1977. WISC and WISC-R correlates of academic achievement. *Psychology in the Schools* 14 (1), 15–18.
- Heckman, J.J., LaFontaine, P.A., 2010. The American high school graduation rate: Trends and levels. *The Review of Economics and Statistics* 92 (2), 244–262 (May).
- Heckman, J.J., Masterov, D.V., 2007. The productivity argument for investing in young children. *Review of Agricultural Economics* 29 (3), 446–493.
- Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A., Shaikh, A., Yavitz, A.Q., 2006a. The Perry Preschool Project: A reanalysis. Unpublished manuscript, University of Chicago, Department of Economics.
- Heckman, J.J., Stixrud, J., Urzúa, S., 2006b. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24 (3), 411–482 (July).
- Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A., Yavitz, A.Q., 2010a. The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics* 94 (1–2), 114–128 (February).
- Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A., Yavitz, A.Q., 2010b. Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics* 1 (1), 1–46.
- Heckman, J.J., Humphries, J.E., Mader, N., 2011a. The GED. In: Hanushek, E.A., Machin, S., Woßmann, L. (Eds.), *Handbook of the Economics of Education*, vol. 3. North Holland, Elsevier, Amsterdam, pp. 423–484 (Chapter 9).
- Heckman, J.J., Humphries, J.E., Urzúa, S., Veramendi, G., 2011b. The effects of educational choices on labor market, health, and social outcomes. Unpublished manuscript, University of Chicago, Department of Economics.
- Heckman, J.J., Humphries, J.E., Kautz, T., 2012a. The GED and the Role of Character in American Life. Unpublished book manuscript, University of Chicago, Department of Economics.
- Heckman, J.J., Malofeeva, L., Pinto, R., Savelyev, P.A., 2012b. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. Unpublished manuscript, University of Chicago, Department of Economics (first draft, 2008). Under revision, *American Economic Review*.
- Herrnstein, R.J., Murray, C.A., 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. Free Press, New York.
- Holt, M.M., Hobbs, T.R., 1979. The effects of token reinforcement, feedback and response cost on standardized test performance. *Behaviour Research and Therapy* 17 (1), 81–83.
- Jencks, C., 1979. *Who Gets Ahead? The Determinants of Economic Success in America*. Basic Books, New York.
- John, O.P., Srivastava, S., 1999. The Big Five trait taxonomy: History, measurement and theoretical perspectives. In: Pervin, L.A., John, O.P. (Eds.), *Handbook of Personality: Theory and Research*. The Guilford Press, New York, pp. 102–138 (Chapter 4).
- Kettner, N., 1976. Armed Services Vocational Aptitude Battery (ASVAB form 5): Comparison with GATB and DAT tests: Final report for period May 1975–October 1976. Technical report, DTIC Document: Department of Defense, Department of the Air Force, Air Force Systems Command, Air Force Human Resources Laboratory.
- Knudsen, E.I., Heckman, J.J., Cameron, J., Shonkoff, J.P., 2006. Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences* (27), 10155–10162 (July).
- Kobrin, J.L., Patterson, B.F., Shaw, E.J., Mattern, K.D., Barbuti, S.M., 2008. Validity of the SAT for predicting first-year college grade point average. *The College Board* 5, 1–10.
- Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E., 2005. Oxytocin increases trust in humans. *Nature* 435 (7042), 673–676.
- Larson, G.E., Saccuzzo, D.P., Brown, J., 1994. Motivation: Cause or confound in information processing/intelligence correlations? *Acta Psychologica* 85 (1), 25–37.
- Lemann, N., 1999. *The Big Test: The Secret History of the American Meritocracy*. Farrar, Straus and Giroux, New York.
- Lillard, A., Else-Quest, N., 2006. The early years: Evaluating Montessori. *Science* 313 (5795), 1893–1894.
- Lindquist, E.F., 1951. Preliminary considerations in objective test construction. In: Lindquist, E.F. (Ed.), *Educational Measurement*. American Council on Education, Washington, DC, pp. 119–184.
- Martins, P.S., 2010. Can targeted, non-cognitive skills programs improve achievement? Discussion Paper 5266. IZA.
- McAdams, D.P., Pals, J.L., 2006. A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist* 61 (3), 204–217.
- McArdle, J.J., Hamagami, F., Meredith, W., Bradway, K.P., 2000. Modeling the dynamic hypotheses of Gf-Gc theory using longitudinal life-span data. *Learning and Individual Differences* 12 (1), 53–79.
- Means, B., Laurence, J.H., 1984. Characteristics and performance of recruits enlisted with General Educational Development (GED) credentials. Technical Report FR-PRD-84-6. Human Resources Research Organization, Alexandria, VA.
- Michalko, K.T., Saklofske, D.H., 1996. A psychometric investigation of the Wechsler Individual Achievement Test with a sample of Saskatchewan schoolchildren. *Canadian Journal of School Psychology* 12 (1), 44–54.
- Mischel, W., 1968. *Personality and Assessment*. Wiley, New York.
- Murray, H.A., 1938. *Explorations in Personality: A Clinical and Experimental Study of Fifty Men of College Age*. Oxford University Press, New York.

- National Center for Education Statistics (Various). Digest of Education Statistics. Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Niolon, R., 2005. Introduction to the WAIS III. Available at http://www.psychpage.com/learning/library/intell/wais_history.html (posted August, 2005).
- Nisbett, R.E., 2009. Intelligence and How to Get It: Why Schools and Cultures Count. W. W. Norton and Company, New York, NY. (February).
- Nisbett, R.E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D.F., Turkheimer, E., 2012. Intelligence: New findings and theoretical developments. *American Psychologist* 67 (2), 130–159.
- O'Leary, U.-M., Rusch, K.M., Guastello, S.J., 1991. Estimating age-stratified WAIS-R IQs from scores on the Raven's Standard Progressive Matrices. *Journal of Clinical Psychology* 47 (2), 277–284.
- Omizo, M.M., 1980. The differential aptitude tests as predictors of success in a high school for engineering program. *Educational and Psychological Measurement* 40 (1), 197–203.
- R. R. Bowker Publishing (Various). The Bowker Annual: Library and Book Trade Almanac.
- Raven, J.C., 1962. Advanced Progressive Matrices: Sets I and II, Revised ed. H.K. Lewis, London.
- Raven, J., Raven, J.C., Court, J., 1988. Manual for Raven's Progressive Matrices and Vocabulary Scales. Harcourt Assessment, San Antonio, TX.
- Roberts, B.W., 2009. Back to the future: Personality and assessment and personality development. *Journal of Research in Personality* 43 (2), 137–145.
- Roberts, B.W., Mroczek, D., 2008. Personality trait change in adulthood. *Current Directions in Psychological Science* 17 (1), 31–35.
- Roberts, R.D., Goff, G.N., Anjou, F., Kyllonen, P.C., Pallier, G., Stankov, L., 2000. The Armed Services Vocational Aptitude Battery (ASVAB): Little more than acculturated learning (Gc)!? *Learning and Individual Differences* 12 (1), 81–103.
- Roberts, B.W., Walton, K.E., Viechtbauer, W., 2006. Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin* 132 (1), 1–25.
- Roberts, B.W., Kuncel, N.R., Shiner, R.L., Caspi, A., Goldberg, L.R., 2007. The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science* 2 (4), 313–345 (December).
- Rosenberg, M., 1965. Society and the Adolescent Self-Image. Princeton University Press, Princeton, NJ.
- Rothlisberg, B., 1987. Comparing the Stanford-Binet, fourth edition to the WISC-R: A concurrent validity study. *Journal of School Psychology* 25 (2), 193–196.
- Rotter, J.B., 1966. Generalized Expectancies for Internal versus External Control of Reinforcement. American Psychological Association, Washington DC.
- Rueda, M.R., Rothbart, M.K., McCandliss, B.D., Saccomanno, L., Posner, M.I., 2005. Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences* 41, 14931–14936.
- Rutter, M., 2006. Implications of resilience concepts for scientific understanding. *Annals of the New York Academy of Sciences* 1094 (1), 1–12.
- Samuel, D.B., Widiger, T.A., 2008. A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: A facet level analysis. *Clinical Psychology Review* 28 (8), 1326–1342.
- Savelyev, P., 2011. Conscientiousness, education, and longevity of high-ability individuals. Unpublished manuscript, University of Chicago, Department of Economics.
- Schmidt, F.L., Hunter, J., 2004. General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology* 86 (1), 162–173.
- Segal, C., 2008. Motivation, test scores, and economic success. Economics Working Papers. Department of Economics and Business, Universitat Pompeu Fabra, Barcelona, Spain. (October).
- Social and Character Development Research Consortium, 2010. Efficacy of schoolwide programs to promote social and character development and reduce problem behavior in elementary school children. Research Report NCER 20112001. National Center for Education Research, Institute of Education Sciences, U.S. Department of Education (October).
- Stevens, C., Fanning, J., Coch, D., Sanders, L., Neville, H., 2008. Neural mechanisms of selective auditory attention are enhanced by computerized training: Electrophysiological evidence from language-impaired and typically developing children. *Brain Research* 1205, 55–69.
- Sylva, K., 1997. The quest for quality in curriculum. In: Schweinhart, L.J., Weikart, D.P. (Eds.), *Lasting Differences: The High/Scope Preschool Curriculum Comparison Study through Age 23*. High/Scope Press, Ypsilanti, pp. 89–93.
- Terman, L.M., Merrill, M.A., 1960. Stanford-Binet Intelligence Scale: Manual for the Third Revision Form L-M. Houghton Mifflin, Boston.
- Terman, L.M., Baldwin, B.T., Bronson, E., DeVoss, J.C., Fuller, F., Lee Kelley, T., Lima, M., Marshall, H., Moore, A.H., Raubenheimer, A.S., Ruch, G.M., Willoughby, R.L., Benson Wyman, J., Hazeltine Yates, D., 1925. *Genetic Studies of Genius: Mental and Physical Traits of a Thousand Gifted Children*, vol. 1. Stanford University Press, Stanford University, CA.
- Terman, L.M., Oden, M.H., Bayley, N., Marshall, H., McNemar, Q., Sullivan, E.B., 1947. *Genetic Studies of Genius: The Gifted Child Grows Up: Twenty-Five Years' Follow-Up of a Superior Group*, vol. 4. Stanford University Press, Stanford University, CA.
- Thaler, R.H., 2008. A short course in behavioral economics. Edge Master Class, Sonoma, CA, July 25–27, 2008 (July).