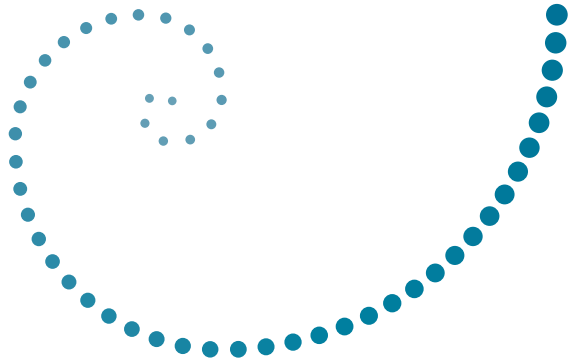
Fostering and Measuring Skills

Heckman, J. and Kautz, T.
Center for the Economics of Human Development,
University of Chicago

Skills for Workforce Success: from Research to Action
The Brookings Institution, Washington, D.C.
Wednesday, June 17th, 2015

THE UNIVERSITY OF CHICAGO



CENTER FOR THE ECONOMICS
OF HUMAN DEVELOPMENT

THE UNIVERSITY OF CHICAGO

Main Points

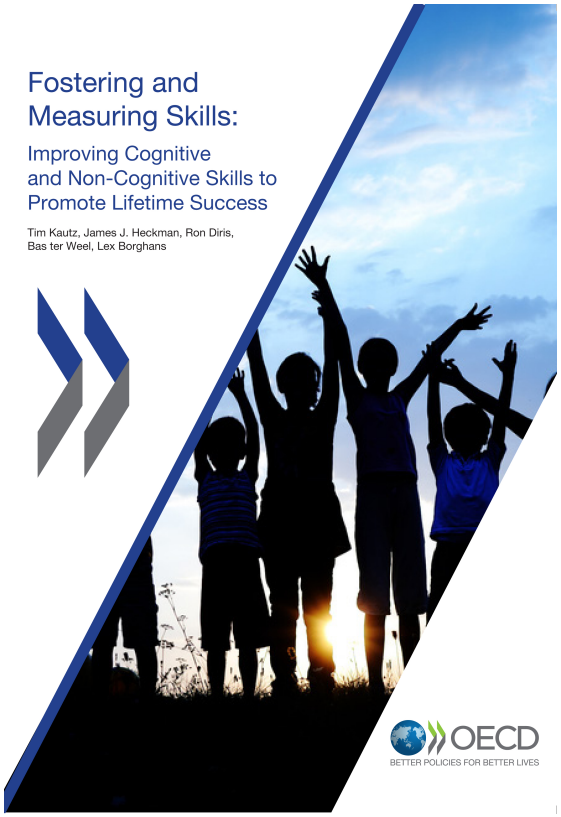
- Many important life skills not captured by scores on achievement tests
- A sole focus on achievement test scores (e.g., NCLB, PISA and Iowa tests) give an incomplete picture of what schools, families, and communities do and how to evaluate schools and other life cycle interventions.
- Socioemotional skills—character, etc.—are important
- These skills can be measured
- They are malleable, and there are effective interventions to promote them
- Soft skills more malleable than cognitive skills at later ages

THE UNIVERSITY OF CHICAGO

- Older ("established") measurement systems such as the Big Five do **not** capture the rich range of behaviors and traits that children and adults exhibit
- Need comprehensive measures of traits
- Instead of relying exclusively or mainly on self-reported "Big Five measures," we should use approaches based on behaviors
  - Teacher reports and assessments as encoded in school system records and interviews
  - Eliciting preference parameters from observed choices in the field and in controlled choice experiments and behaviors:
    - Risk aversion
    - Time preference
    - Ambiguity aversion
    - Trust
    - Reciprocity (positive and negative)
- For all measurement systems, we should adjust for incentives and other traits

THE UNIVERSITY OF CHICAGO

James J. Heckman,Tim Kautz, Ron Diris, Bas ter Weel, and Lex Borghans
Fostering and Measuring Skills: Improving Cognitive and Noncognitive Skills to Promote Lifetime Success
OECD, 2014



THE UNIVERSITY OF CHICAGO

"To value schools, by length instead of quality, is a matchless absurdity. Arithmetic, grammar, and the other rudiments, as they are called, comprise but a small part of the teachings in a school. The rudiments of feeling are taught not less than the rudiments of thinking. The sentiments and passions get more lessons than the intellect. Though their open recitations may be less, their secret rehearsals are more."

—Horace Mann (1838)

THE UNIVERSITY OF CHICAGO

Modern History of Testing

- Starts with IQ tests

THE UNIVERSITY OF CHICAGO

General Knowledge: The Achievement Test

- What schools add to the capability of students to perform tasks

THE UNIVERSITY OF CHICAGO

"We lean heavily on written examinations, on a few types of objective tests, and on the subjective impressions of teachers. Many other appraisal devices could be used, such as records of activities in which pupils participate, questionnaires, check lists, anecdotal records and observational records, interviews, reports made by parents, products made by the pupils, and records made by instruments (motion pictures, eye-movement records, sound recordings, and the like)."

—Ralph Tyler (1940)

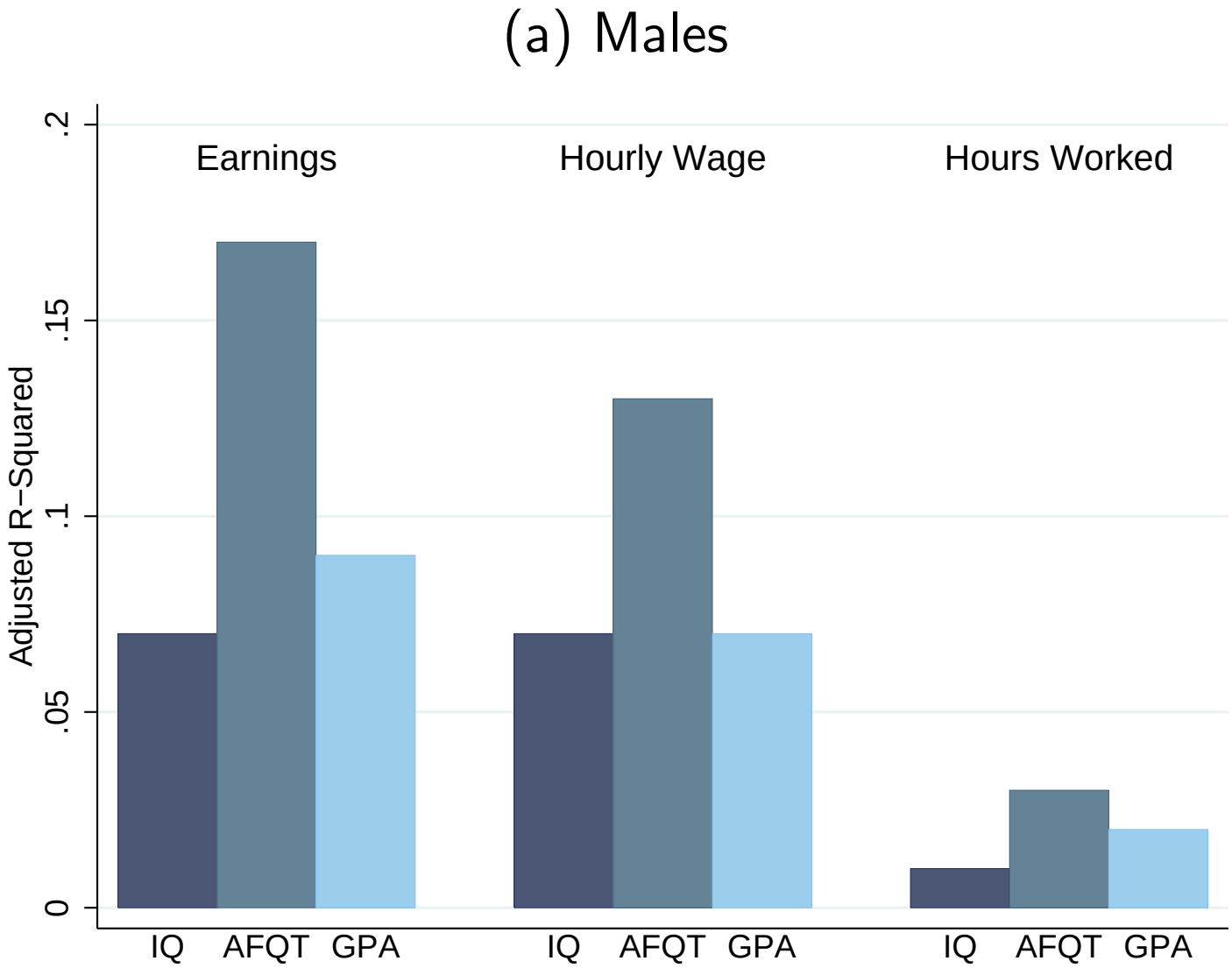
Testing the Tests

THE UNIVERSITY OF CHICAGO

- How well do test scores predict later-life outcomes that matter?
- Prediction is the hallmark of success of any measurement system.

THE UNIVERSITY OF CHICAGO

Figure 1: Validities of Cognitive Measures in Age-35 Labor Market Outcomes (Adjusted R-Squared)



NIVERSITY OF
ICAGO

- Much of the variance in outcomes is not explained. Lots of room for improvement.

THE UNIVERSITY OF CHICAGO

Character skills are a missing ingredient

THE UNIVERSITY OF CHICAGO

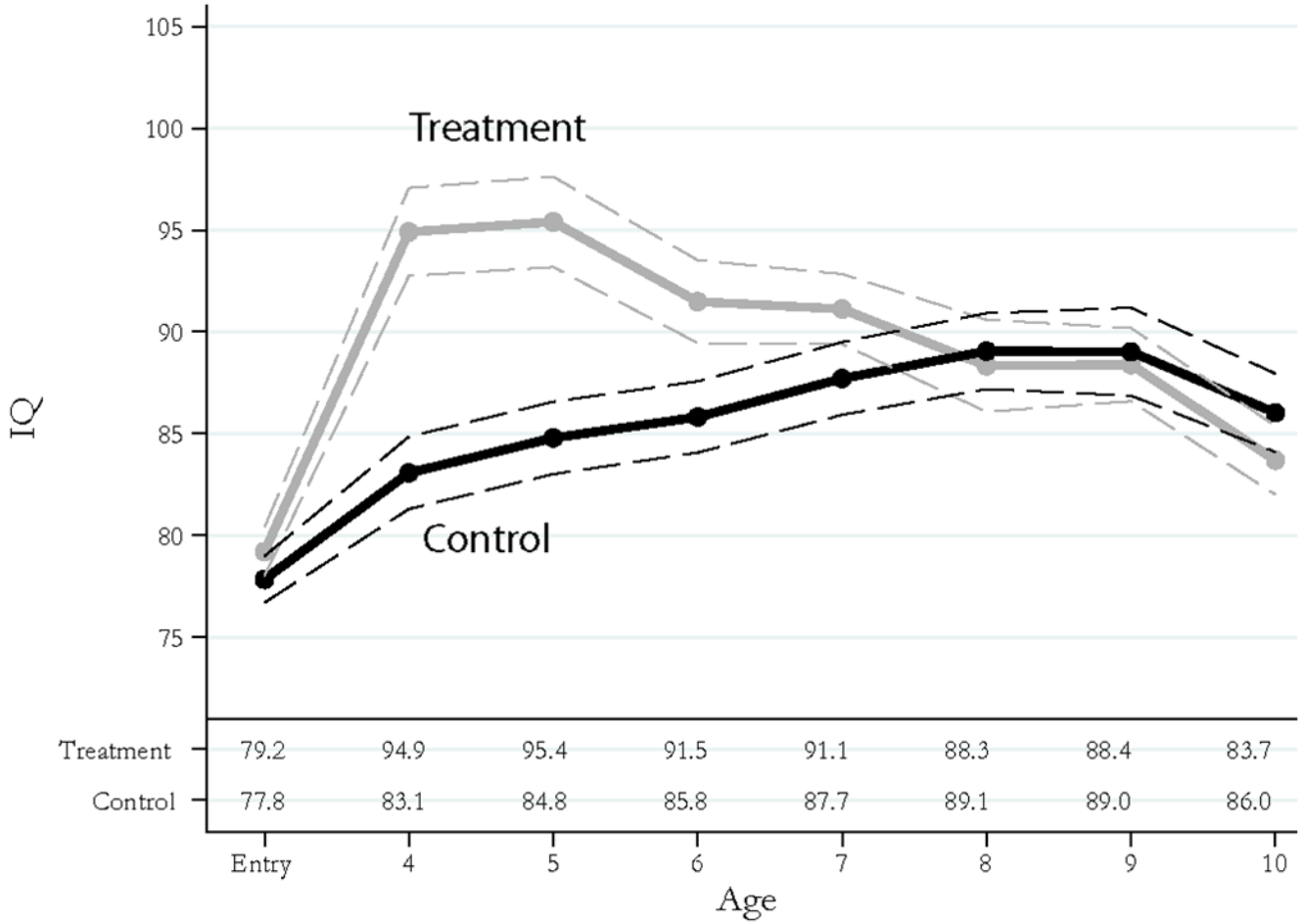
Early childhood programs are the most well-studied and many are effective

- They show the importance of noncognitive skills

THE UNIVERSITY OF CHICAGO

Cognitive Evolution through time, Perry Males

Figure 2: Male Cognitive Dynamics



Yet the Perry Program has a statistically significant annual rate of return of around 6%-10% per annum– for both boys and girls– in the range of the post-World War II stock market returns to equity in the U.S. labor market, estimated to be 6.9%.

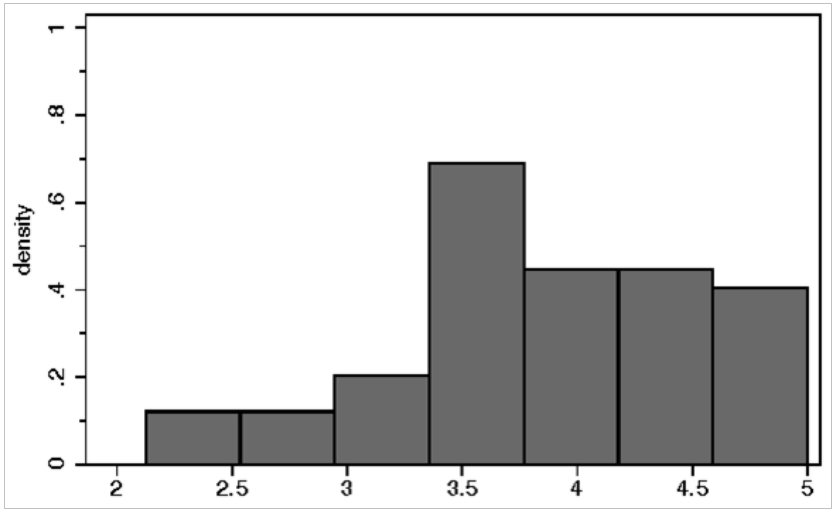
THE UNIVERSITY OF CHICAGO

- It worked primarily through noncognitive and character channels
- Early interventions reducing problem behavior in children reduce unhealthy behavior in adulthood
- They promote adult health

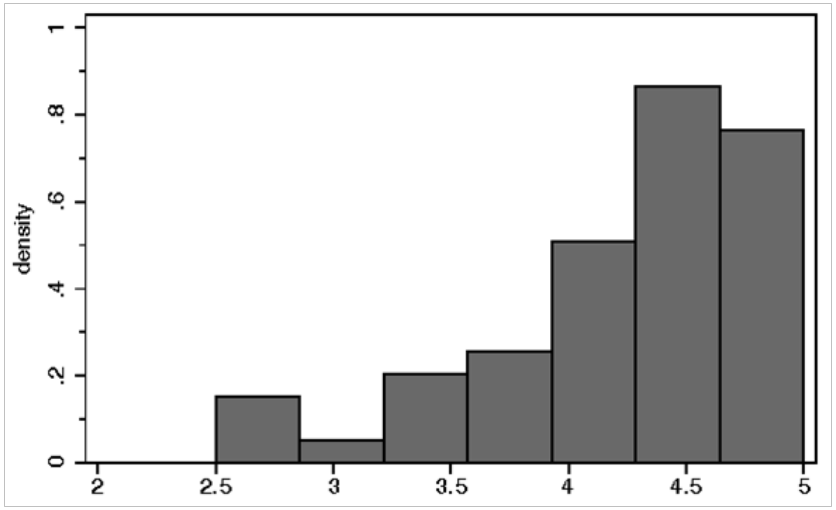
THE UNIVERSITY OF CHICAGO

Figure 3: Mechanisms: Externalizing Behavior, Males

(a) Control,



(b) Treatment,

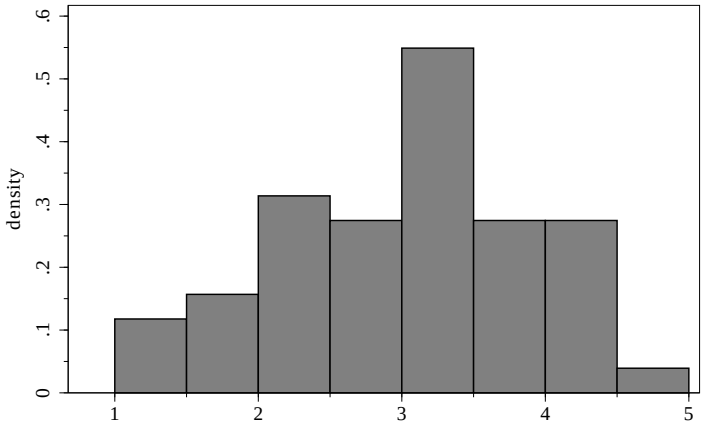


Data: Perry Preschool Program.
Source: Heckman, Pinto, Savelyev (2013).

Figure 4: Histograms of Indices of Noncognitive Skills and CAT Scores

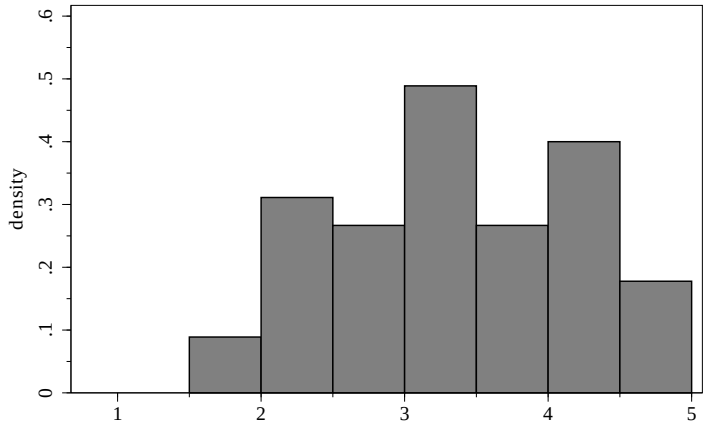
(a) Academic Motivation

Control Group



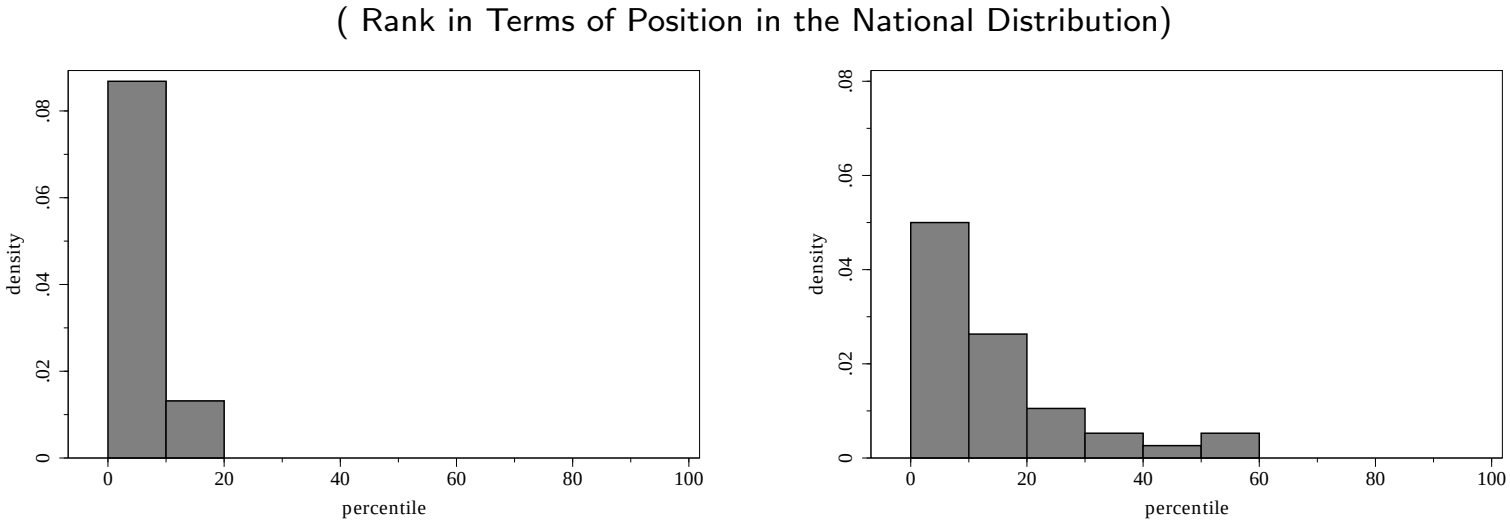
(b) Academic Motivation

Treatment Group



Source: Heckman et al. (2013).

Figure 5: Histograms of Indices of Noncognitive Skills and CAT Scores



Source: Heckman et al. (2013).

- Data on socioemotional skills come from **teacher reports** in early grade school years
- Self reports and parental reports are far less accurate predictors

THE UNIVERSITY OF CHICAGO

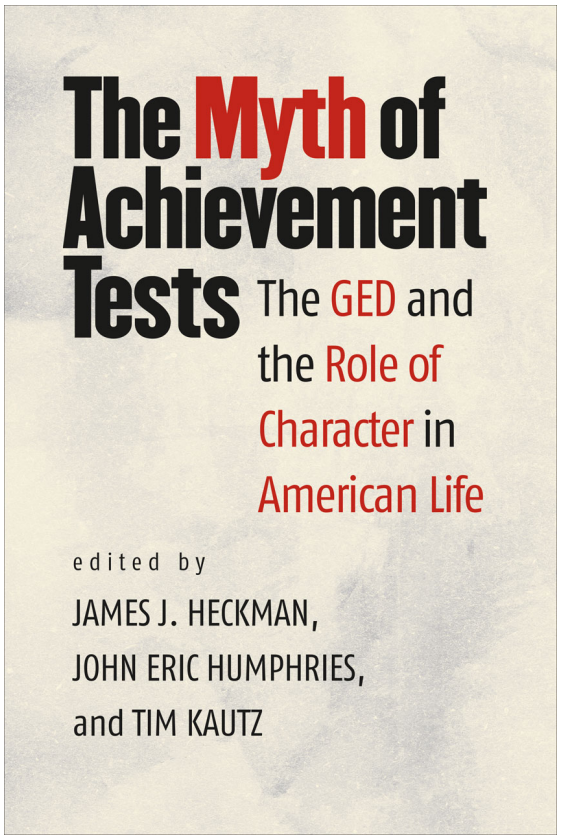
Figure 6: Abecedarian Project, Health Effects at Age 35 (Males)

	Treatment Mean	Control Mean	Treatment p-value
Systolic Blood Pressure	125.79	143.33	0.018
Diastolic Blood Pressure	78.53	92.00	0.024
Pre-Hypertension	0.68	0.78	0.235
Hypertension	0.10	0.44	0.011
HDL Cholesterol	53.21	42.00	0.067
Cholesterol/HDL-C	3.89	4.69	0.057
Abdominal Obesity	065	0.87	0.136
Metabolic Syndrome	0.00	0.25	0.009

Source: Campbell et al. (2014).

The GED as a case study of the power of character skills and the costs of neglecting them

James J. Heckman, John E. Humphries, and Tim Kautz
The Myth of Achievement Tests: The GED and the Role of Character in American Life
University of Chicago Press, 2014



THE UNIVERSITY OF CHICAGO

Figure 7: Distribution of Cognitive Ability by Educational Status

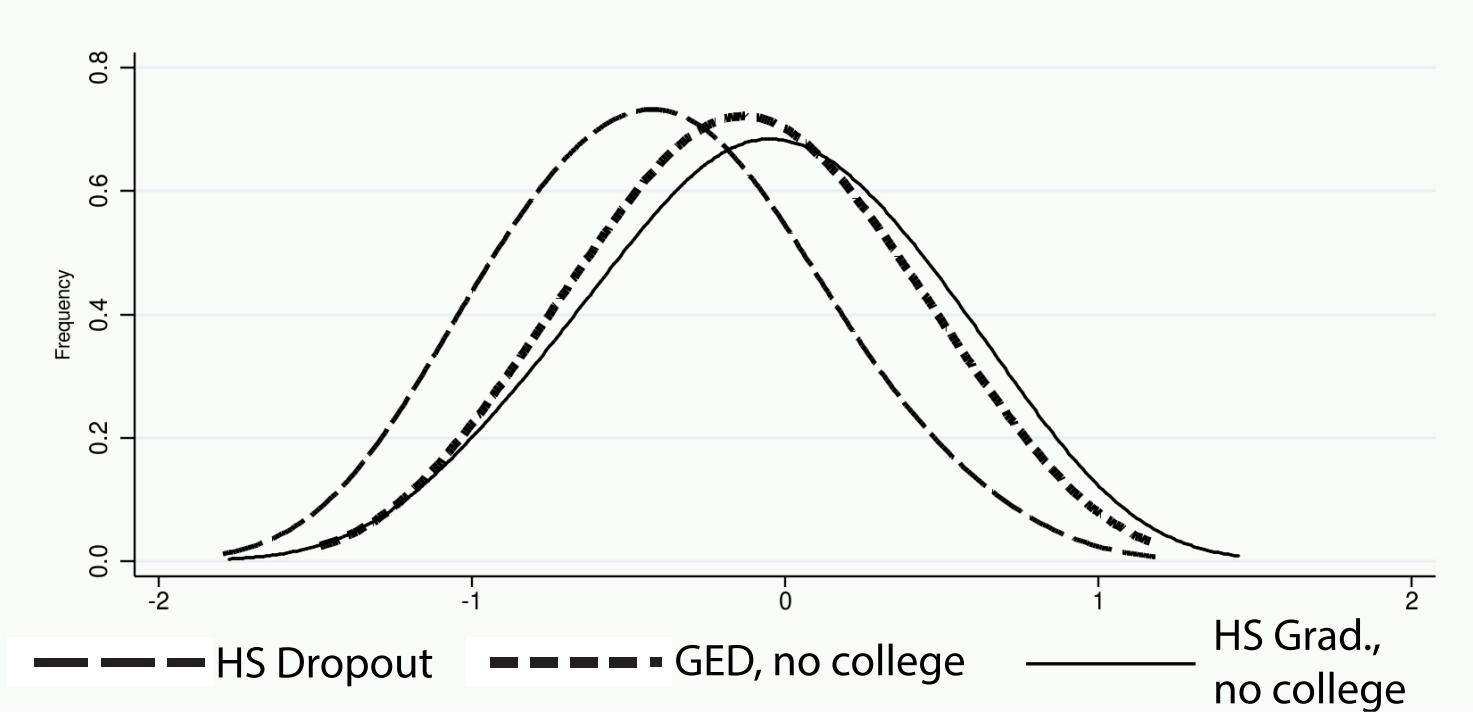
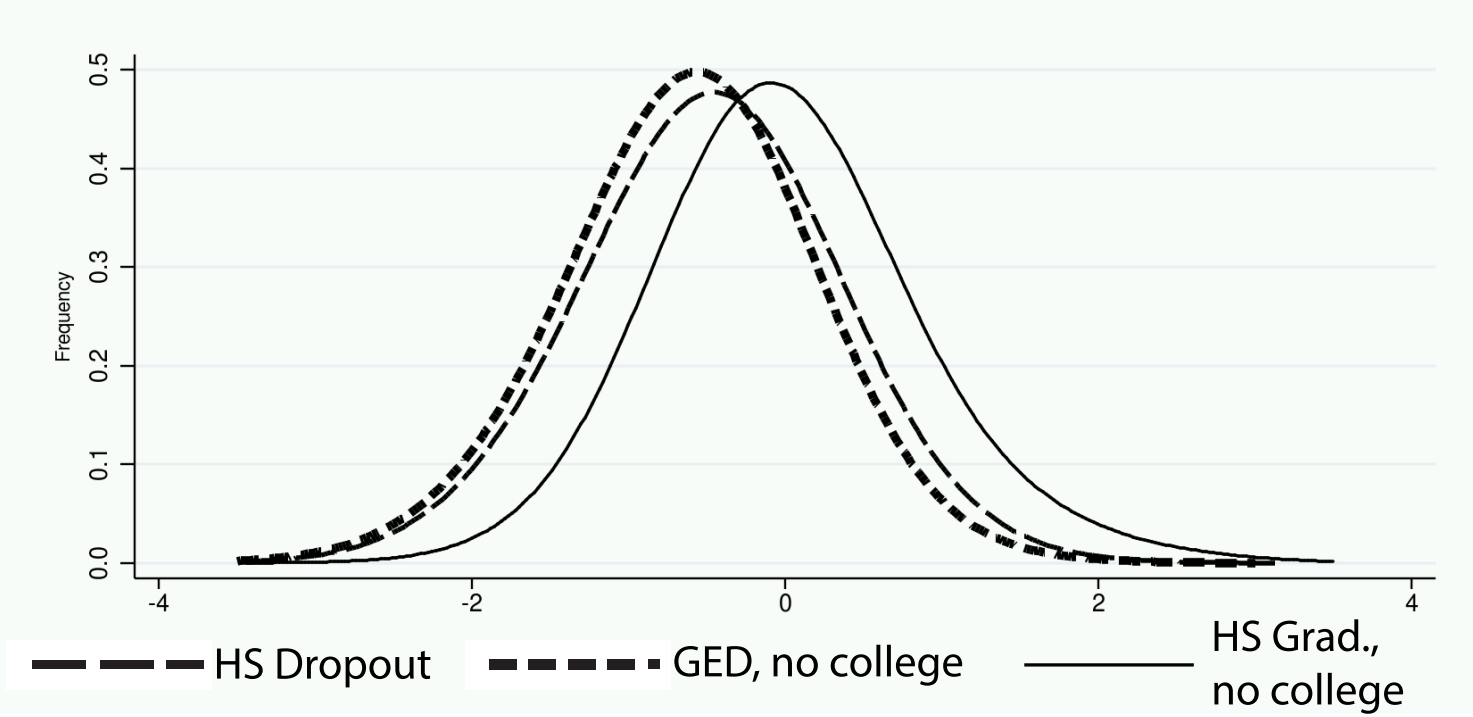
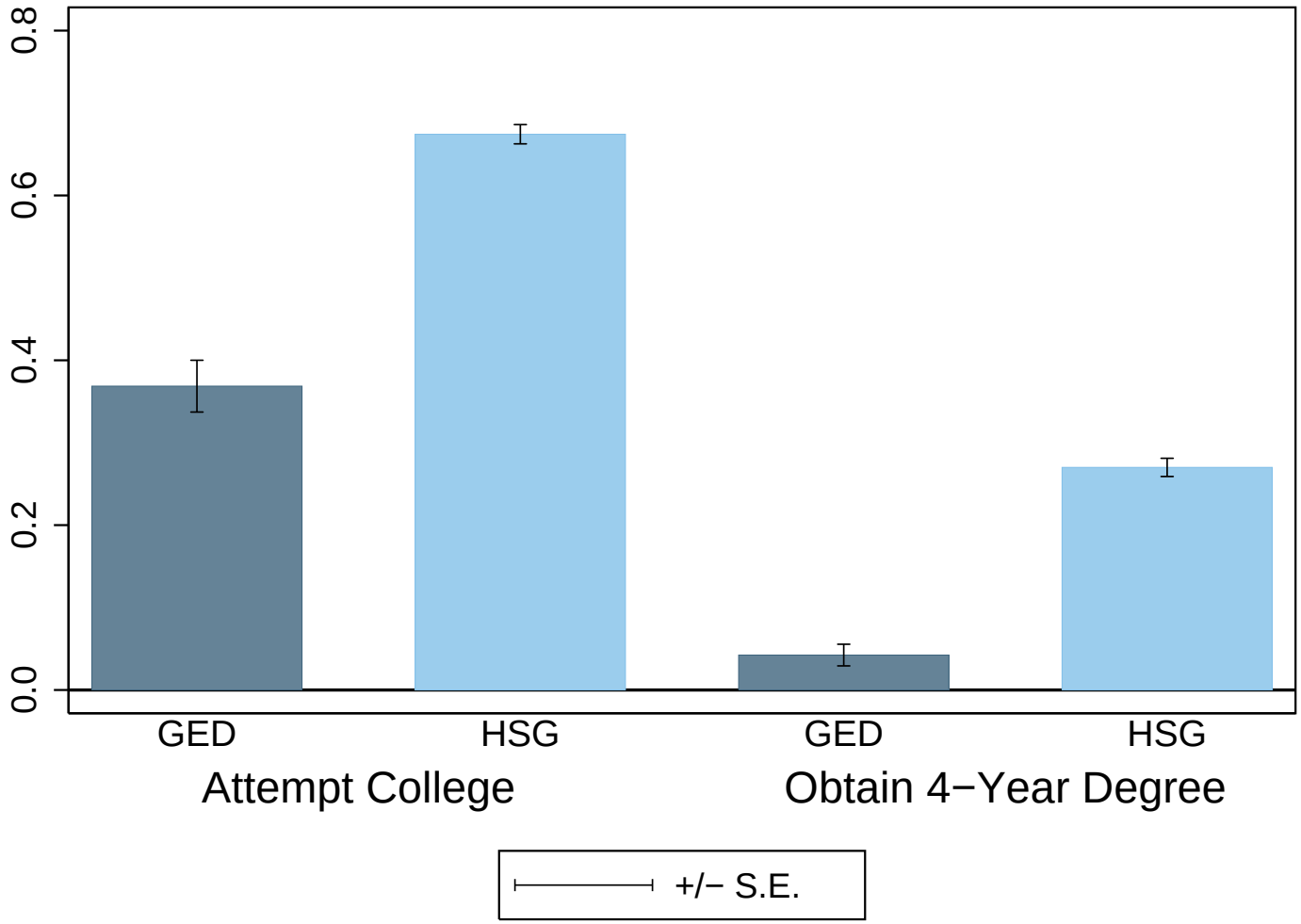


Figure 8: Distribution of Character Skills by Education Group



Measures of Character Skills Based on **Behaviors** in Early Teenage Years: Outcomes are for Adult Years

Figure 9: Postsecondary Educational Attainment across Education Groups through Age 40



- **GEDs earn wages of dropouts controlling for their greater cognitive ability**
- **Drop out of marriage, jobs, military – same rates as dropouts**

THE UNIVERSITY OF CHICAGO

Traditional Approach to Measurement: *The Big Five*

Table 1: The Big Five Traits

OCEAN

	Trait	Definition of Trait
I.	**O**penness to Experience	The tendency to be open to new aesthetic, cultural, or intellectual experiences.
II.	**C**onscientiousness	The tendency to be organized, responsible, and hardworking.
III.	**E**xtraversion	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.
IV.	**A**greeableness	The tendency to act in a cooperative, unselfish manner.
V.	**N**euroticism	Neuroticism is a chronic level of emotional instability and proneness to psychological distress. Emotional stability is predictability and consistency in emotional reactions, with absence of rapid mood changes.

THE UNIVERSITY OF CHICAGO

New approaches to measuring character and cognitive skills go well beyond the Big Five

THE UNIVERSITY OF CHICAGO

Definition of Personality

Personality traits are the relatively enduring patterns of thoughts, feelings, and behaviors that reflect the tendency to respond in certain ways under certain circumstances.

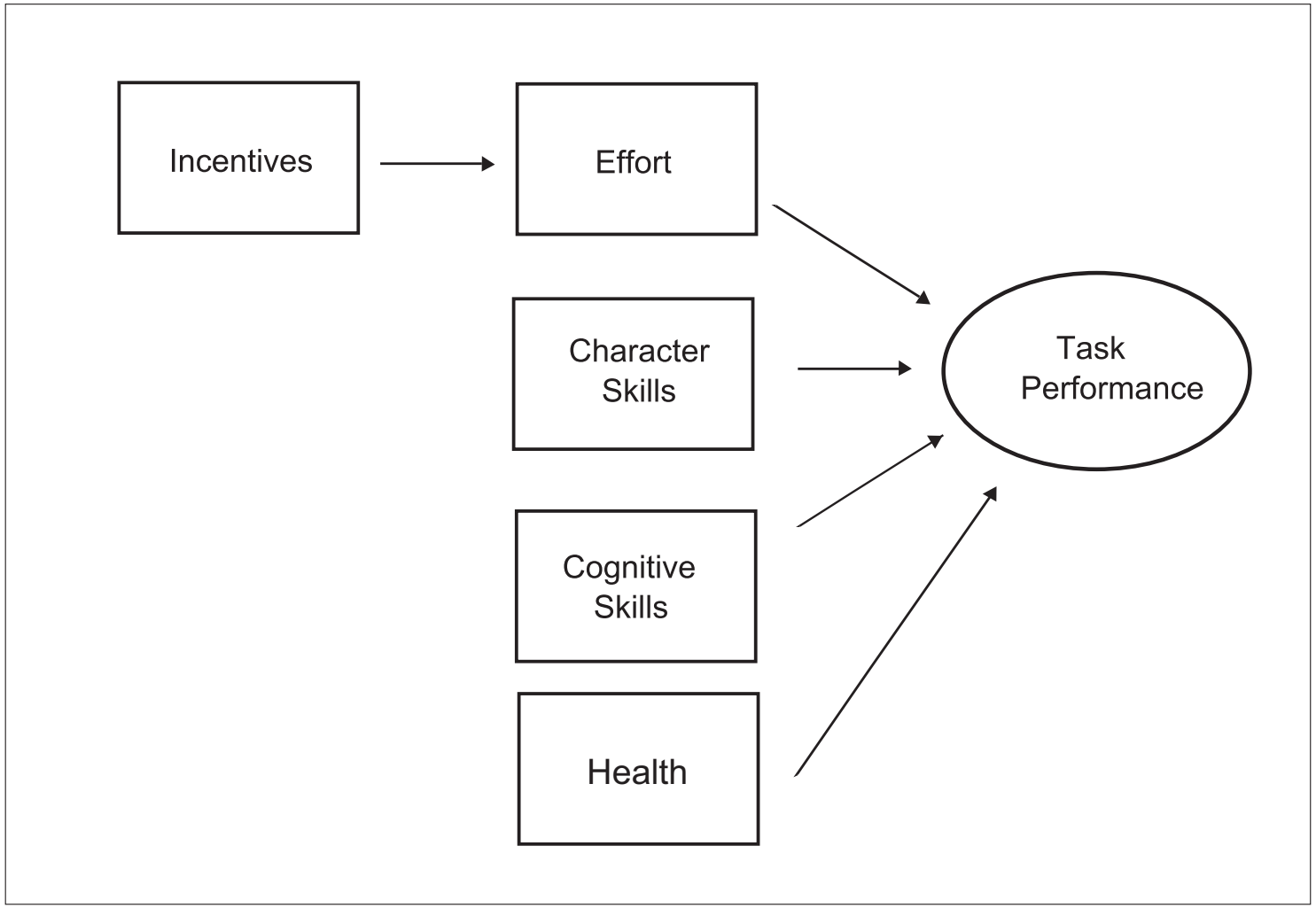
—Roberts (2009, p. 140)

THE UNIVERSITY OF CHICAGO

Distinction between tasks & tests artificial

All tests are tasks

Figure 10: **Determinants of Task Performance**



Broader Notions of Personality Skills

(a) Risk aversion
(b) Trust
(c) Empathy and social preference
(d) Ambiguity aversion
(e) Time preference
(f) Positive and negative reciprocity

THE UNIVERSITY OF CHICAGO

- These are predictive of numerous life outcomes
- They are not closely related to Big Five:
  - They capture dimensions of human capabilities

THE UNIVERSITY OF CHICAGO

- Correlations between Big Five and broader notions of preference

THE UNIVERSITY OF CHICAGO

Table 2: Pearson correlation structure experimental data set

	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism	LoC
Time	0.0370	0.0057	−0.0084	0.1026**	−0.0518	0.0847
Risk	−0.0379	−0.0611	0.0762*	0.0202	−0.1201***	0.0434
Positive reciprocity	0.1724***	0.0140	0.0211	0.2042***	0.0361	0.0152
Negative reciprocity	−0.0885*	−0.0393	0.0943*	−0.1451***	−0.0136	−0.1418**
Trust	0.1232***	−0.1300***	0.0004	0.1665***	−0.0134	−0.0140
Altruism	0.1242**	−0.0979*	0.0249	0.1911***	0.0847*	0.0480

The asterisks indicate significance at the 10% (*), 5% (**), and 1% (***) levels. Correlations between economic preferences and the Big Five were calculated using 394–477 observations. Correlations between economic preferences and the locus of control (LoC) were calculated using 254–315 observations. All measures are standardized.

Source: Becker et al. (2012).

Table 3: Correlation structure between personality measures and economic preferences from SOEP observations

	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism	LoC
Time	0.0183**	0.1122***	−0.0415***	0.3122***	−0.0584***	0.0681***
Risk	0.2793***	−0.0400***	0.2601***	−0.1454***	−0.0996***	0.1521***
Positive reciprocity	0.1814***	0.2520***	0.1473***	0.1842***	0.0872***	0.0954***
Negative reciprocity	−0.0522***	−0.1558***	−0.0264***	−0.3756***	0.0612***	−0.2154***
Trust	0.1272***	−0.0680***	0.0575***	0.0945***	−0.1919***	0.2094***
Altruism	0.1756***	0.1495***	0.1670***	0.2557***	0.0908***	0.0874***

The asterisks indicate significance at the 10% (*), 5% (**), and 1% (***) levels. Correlations are calculated using 14,243 observations. All measures are standardized. Abbreviation: LoC, locus of control.

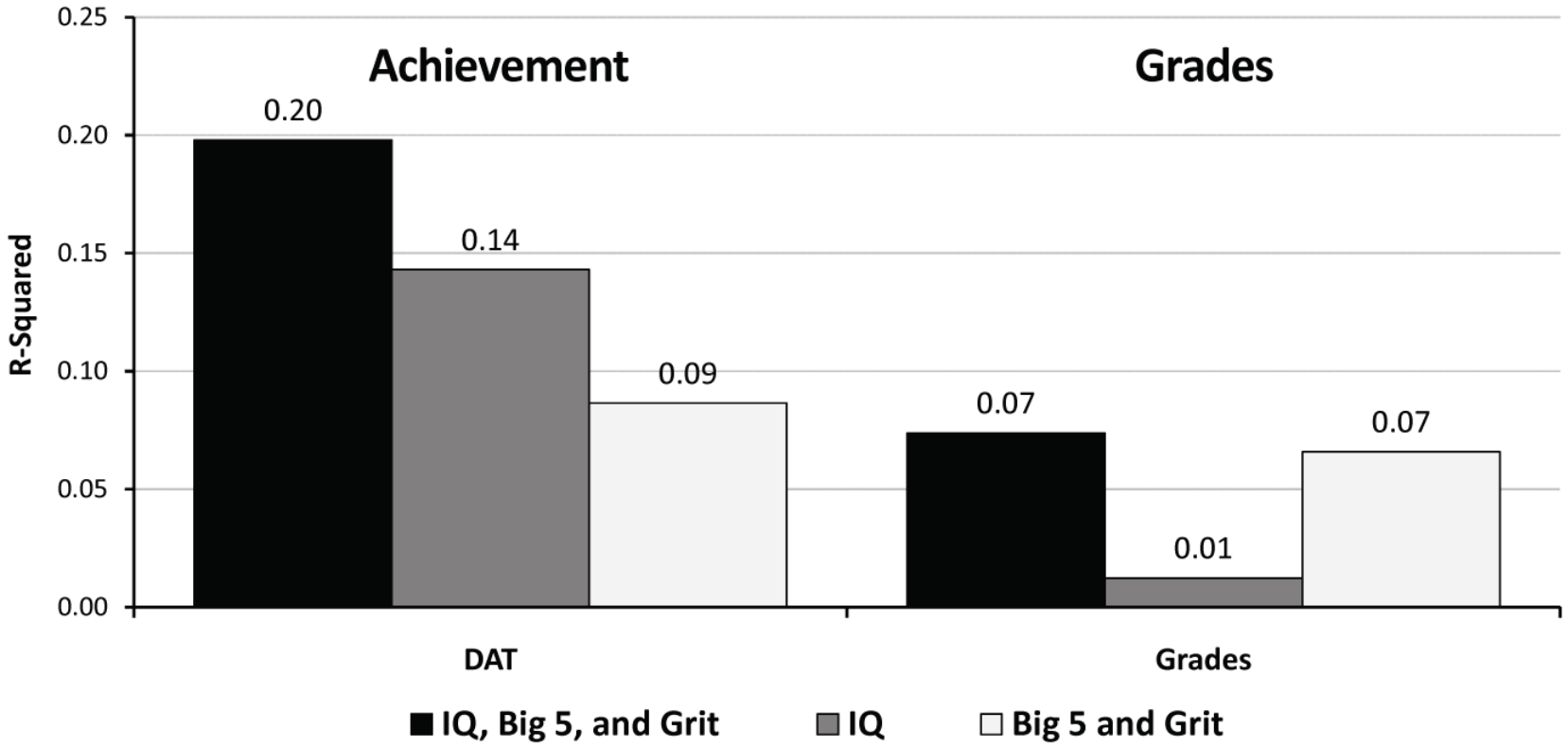
Source: Becker et al. (2012).

Better and often more easily implemented ways to measure personality have been developed

- GPA is a better predictor of success in college than SAT (Bowen et al., 2009)
- Grades capture personality

THE UNIVERSITY OF CHICAGO

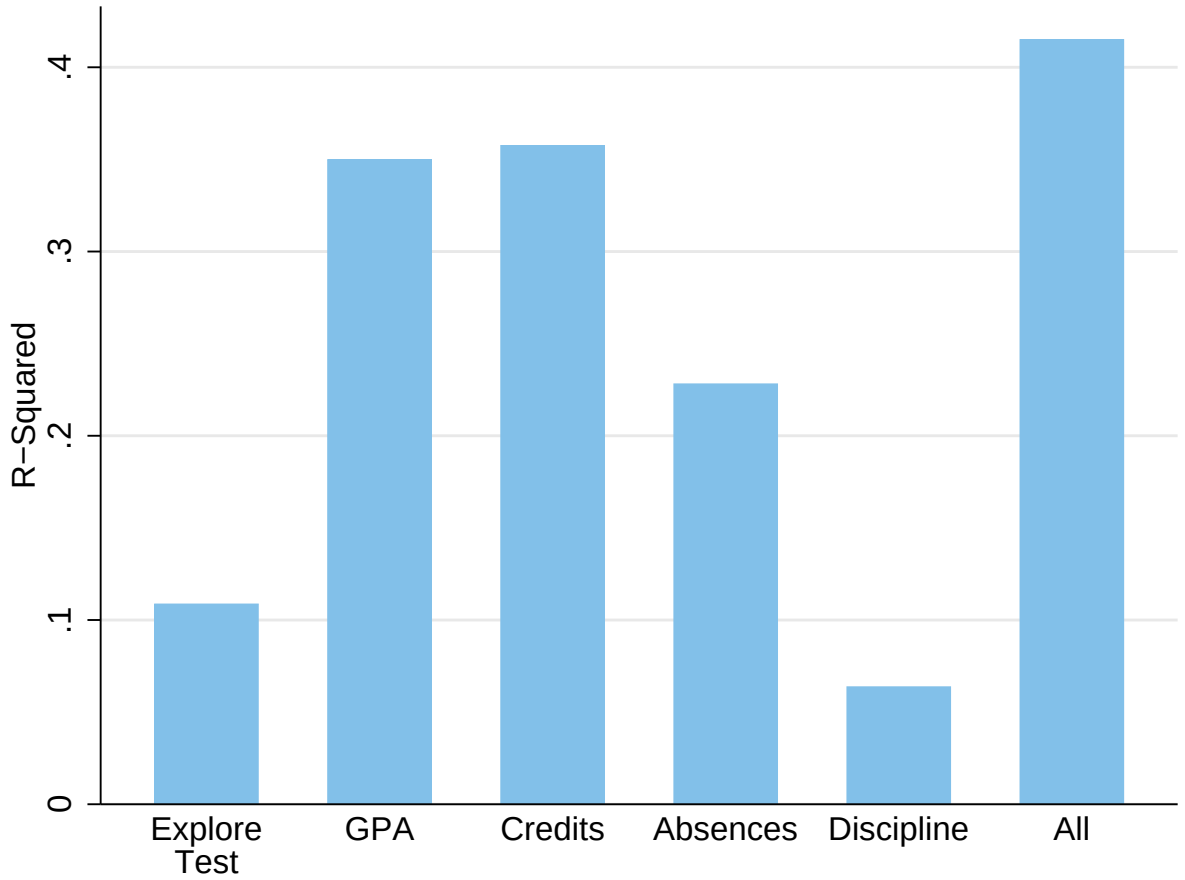
Figure 11: Decomposing Variance Explained for Achievement Tests and Grades into IQ and Character: Stella Maris Secondary School, Maastricht, Holland



Source: Borghans et al. (2011). *Note*: Grit is a measure of persistence on tasks (Duckworth et al., 2007).

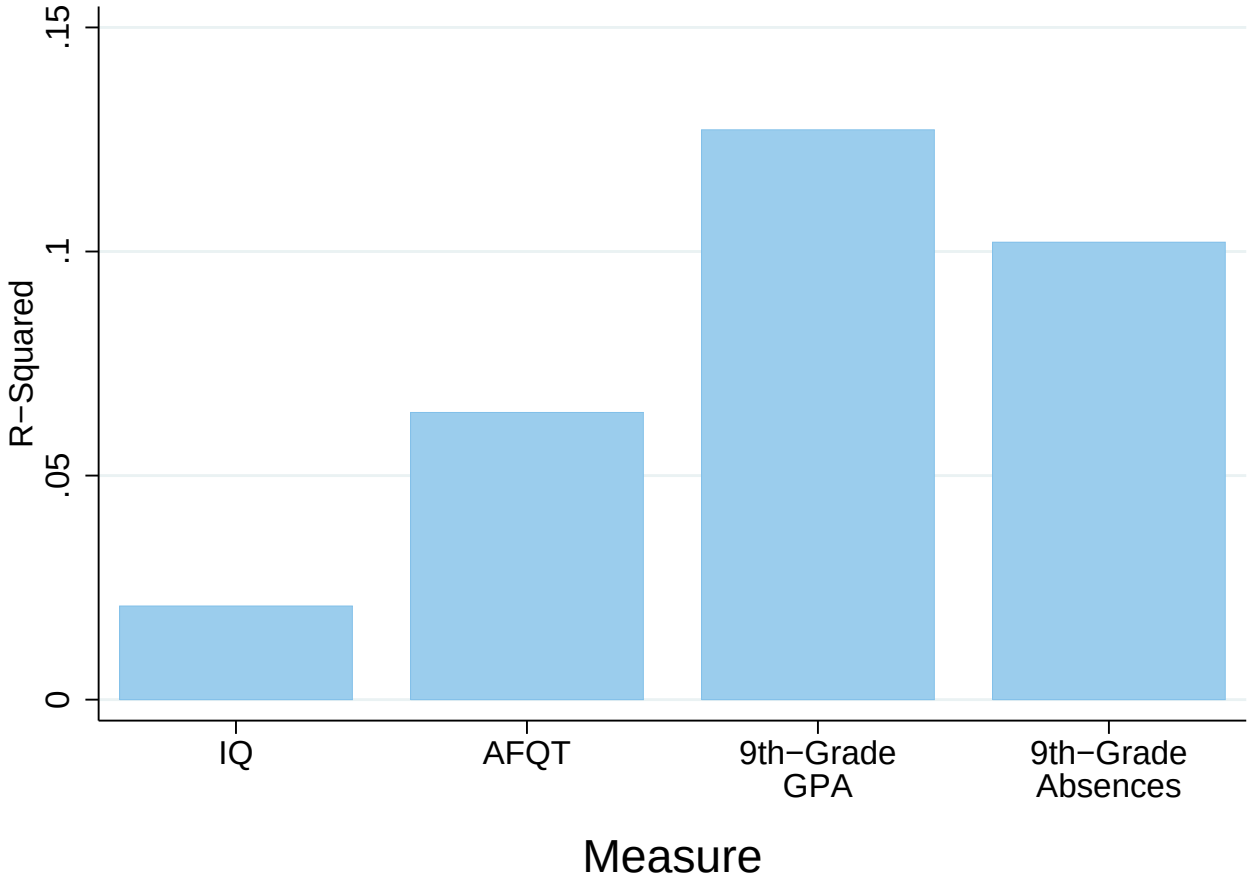
THE UNIVERSITY OF CHICAGO

Figure 12: Predictive Validities of Measures of Cognition and Character in Secondary School Graduation (Explained Variance)



Source: Kautz and Zanoni (2015)

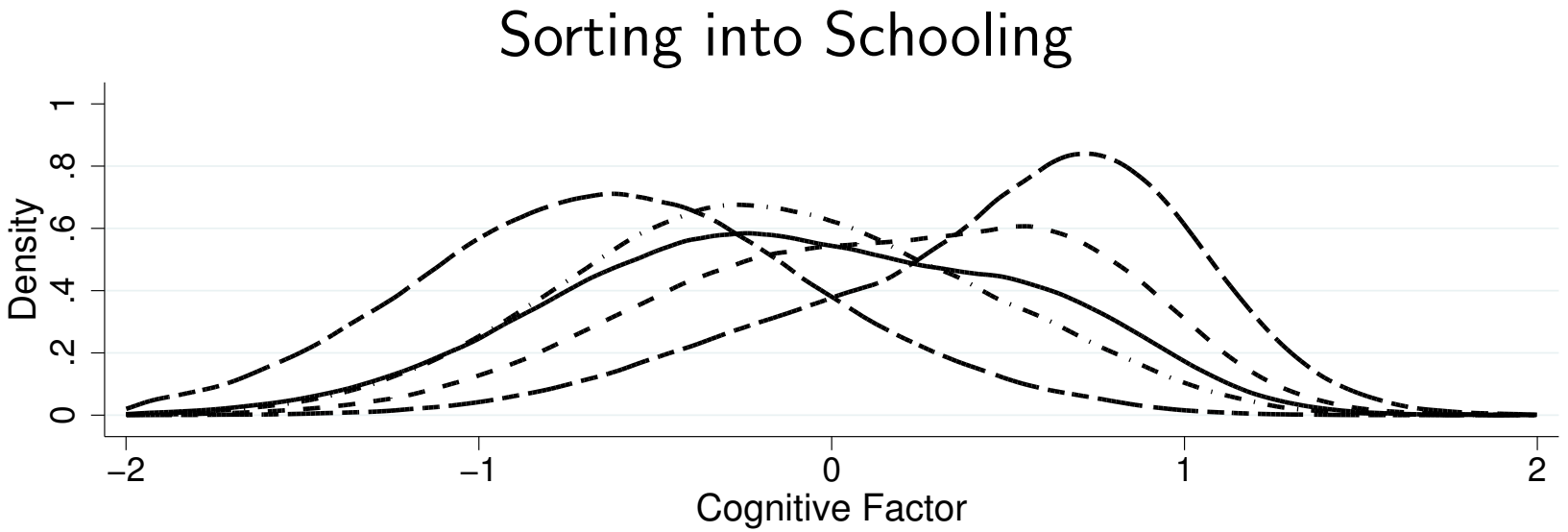
Figure 13: Predictive Validities of Measures of Cognition and Character in Secondary School Graduation (Explained Variance)



Source: Heckman et al. (2015).

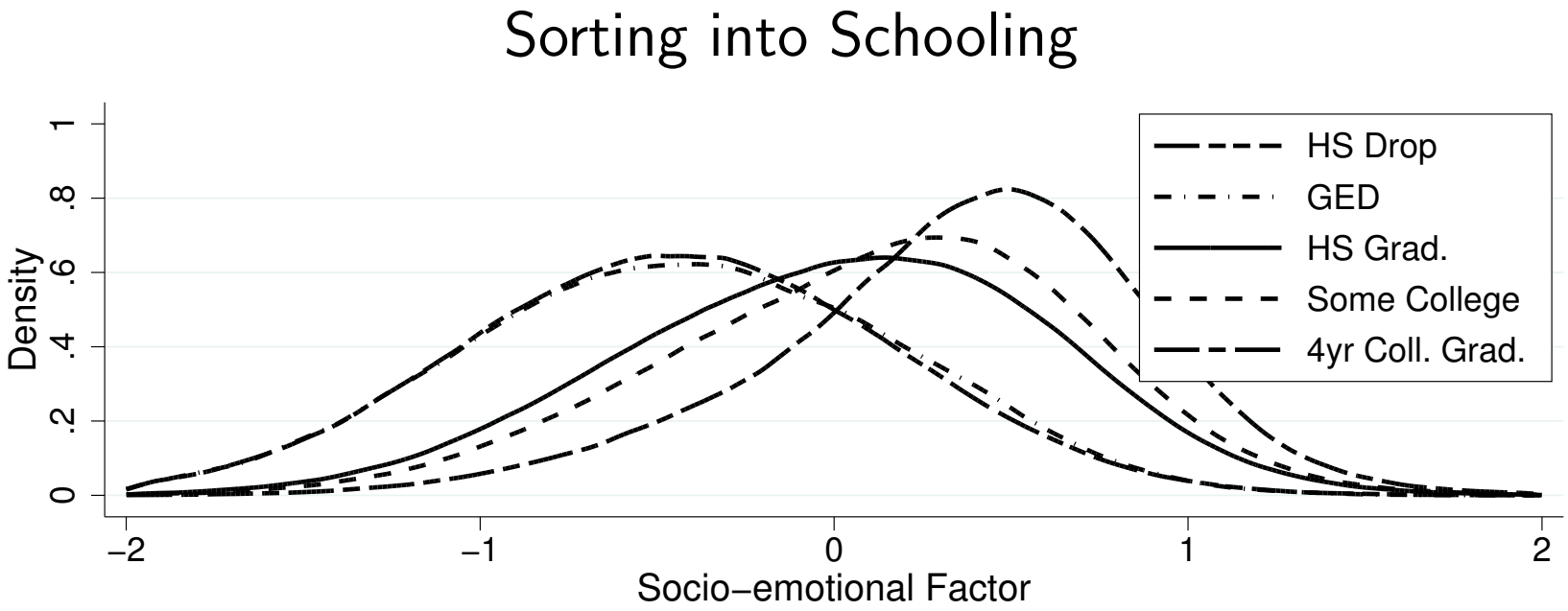
Early Behaviors as Predictors of Later Behaviors

Figure 14: Distribution of factors by schooling level



Note: The factors are simulated from the estimates of the model. The simulated data contain 1 million observations.
Source: Heckman et al. (2015).

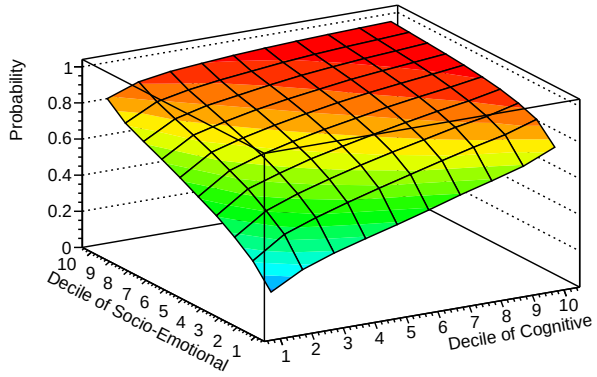
Figure 14: Distribution of factors by schooling level (cont.)

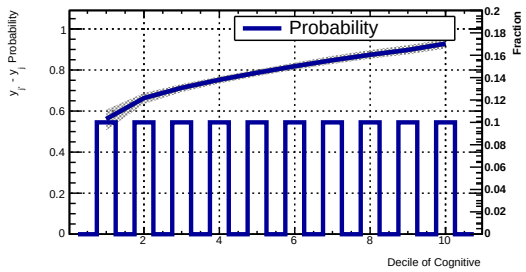


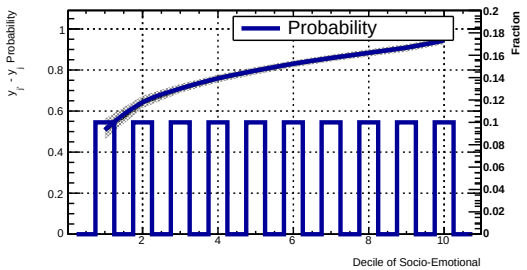
Note: The factors are simulated from the estimates of the model. The simulated data contain 1 million observations.
Source: Heckman et al. (2015).

Figure 15: The Probability of Educational Decisions, by Endowment Levels

A. Dropping from HS vs. Graduating from HS



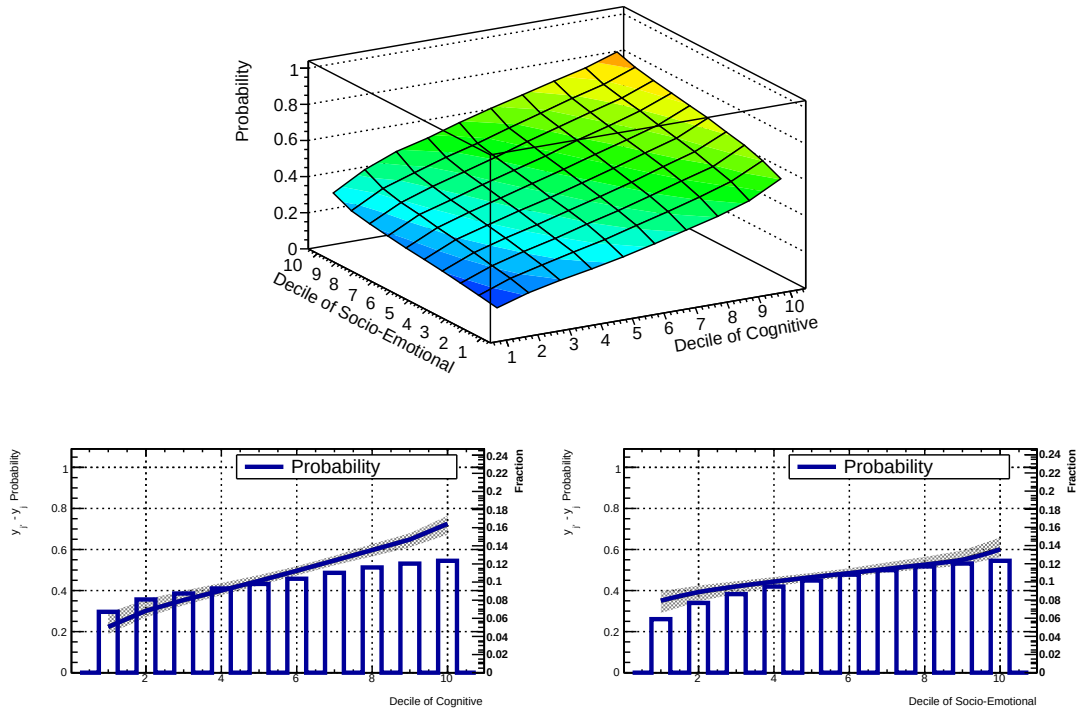




Source: Heckman et al. (2015).

Figure 15: The Probability of Educational Decisions, by Endowment Levels (cont.) (Final Schooling Levels are Highlighted Using Bold Letters)

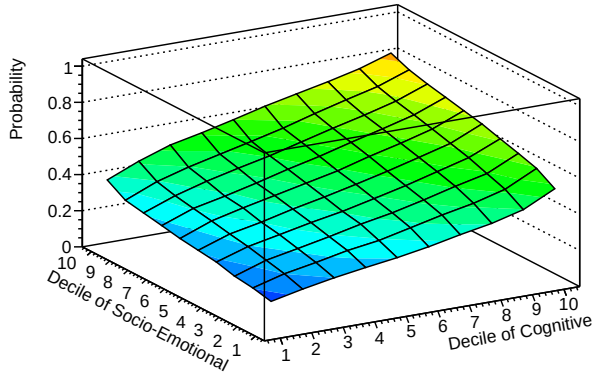
C. **HS Graduate** vs. College Enrollment

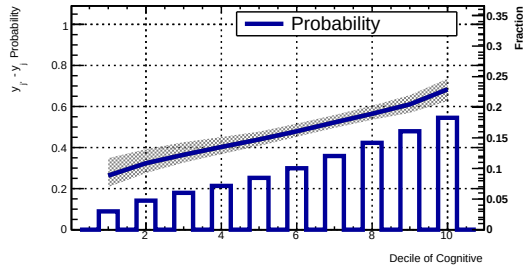


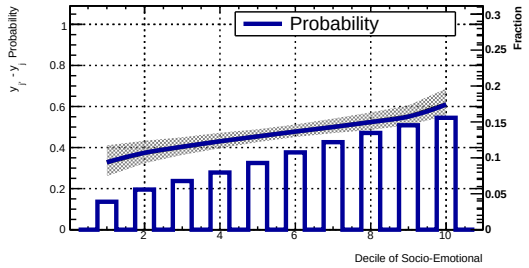
Source: Heckman et al. (2015).

Figure 15: The Probability of Educational Decisions, by Endowment Levels (cont.) (Final Schooling Levels are Highlighted Using Bold Letters)

D. **Some College** vs. **4-year college degree**







Source: Heckman et al. (2015).

Character Can Be Fostered

- See OECD report

Conclusion #1: There is More Evidence that Early Programs are Effective

- Early childhood programs have been shown to be more effective
- Adolescent programs are less well-studied and the evidence is mixed
- Adolescent interventions that teach personality skills in the workplace (or specific context) are promising

THE UNIVERSITY OF CHICAGO

Conclusion #2: Long-Term Follow-Ups are Vital

- Many programs have short-term effects but no long-term effects
- Others have no short-term effects (for some measures) but long-term effects

THE UNIVERSITY OF CHICAGO

Conclusion #3: Non-Cognitive Skills are an Important Channel Throughout Early Childhood and Adolescence

- Only interventions that started before age 3 had a long-term effect on IQ
- Many interventions starting after age 3 have effectively improved outcomes by improving non-cognitive skills
- Adolescent interventions that teach personality skills in the workplace (or specific context) are promising

THE UNIVERSITY OF CHICAGO

Summary

THE UNIVERSITY OF CHICAGO

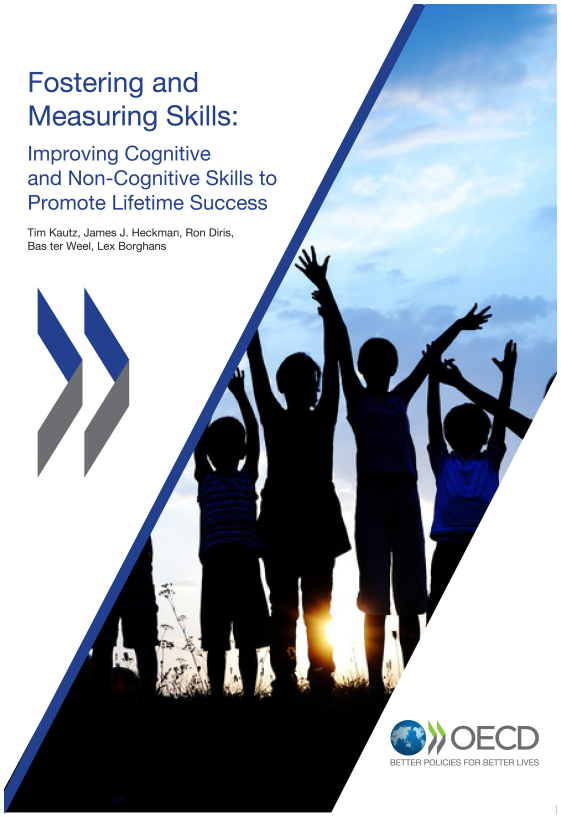
- Many important life skills not captured by scores on achievement tests
- A sole focus on achievement test scores (e.g., NCLB, PISA and Iowa tests) give an incomplete picture of what schools, families, and communities do and how to evaluate schools and other life cycle interventions.
- Socioemotional skills—character, etc.—are important
- These skills can be measured
- They are malleable, and there are effective interventions to promote them
- Soft skills more malleable than cognitive skills at later ages

THE UNIVERSITY OF CHICAGO

- Older ("established") measurement systems such as the Big Five do **not** capture the rich range of behaviors and traits that children and adults exhibit
- Need comprehensive measures of traits
- Instead of relying exclusively or mainly on self-reported "Big Five measures," we should use approaches based on behaviors
  - Teacher reports and assessments as encoded in school system records and interviews
  - Eliciting preference parameters from observed choices in the field and in controlled choice experiments and behaviors:
    - Risk aversion
    - Time preference
    - Ambiguity aversion
    - Trust
    - Reciprocity (positive and negative)
- For all measurement systems, we should adjust for incentives and other traits

THE UNIVERSITY OF CHICAGO

James J. Heckman,Tim Kautz, Ron Diris, Bas ter Weel, and Lex Borghans
Fostering and Measuring Skills: Improving Cognitive and Noncognitive Skills to Promote Lifetime Success
OECD, 2014



THE UNIVERSITY OF CHICAGO