Interpreting Personality Measurement

James J. Heckman
University of Chicago

Bridget Galaty

Econ 350, Winter 2023

THE UNIVERSITY OF CHICAGO

How do Psychologists Define Personality Traits?

- *"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)

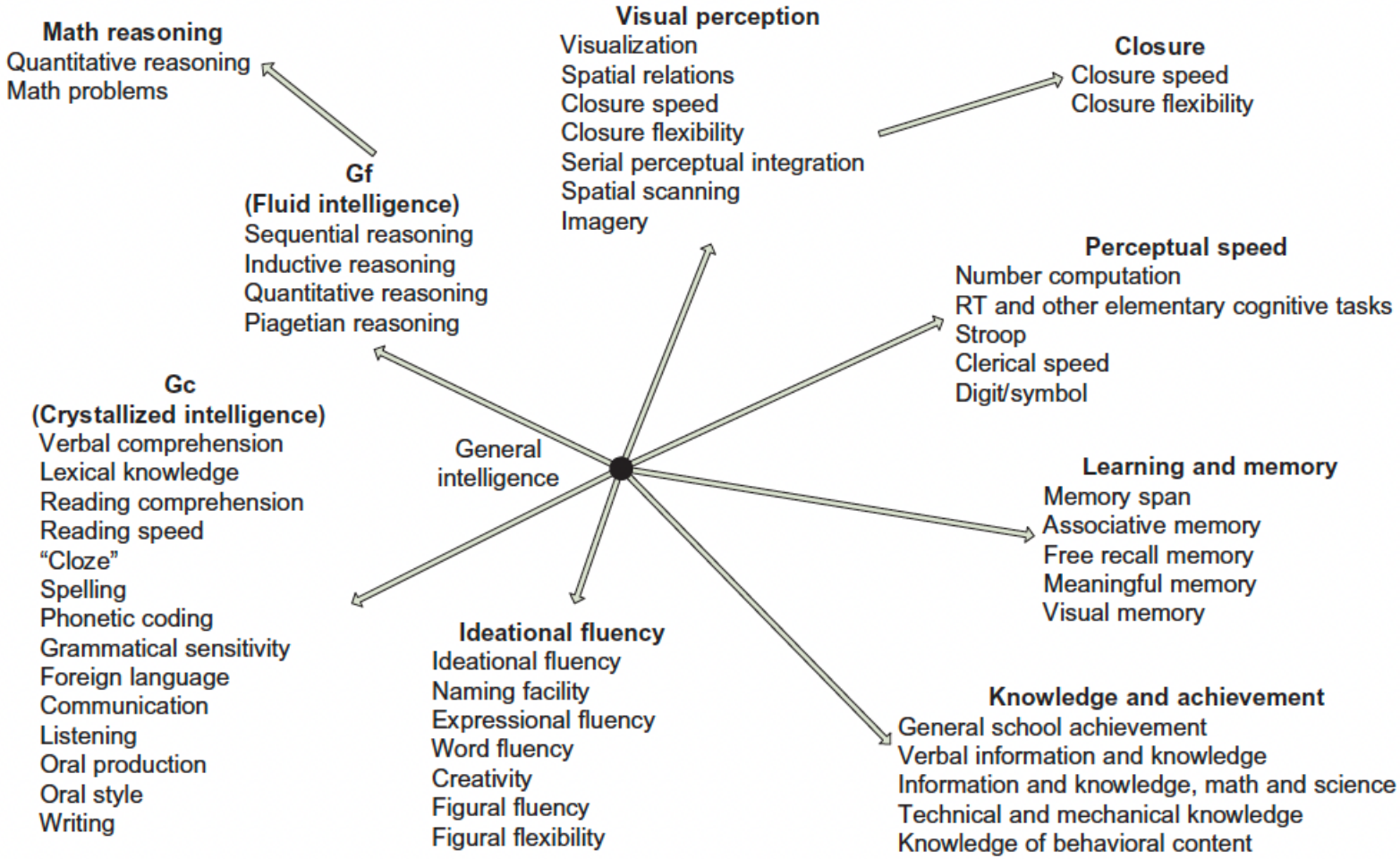


Figure 1: A Hierarchical Scheme of General Intelligence and its Components

Source: Almlund et al. (2011)

The Big 5 Personality Measure

Trait	Definition of Trait*
I. Openness to Experience (Intellect)	The tendency to be open to new aesthetic, cultural, or intellectual experiences.
II. Conscientiousness	The tendency to be organized, responsible, and hardworking.
III. 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.
IV. Agreeableness	The tendency to act in a cooperative, unselfish manner.
V. Neuroticism (Emotional Stability)	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.

*From the American Psychological Association Dictionary (2007).

Figure 2: The Big Five Traits

Source: Almlund et al. (2011)

Big 5 Domains and Facets

The Big Five Domains and their Facets

Factor	Facets	Definition of Factor	ACL[a] Marker Items for Factor
I. Openness to Experience (Intellect)	Fantasy, Aesthetics, Feelings, Actions, Ideas, Values	The degree to which a person needs intellectual stimulation, change, and variety.	Commonplace, Narrow-interest, Simple- vs. Wide-interest, Imaginative, Intelligent
II. Conscientiousness	Competence, Order, Dutifulness, Achievement striving, Self-discipline, Deliberation	The degree to which a person is willing to comply with conventional rules, norms, and standards.	Careless, Disorderly, Frivolous vs. Organized, Thorough, Precise
III. Extraversion	Warmth, Gregariousness, Assertiveness, Activity, Excitement seeking, Positive emotions	The degree to which a person needs attention and social interaction.	Quiet, Reserved, Shy vs. Talkative, Assertive, Active
IV. Agreeableness	Trust, Straight-forwardness, Altruism, Compliance, Modesty, Tender-mindedness	The degree to which a person needs pleasant and harmonious relations with others.	Fault-finding, Cold, Unfriendly vs. Sympathetic, Kind, Friendly
V. Neuroticism (Emotional Stability)	Anxiety, Angry hostility, Depression, Self-consciousness, Impulsiveness, Vulnerability	The degree to which a person experiences the world as threatening and beyond his/her control.	Tense, Anxious, Nervous vs. Stable, Calm, Contented

Source: Costa and McCrae (1992b) and Hogan and Hogan (2007).
Note: a. ACL = Adjective Check List (Gough and Heilbrun 1983).

Figure 3: The Big Five Domains and Facets

Borghans et al. (2008)

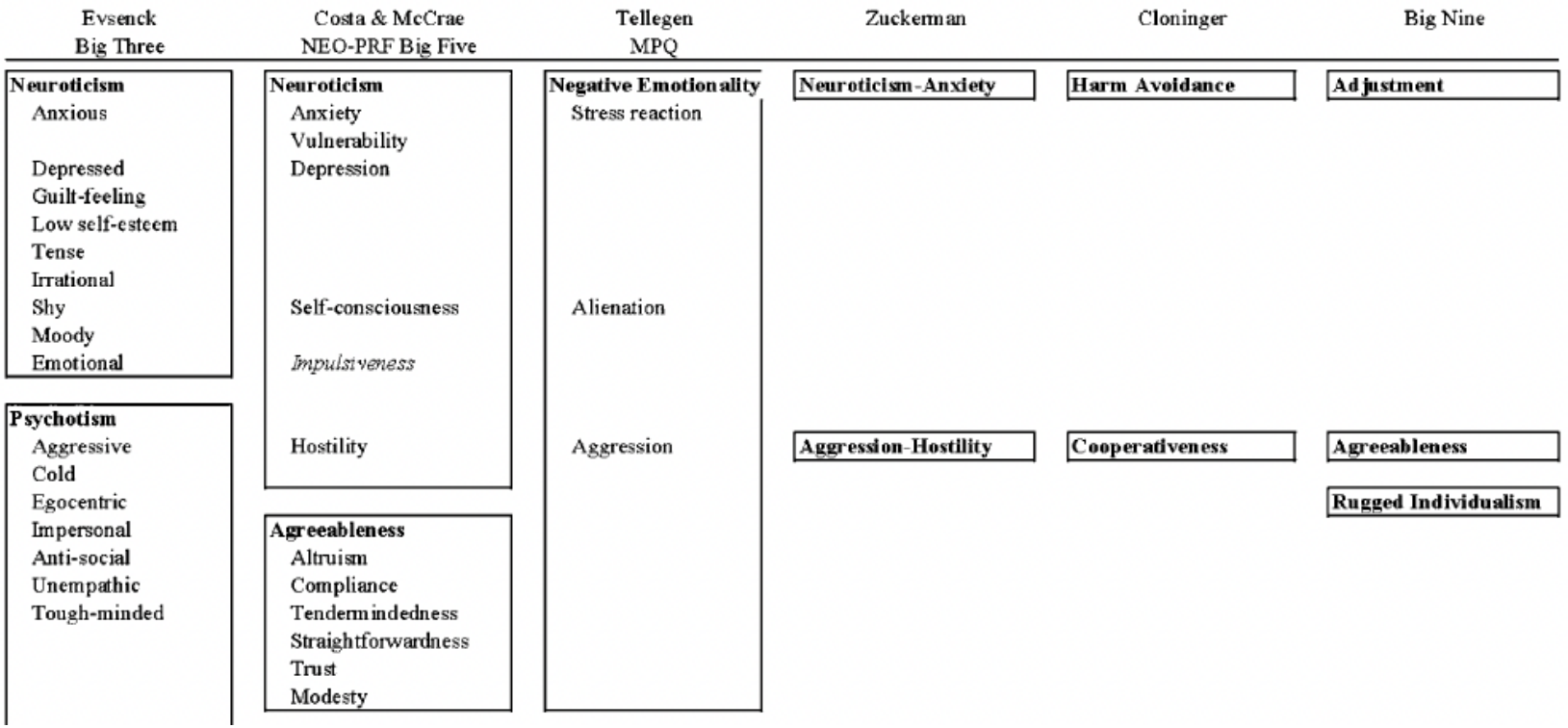


Figure 4: Competing Taxonomies of Personality

Borghans et al. (2008), Cont'd

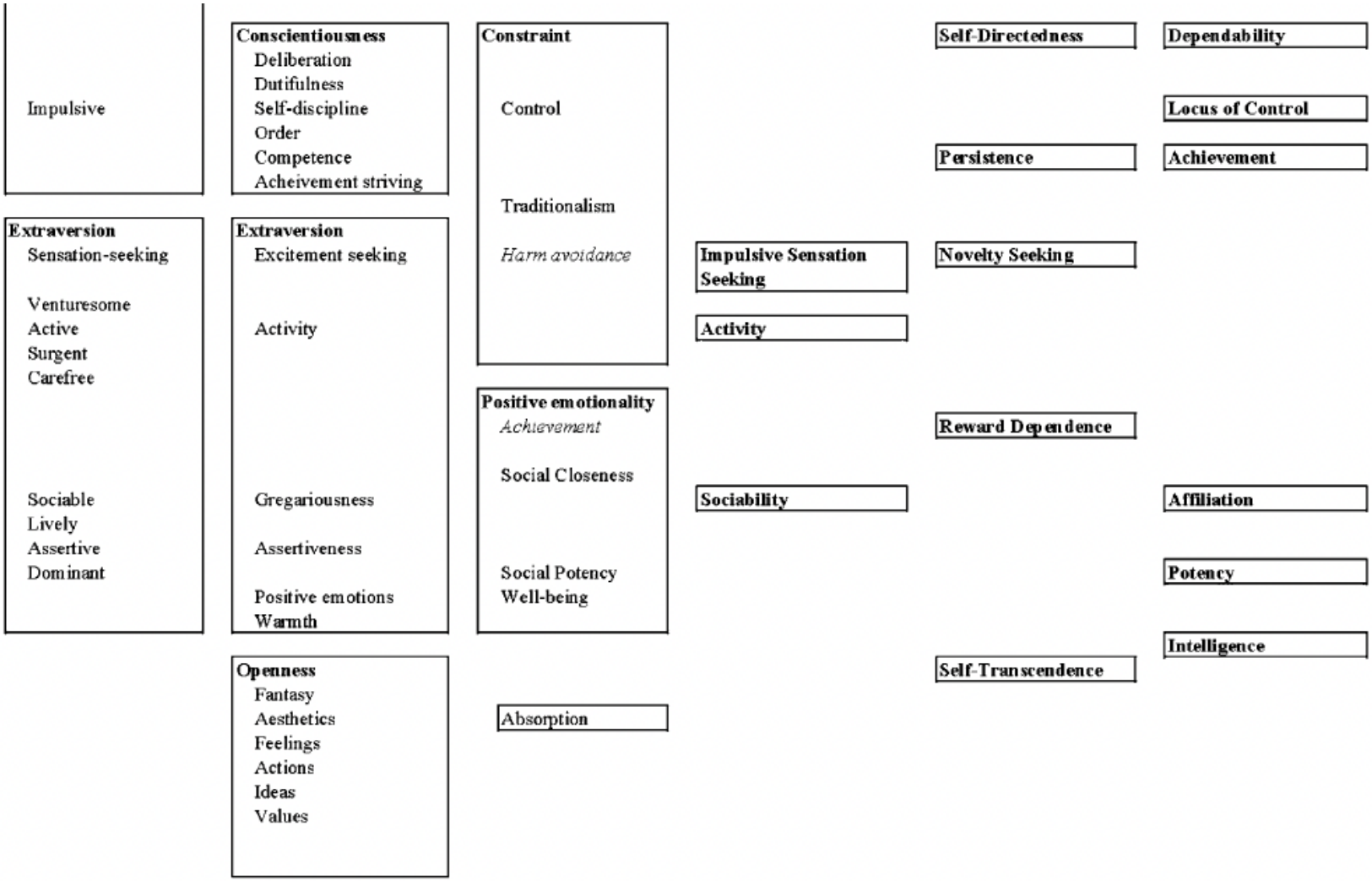


Figure 4: Competing Taxonomies of Personality (Cont'd)

HEXACO

- The HEXACO model of personality structure was first proposed in the early 2000s, and it has been increasingly widely used as an organizing framework in personality research.
- This model posits that personality traits can be summarized by six dimensions: Honesty–Humility (H), Emotionality (E), eXtraversion (X), Agreeableness (A), Conscientiousness (C), and Openness to Experience (O).
- As with the five-factor model, the HEXACO model originated from research based on the lexical approach to personality structure.

- The theoretical interpretation of the six HEXACO personality factors categorizes them into two broad conceptual groups.
  1. The Extraversion, Conscientiousness, and Openness to Experience dimensions represent individual differences in engagement within three different domains of endeavor: social, work-related, and idea-related.
  2. The Honesty–Humility, Emotionality, and Agreeableness dimensions represent individual differences in three different forms of altruistic tendencies.

Behavioral, Emotional, and Social Skills Inventory (BESSI)

Skill facet	Primary Big Five domain (Study 1)	Primary BESSI domain (Study 3)	Capacities used to...
Social Engagement Skills			Actively engage with other people.
Leadership Skill	E	SE	Assert one's views and speak in a group.
Conversational Skill	E	SE	Initiate and maintain social interactions.
Expressive Skill	E	SE	Communicate one's thoughts and feelings to other people.
Persuasive Skill	A-	SE	Present arguments effectively.
Energy Regulation	E	SE/SM	Channel energy in a productive way.
Cooperation Skills			Maintain positive social relationships.
Perspective-Taking Skill	A	C	Understand other people's thoughts and feelings.
Capacity for Social Warmth	A	C	Evoke positive social responses from other people.
(Capacity for Friendliness)	E	—	(Merged with capacity for social warmth.)
Capacity for Trust	A	C	Trust and forgive other people.
Teamwork Skill	A	C	Work with others to achieve shared goals.
Ethical Competence	A	C/SM	Behave ethically, even in difficult circumstances.

Figure 5: A Hierarchical Taxonomy of Behavioral, Emotional, and Social Skills

Self-Management Skills			Effectively pursue goals and complete tasks.
Task Management	C	SM	Work persistently to complete tasks and achieve goals.
Responsibility Management	C	SM	Fulfill promises and commitments.
Organizational Skill	C	SM	Organize personal spaces and objects.
[Time Management]	—	SM	Use time effectively while accomplishing goals.
Detail Management	C	SM	Do careful and thorough work.
Goal Regulation	C	SM	Set clear and ambitious personal goals.
Rule-Following Skill	C	SM	Follow instructions, rules, and norms.
Decision-Making Skill	C	SM	Make well-reasoned decisions.
Capacity for Consistency	O-	SM	Reliably perform routine tasks.

Figure 5: A Hierarchical Taxonomy of Behavioral, Emotional, and Social Skills (Cont'd)

Source: Soto et al. (in press)

Temperament

	Definition of temperament =	Emotionality	Extraversion	Activity	Persistence
Thomas and Chess	Stylistic aspects of behaviour	Negative emotionality	Social inhibition	Activity level	Task persistence
Buss and Plomin	Early-appearing, heritable aspects of personality	Emotionality	Sociability Shyness	Activity	
Rothbart	Reactive and self-regulatory aspects of behaviour	Negative affectivity	Surgency	Surgency	Effortful control

Source: Adapted from De Pauw and Mervielde (2010[171]), "Temperament, personality and developmental psychopathology: A review based on the conceptual dimensions underlying childhood traits",

Figure 6: Common Dimensions of Temperament

Source: Chernyshenko et al. (2018)

Non-taxonomized Personality attributes

- **Creativity:** Creativity and curiosity have been shown to cultivate skills that contribute to the invention and adoption of new processes and products.
- **Grit:** Grit is characterized by a combination of passion and perseverance for a singularly important goal.
- **Change:** The ability to adapt to changing situations or incentive schemes.

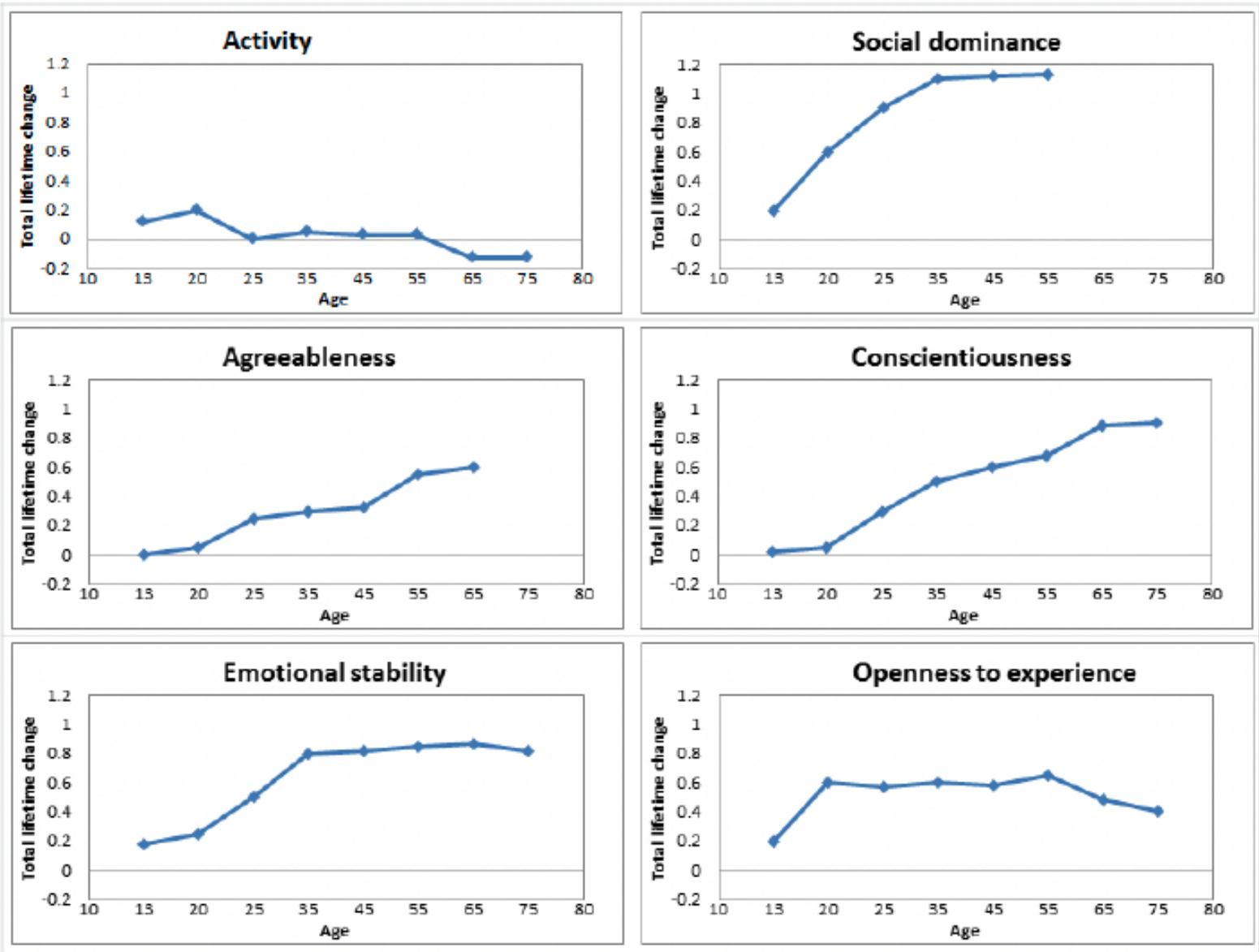


Figure 7: Cumulative Average-Level Changes in Personality Throughout the Life Span

- Figure 8 shows mean-level changes in cognitive skills using a longitudinal analysis, and the bottom panel of Figure 8 shows mean-level changes using a cross-sectional analysis.

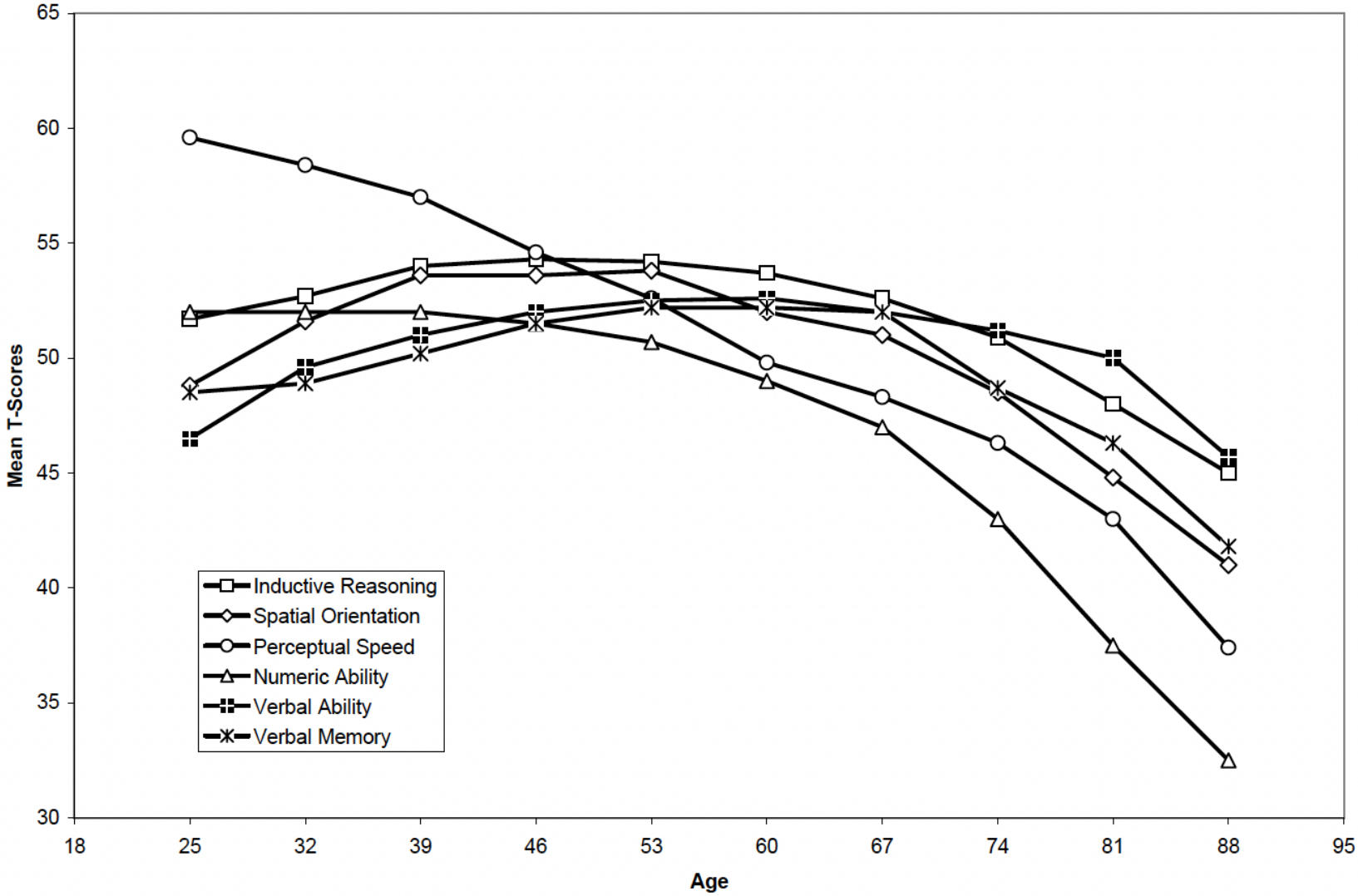


Figure 8: Lifecycle Trends in Cognition

Source: Borghans at al. (2008)

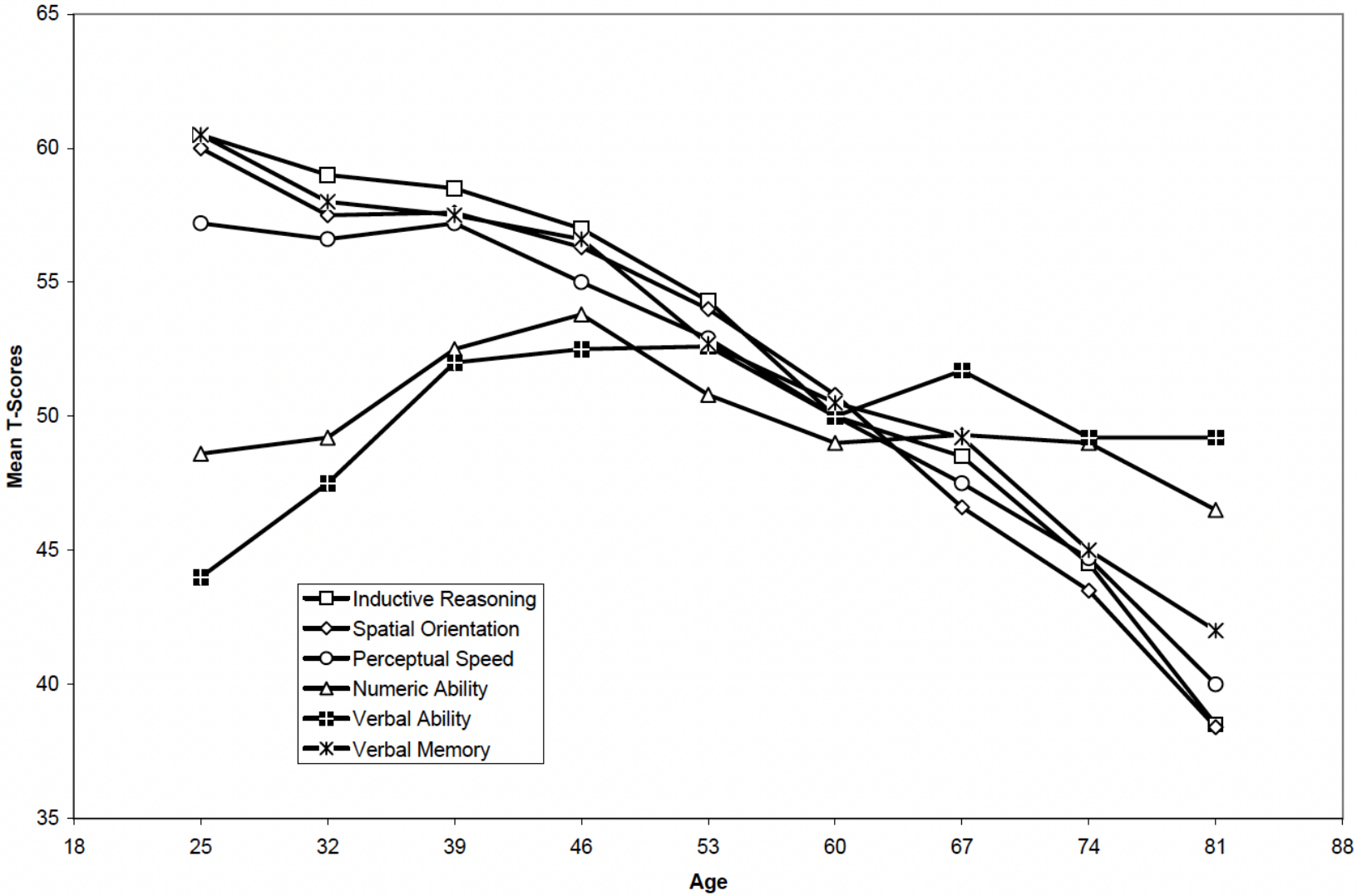
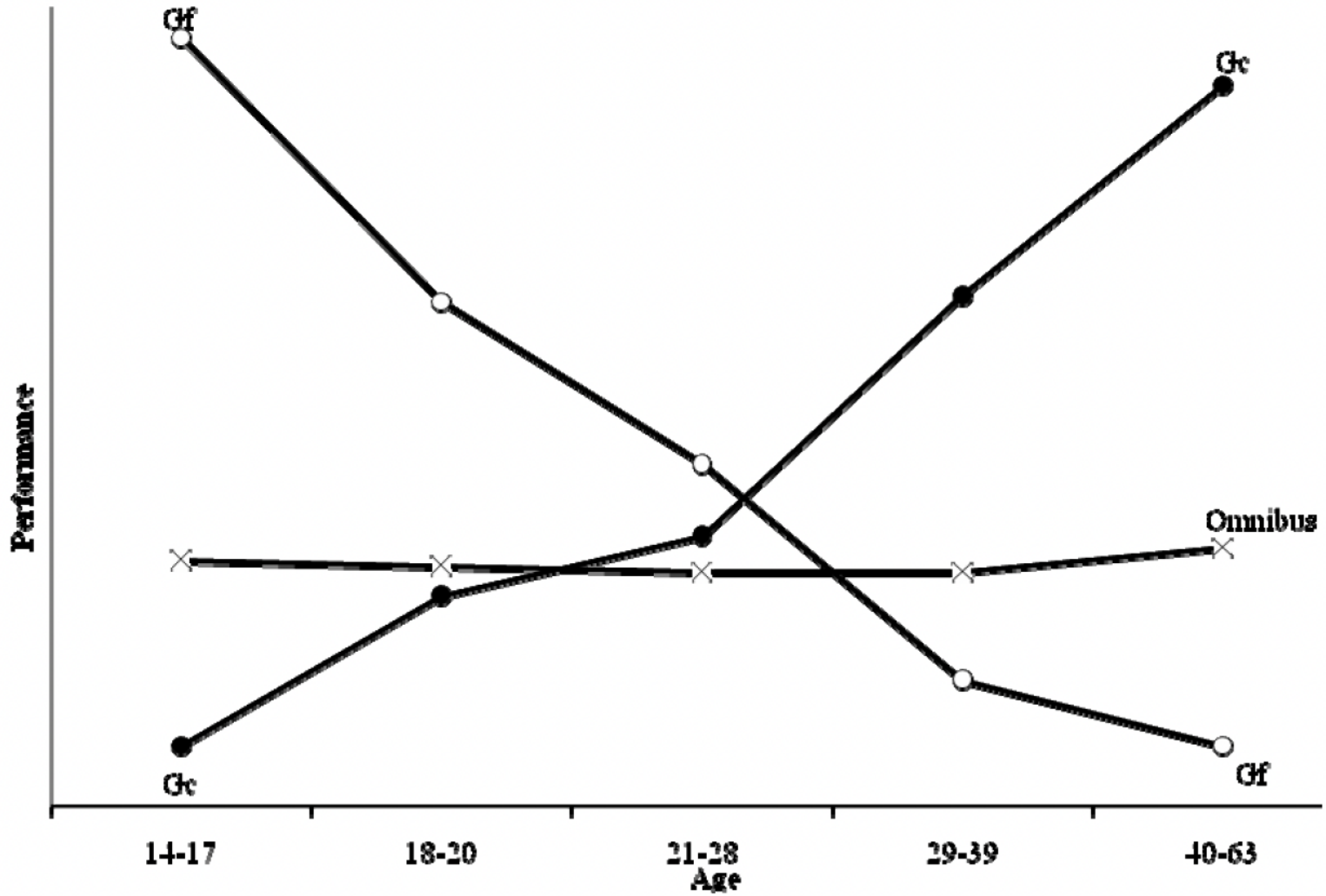


Figure 8: Lifecycle Trends in Cognition (Cont'd)

Source: Borghans at al. (2008)



Source: Borghans at al. (2008)

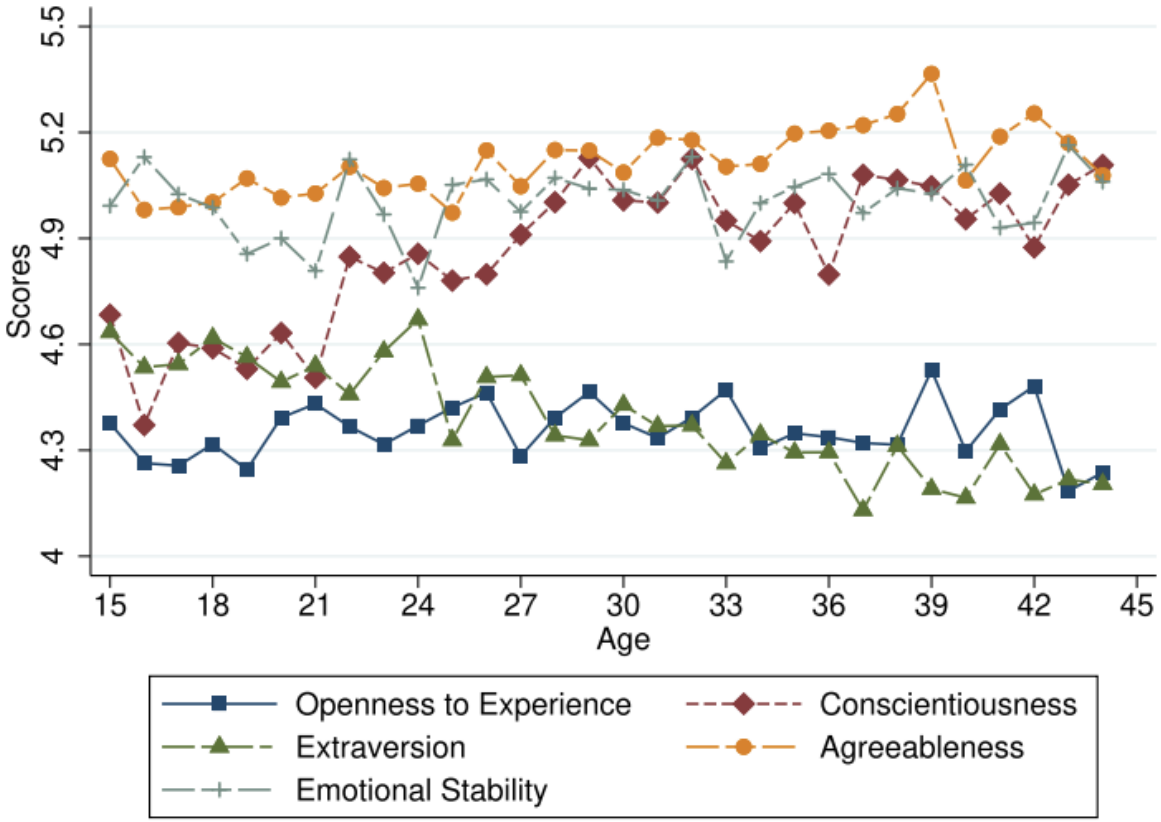


Figure 9: The "Big Five" Personality Trait Scores by Age

Source: Todd & Zhang (2020)

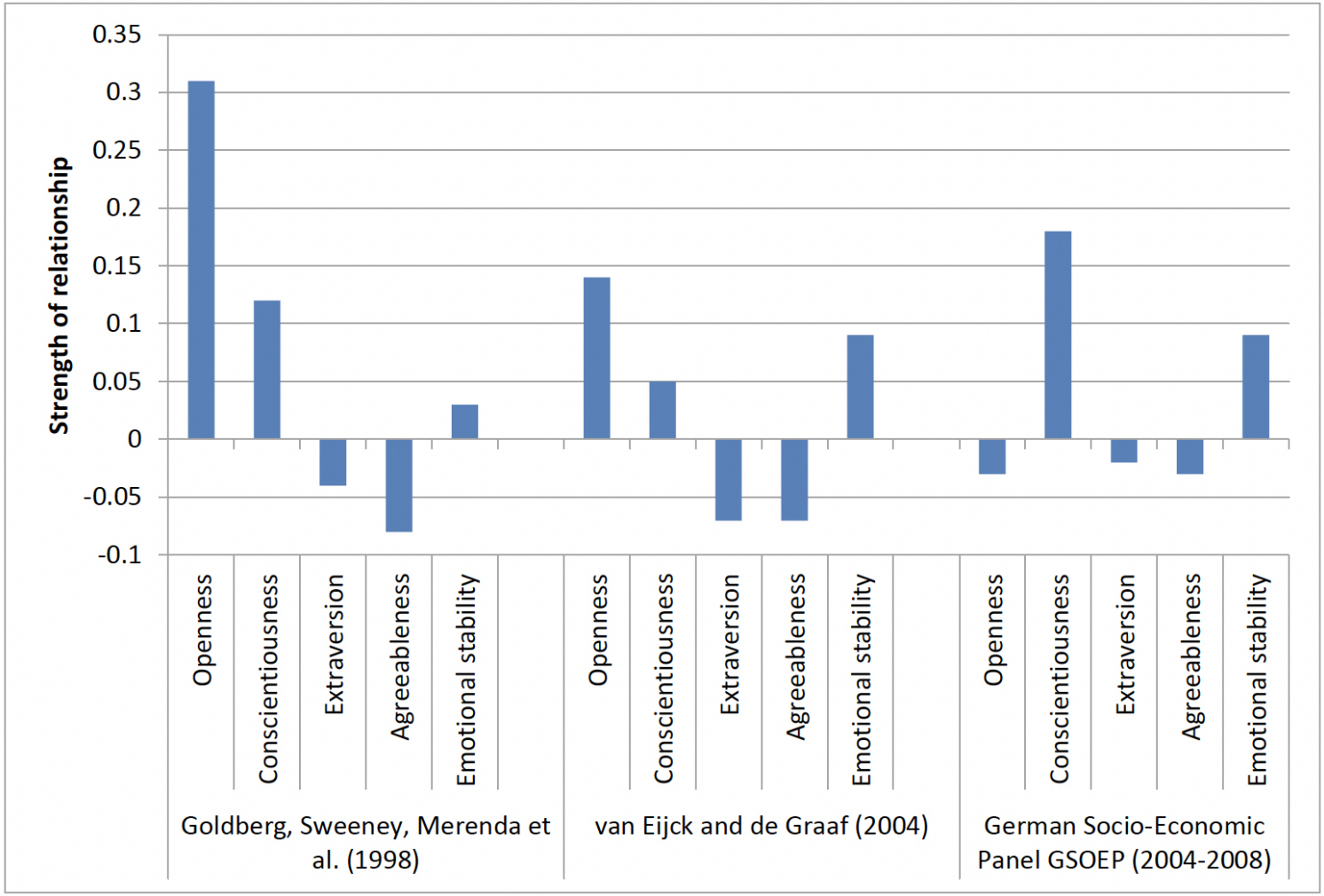
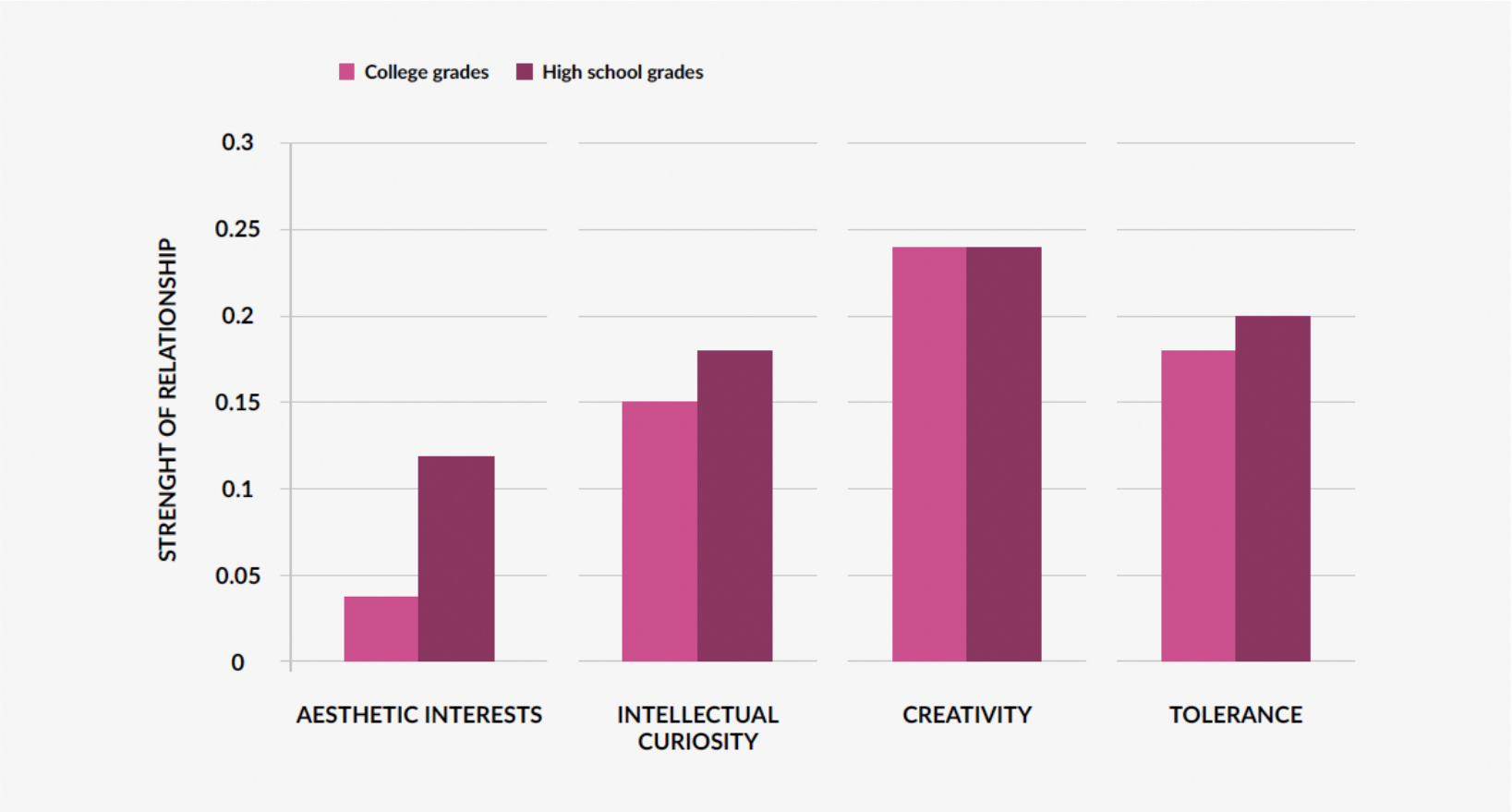


Figure 10: The relationship between Years of Schooling and the Big Five Dimensions

Source: Chernyshenko et al. (2018)

Creativity



Note: Strength of relationships represents average correlation across studies. *Source*: Dudley et al. (2006).

Figure 11: Relationship between Openness-Related Skills and College and High-School Grades

Perseverance

Variable	With control variables		Without control variables	
	β	*SE*	β	*SE*
Dependent Variable: Math Achievement				
NEG	-.24*	0.00	-.26*	0.00
POS	-.04*	0.00	-.05*	0.00
SES	.23*	0.00		
Gender[1]	.05*	0.00		
Dependent Variable: Truancy				
NEG	.09*	0.00	.09*	0.00
POS	-.07*	0.00	-.07*	0.00
SES	-.01*	0.00		
Gender[1]	.05*	0.00		

Note. *SE* = standard error of the standardized regression coefficient; NEG = the Negative facet of the Perseverance scale; POS = the Positive facet of the Perseverance scale.

* $p < .01$.

[1] 0 = woman, 1 = man.

Figure 12: Linear Regressions of Perseverance Facets

Source: Zhang et al. (2022)

THE UNIVERSITY OF CHICAGO

Regio	*N*	β (*SE*)					
		NEG_with	POS_with	Gender[1]	SES	NEG_without	POS_without
NA	31,296	-.26* (.01)	.08* (.01)	.04* (.01)	.21* (.01)	-.28* (.01)	.10* (.01)
ME	26,243	-.37* (.01)	-.01* (.01)	.02* (.01)	.19* (.01)	-.38* (.01)	-.01 (.01)
LA	55,260	-.25* (.00)	.02* (.00)	.10* (.00)	.23* (.00)	-.26* (.00)	.03* (.01)
SE	22,034	-.19* (.01)	.02* (.01)	.05* (.01)	.14* (.01)	-.20* (.01)	.02 (.01)
WE	72,840	-.27* (.01)	.00* (.01)	.06* (.00)	.20* (00)	-.29* (.00)	.01 (.01)
FCC	35,929	-.15* (.01)	-.01* (.01)	.02* (.01)	.19* (.01)	-.16* (.01)	.02 (.01)
NOR	17,775	-.33* (.01)	.07* (.01)	-.02 (.01)	.17* (.01)	-.34* (01)	.08* (.01)
EA	21,086	-.20* (.01)	.12* (.01)	.02* (.01)	.19* (.01)	-.21* (.01)	.12* (.01)
SA	17,985	-.28* (.01)	-.02* (.01)	-.01 (.05)	.27* (.01)	-.29* (.01)	-.01 (.01)

Note. *N* = sample size; *SE* = standard error of the standardized regression coefficient; NEG_with = the negative perseverance facet controlling for gender and SES; POS_with = the positive perseverance facet controlling for gender and SES; NEG_without = the negative perseverance facet without controlling for gender and SES; POS_without = the positive perseverance facet without controlling for gender and SES; NA = North America/Oceania; ME = The Middle East; LA = Latin America; SE = Southern Europe; WE = Western Europe; FCC = Former Communist Countries; NOR = The Nordic countries; EA = East Asia; SA = Southeast Asia.

* $p < .01$.

[1] 0 = woman, 1 = man.

Figure 13: Results of Linear Regressions with Partial Measurement Invariance Across 9 Cultural Regions (Dependent Variable: Math Achievement)

Source: Zhang et al. (2022)

Region	*N*	β (*SE*)					
		NEG_with	POS_with	Gender[1]	SES	NEG_without	POS_without
NA	31,296	.12* (.01)	-.12* (.01)	.00 (.01)	-.06* (.01)	.13* (.01)	-.13* (.01)
ME	26,243	.10* (.01)	-.12* (.01)	.08* (.01)	-.05* (.01)	.11* (.01)	-.12* (.01)
LA	55,260	.10* (.01)	-.12* (.01)	.03* (.00)	.03* (.00)	.09* (.01)	-.12* (.01)
SE	22,034	.12* (.01)	-.11* (.01)	.07* (.01)	.02* (.01)	.12* (.01)	-.11* (.01)
WE	72,840	.07* (.01)	-.07* (.01)	.04* (.00)	-.04* (.00)	.07* (.01)	-.07* (.01)
FCC	35,929	.03* (.01)	-.06* (.01)	.07* (.01)	.00 (.01)	.03* (.01)	-.06* (.01)
NOR	17,775	.14* (.01)	-.09* (.01)	.04* (.01)	-.04* (.01)	.14* (.01)	-.10* (.01)
EA	21,086	.04* (.01)	-.06* (.01)	.06* (.01)	-.05* (.01)	.04* (.01)	-.06* (.01)
SA	17,985	.17* (.01)	-.05* (.01)	.12 (.01)	.02* (.01)	.16* (.01)	-.05* (.01)

Figure 14: Results of Linear Regressions with Partial Measurement Invariance Across 9 Cultural Regions (Dependent Variable: Truancy)

Source: Zhang et al. (2022)

Grit

	Overall Grit							Perseverance of Effort						Consistency of Interest					
Correlate/Criterion	k	N	r_{obs}	ρ	SD_ρ	10% CV	90% CV	K	N	r_{obs}	ρ	SD_ρ	10% CV	90% CV	k	N	r_{obs}	ρ	SD_ρ	10% CV	90% CV
Academic Performance	39	13,141	.15	.18	0.11	.04	.31	11	5,221	.20	.26	0.12	.11	.41	11	5,221	.08	.10	0.02	.07	.13
GPA (All Levels)	37	12,601	.14	.17	0.10	.04	.30	10	5,065	.20	.26	0.12	.11	.41	10	5,065	.08	.10	0.03	.06	.13
Undergraduate GPA	30	10,526	.14	.17	0.10	.04	.30	8	4,595	.20	.25	0.12	.11	.40	8	4,595	.08	.09	0.03	.05	.14
High School GPA	17	6,364	.13	.16	0.14	-.02	.34	7	3,313	.22	.29	0.08	.19	.39	7	3,313	.11	.13	0.06	.05	.21
Graduate School GPA	3	1,141	.07	.08	0.00	.08	.08														
Retention (General)	5	2,705	.14	.16	0.06	.08	.23														
Retention (General)*	11	17,525	.10	.12	0.09	.00	.24														
Retention (w/o Marriage)*	10	11,163	.16	.18	0.03	.14	.23														
Non-Academic Performance	7	4,116	.19	.21	0.00	.21	.21														
Intent to Persist in College	5	3,967	.13	.18	0.00	.18	.18	4	2,959	.15	.22	0.00	.22	.22	4	2,959	.09	.12	0.05	.06	.19
Intent to Persist with Current Employer	4	519	.13	.15	0.00	.15	.15														
Grit – Perseverance															17	22,048	.44	.60	0.21	.34	.87

Note. k = number of studies, N =number of subjects, r_{obs} = sample size weighted mean observed correlation, ρ =true score correlation, SD_ρ = standard deviation of true score correlation, 10%CV and 90%CV = lower and upper bound of 80% credibility intervals. * denotes that these estimates include estimates of point-biserial correlations computed by taking root of Nagelkerke R^2 values. Confidence intervals are not shown here due to space limitations but can be obtained from first author by request.

Figure 15: Meta-Analytic Estimates of the Relations between Grit and Performance Criteria

Source: Crede et al. (2017)

THE UNIVERSITY OF CHICAGO

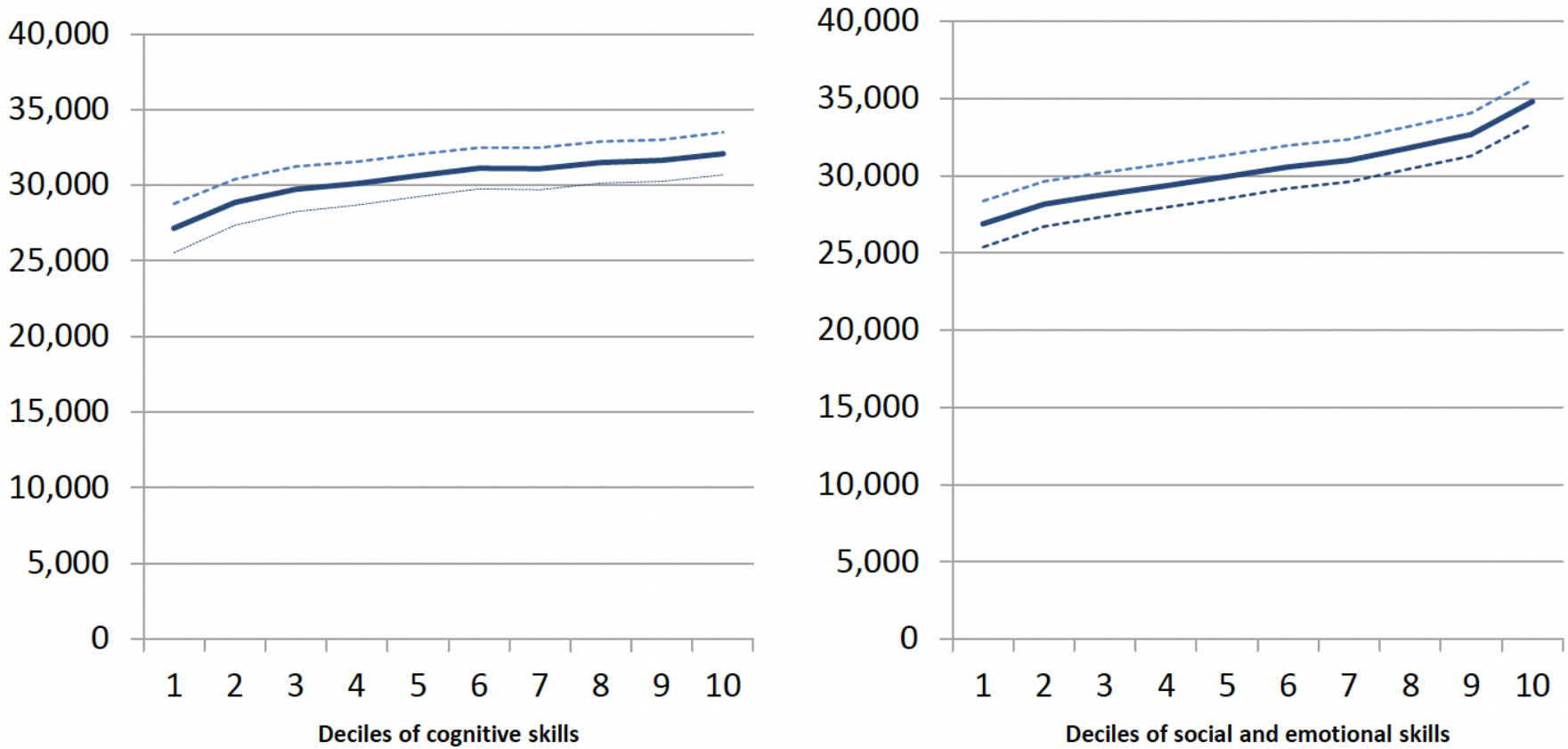


Figure 16: Income as a Function of Cognitive or Social and Emotional Skills

Source: Chernyshenko et al. (2018)

The Person-Situation Debate

- *"...with the possible exception of intelligence, highly generalized behavioral consistencies have not been demonstrated, and the concept of personality traits as broad dispositions is thus untenable"*
  –Mischel (1968)
- *"Manipulations of the immediate social situation can overwhelm in importance the type of individual differences in personal traits or dispositions that people normally think of as being determinative of social behavior."*
  –Ross and Nisbet (1991)
- *"The great contribution to psychology by Walter Mischel [...] is to show that there is no such thing as a stable personality trait"*
  –Thaler (2008)

Executive Function

- Executive function (EF) skills are a set of neurocognitive skills that support the conscious, top-down attentional control of thought, action, and emotion.
- These skills vary along a continuum of motivational significance from "hot EF" to "cool EF," as demonstrated by lesion studies, neuroimaging studies, and research using transcranial direct stimulation (tCDS).
- Together, cool and hot EF skills make it possible to sustain attentional focus, keep goals and information in mind, refrain from responding impulsively, resist distraction, tolerate frustration, consider the consequences of different behaviors, reflect on past experiences, and plan for the future.
- EF skills are essential for goal-directed problem solving, flexible adaptation to changing circumstances, effective social functioning, and intentional learning.

- Methodological advances in the assessment of cool EF skills have resulted in standardized direct behavioral assessments that can be administered repeatedly to children as young as 2 years and across the life span.
- Examples include the National Institutes of Health (NIH) Toolbox measures of EF and the Minnesota Executive Function Scale (MEFS).
- These computer adaptive, tablet-based measures indicate that EF skills, which emerge in infancy (e.g., in the context of search for hidden objects), develop rapidly during early childhood and the transition to adolescence, and continue to improve into early adulthood (see Figure 17).

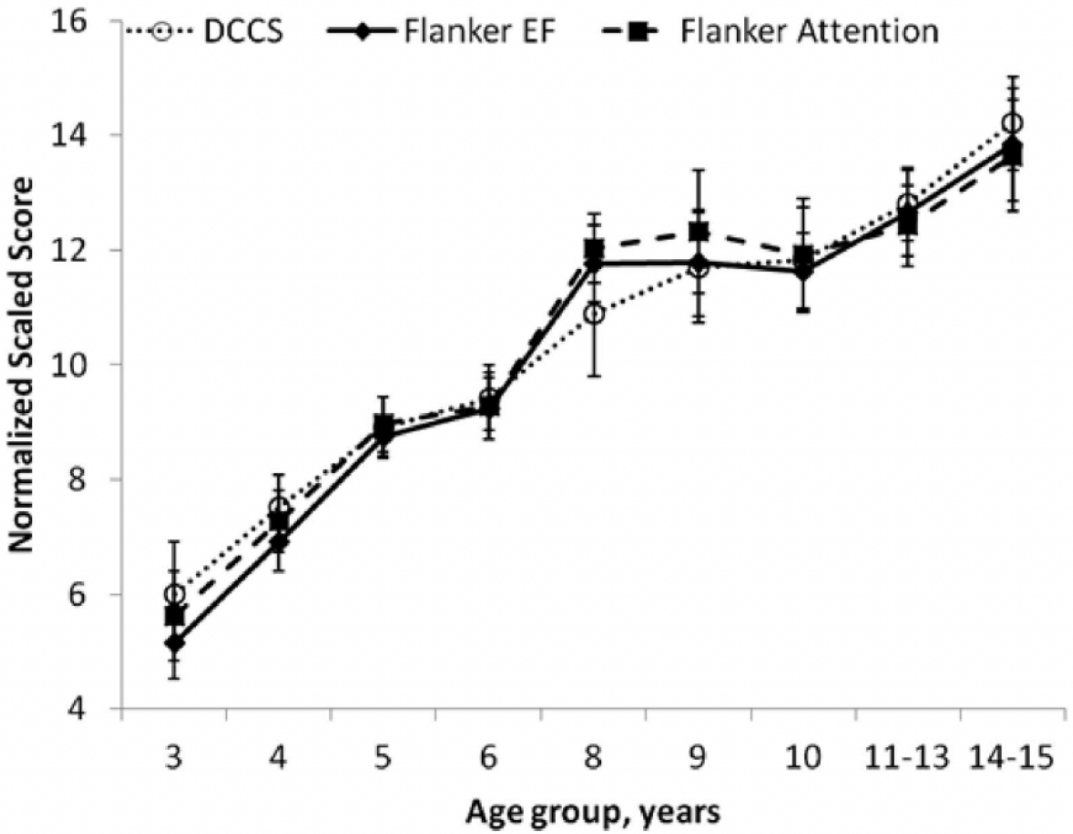


Figure 17: Performance on the NIH Toolbox Dimensional Change Card Sort Test and the Toolbox Flanker Inhibitory Control and Attention Test across age groups

Source: Zelazo & Carlson (2020)

- Although EF skills remain sensitive to the influence of experience across the life span, there may be periods of relatively high plasticity, during which experiential influences are particularly strong.
- As shown in Figure 17, EF skills appear to develop especially rapidly during early childhood and the transition to adolescence, indicating that underlying neural networks are adapting to correspondingly salient environmental challenges.
- Sociocultural practices such as the transitions to formal schooling, middle school, and high school, place new demands on children's EF skills, thereby growing these skills.
- In turn, growing EF skills increase children's readiness to learn (i.e., their plasticity); thus, setting up the conditions for major developmental transitions associated with these stages of life.

Bias

- Cognitive biases refer to the immediate reactions of subjects to stimuli. Often this involves subjects providing a wrong first answer to a question rather than the correct reasoned answer.
- These biases exert effects on human cognition and behavior, are largely ubiquitous, and are quite resistant to attempts to mitigate or eliminate them.

Biases can take many forms. Gertner et al. (2016) identify 6 main types of bias:

- **Confirmation bias**. The tendency to search for or interpret information in a way that confirms one's preconceptions.
- **Fundamental attribution error**. The tendency for people to overemphasize personality based explanations for behaviors observed in others while underemphasizing the role and power of situational influences on the same behavior (also called attribution bias).
- **Bias blind spot**. The tendency for an individual to be unaware of their own cognitive biases, even when the individual can recognize cognitive biases in others.

- **Anchoring bias**. The tendency to rely too heavily or overly restrict one's attention to one trait or piece of information when making judgments.
- **Representativeness bias**. The tendency for people to judge the probability or frequency of a hypothesis by considering how much the hypothesis resembles available data.
- **Projection bias**. The tendency to unconsciously assume that others share one's current emotional states, thoughts, and values.

Implicit Association Test

- The Implicit Association Test (IAT) is a measure of associative knowledge that has been used to measure implicit bias.
- The IAT measures differential association of 2 target concepts with an attribute. The 2 concepts appear in a 2-choice task (e.g., flower vs. insect names), and the attribute in a 2nd task (e.g., pleasant vs. unpleasant words for an evaluation attribute).
- When instructions oblige highly associated categories (e.g., flower + pleasant) to share a response key, performance is faster than when less associated categories (e.g., insect + pleasant) share a key.
- This performance difference implicitly measures differential association of the 2 concepts with the attribute.
- It has been suggested that the IAT may be used to measure individual differences in bias.

Implicit vs Explicit Measurements

- **Implicit measures** provide estimates of individuals' attitudes without researchers directly asking them for this information.
- Participants may be unaware that their attitudes are being assessed, but they may not necessarily be unaware that they possess those attitudes.
- **Explicit measures** ask subjects to self-report their attitudes and feelings.

Predictive Power of IAT (from meta-analyses)

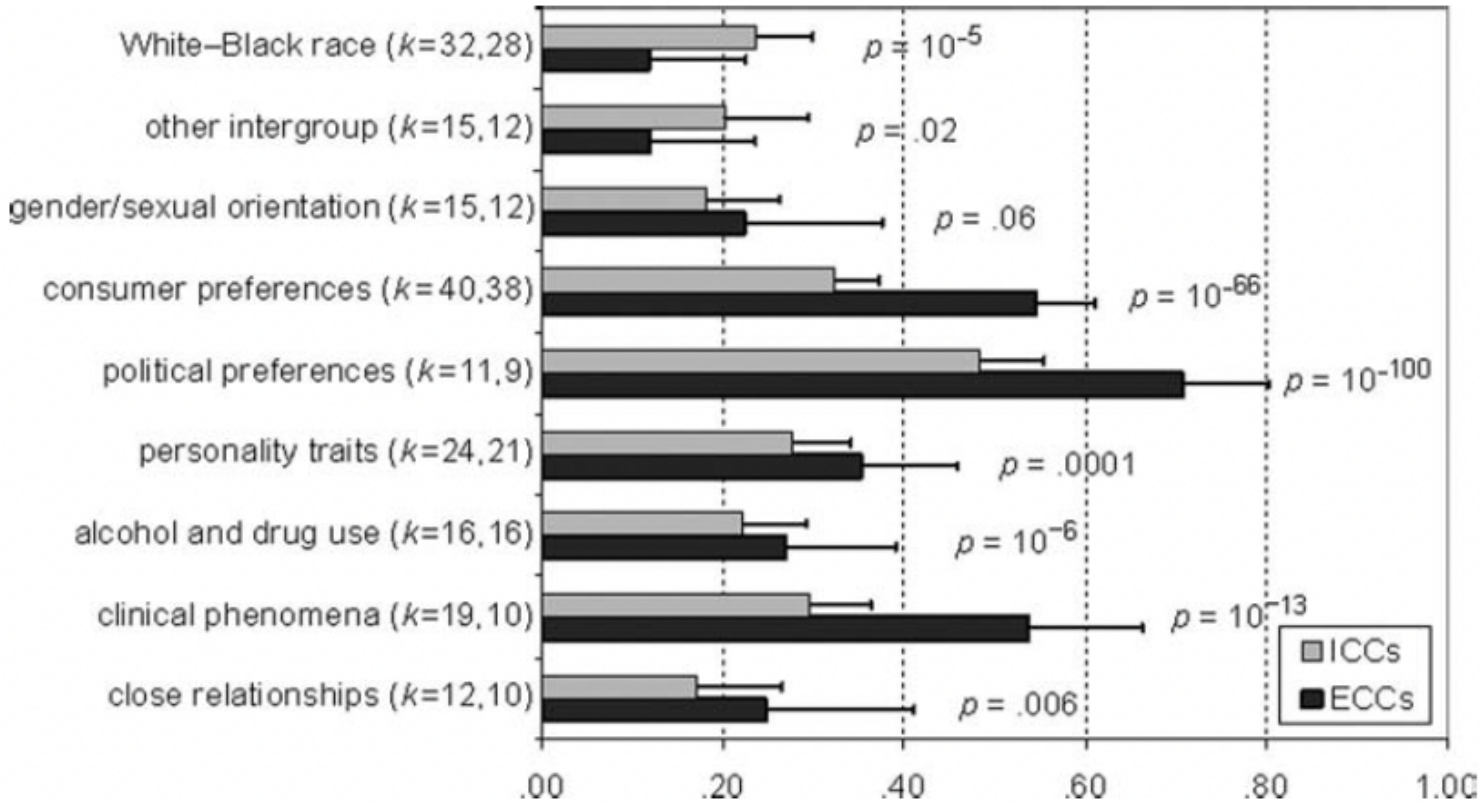


Figure 18: Effect sizes for IAT–criterion (ICC) and explicit– criterion (ECC) correlations

Source: Greenwald et al. (2009)

THE UNIVERSITY OF CHICAGO

Implicit Association Test (Cont'd)

- Effect sizes for both ICCs and ECCs vary widely across domains.
- This across-domain variation in effect sizes is much greater for ECCs than for ICCs (i.e., lengths of the black bars in Figure 18 vary much more than do those of the gray bars)
- This greater heterogeneity of ECCs than ICCs can also be seen in the wider 95% confidence intervals for black than gray bars in Figure 18.
- Although average ECCs are significantly greater than ICCs in six criterion domains, the reverse is true for the two domains that involved intergroup behavior (the top two pairs of bars in Figure 18).

Economic Preferences

- In economics, choices are determined by constraints, prices, information, and preferences.
- If constraints, prices, and information are the same, differences in behavior are attributed to preferences.
- They are usually measured by presenting individuals with choices and observing decisions in different situations.

Economic Preferences (Cont'd)

- **Risk Aversion**: Preference for certainty over uncertainty.
- **Ambiguity Aversion**: Preference for "known" uncertainty, over unknown uncertainty.
- **Loss Aversion**: Higher sensitivity to losses when compared to gains of the same scale.
- **Time Preference**: Preference over consumption in different time periods.
- **Altruism**: Unconditional kindness.
- **Trust**: Willingness to make oneself vulnerable to opportunistic individuals.

Economic Preferences (Cont'd)

- **Positive Reciprocity**: Tendency to reward kind actions.
- **Negative Reciprocity**: Tendency to punish others for unkind actions.
- **Cooperation**: Preferences for working with others toward mutual gain.
- **Complexity**: Ability to scope out and then break down solutions to problems.
- **Leisure**: Preference over leisure and consumption.

Risk acceptance (gains)							
0.47***	Risk acceptance (losses)						
0.07**	0.01	Delay acceptance					
0.10***	–0.14***	0.22***	Cognitive skill				
–0.01	–0.02	0.00	–0.04	C			
0.01	–0.03	–0.01	0.00	0.27***	E		
–0.04	0.01	0.07**	–0.03	0.39***	0.05*	A	
–0.05*	–0.02	0.02	–0.06*	–0.40***	–0.32***	–0.36***	N

Figure 19: Pairwise Correlations between Risk Acceptance, Delay Acceptance, Cognitive Ability, and Personality

Source: Adapted from Anderson, Burks, DeYoung, and Rustichini (2011)

- The recent literature shows that non-cognitive skills predict standardized achievement test scores, which some psychologists assume are good measures of intelligence.
- Non-cognitive skills explain a substantial portion of the variability across persons in standardized achievement test scores.
- Non-cognitive skills explain the variance in achievement scores above and beyond the variance that IQ explains when both measures of non-cognitive skill and IQ are included in a regression.
- These findings caution the interpretation that standardized achievement tests only measure cognitive ability, they also capture non-cognitive skills.

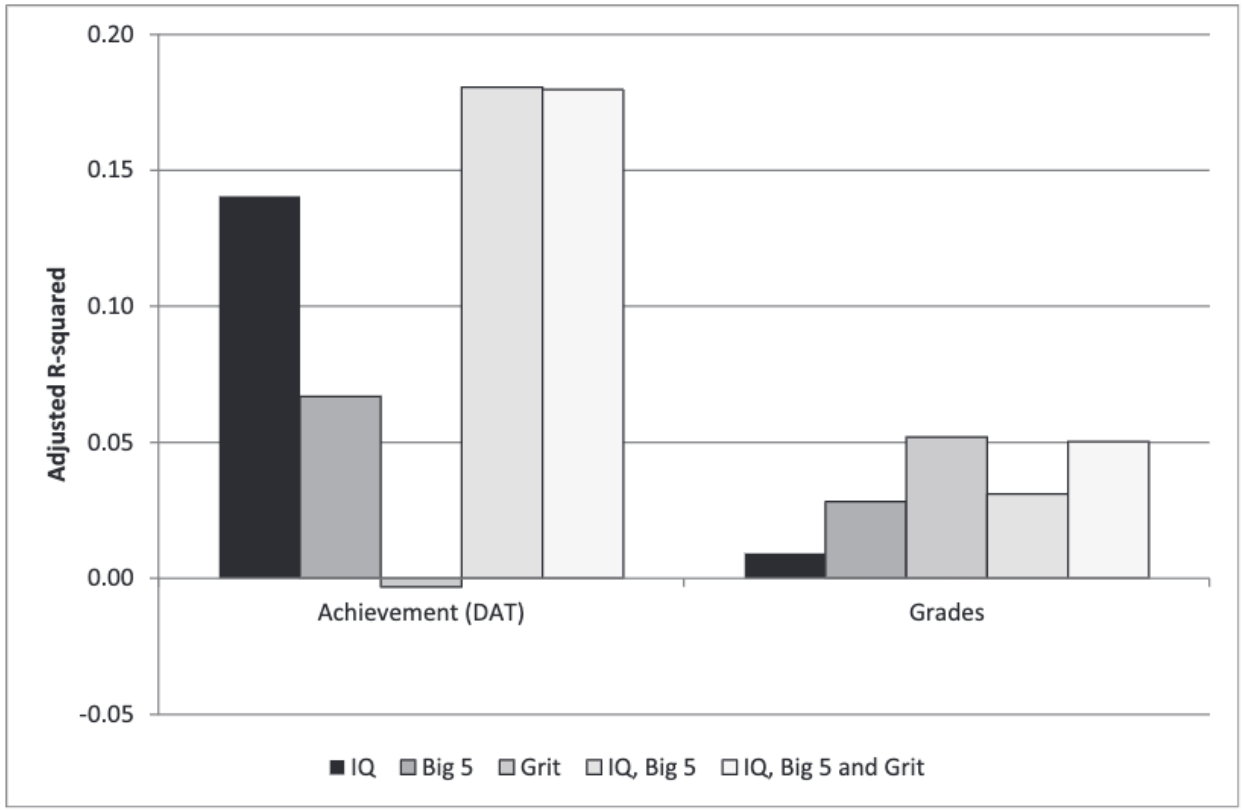


Figure 20: Decomposing achievement tests and grades into IQ and personality (Stella Maris)

Source: Borghans at al. (2016)

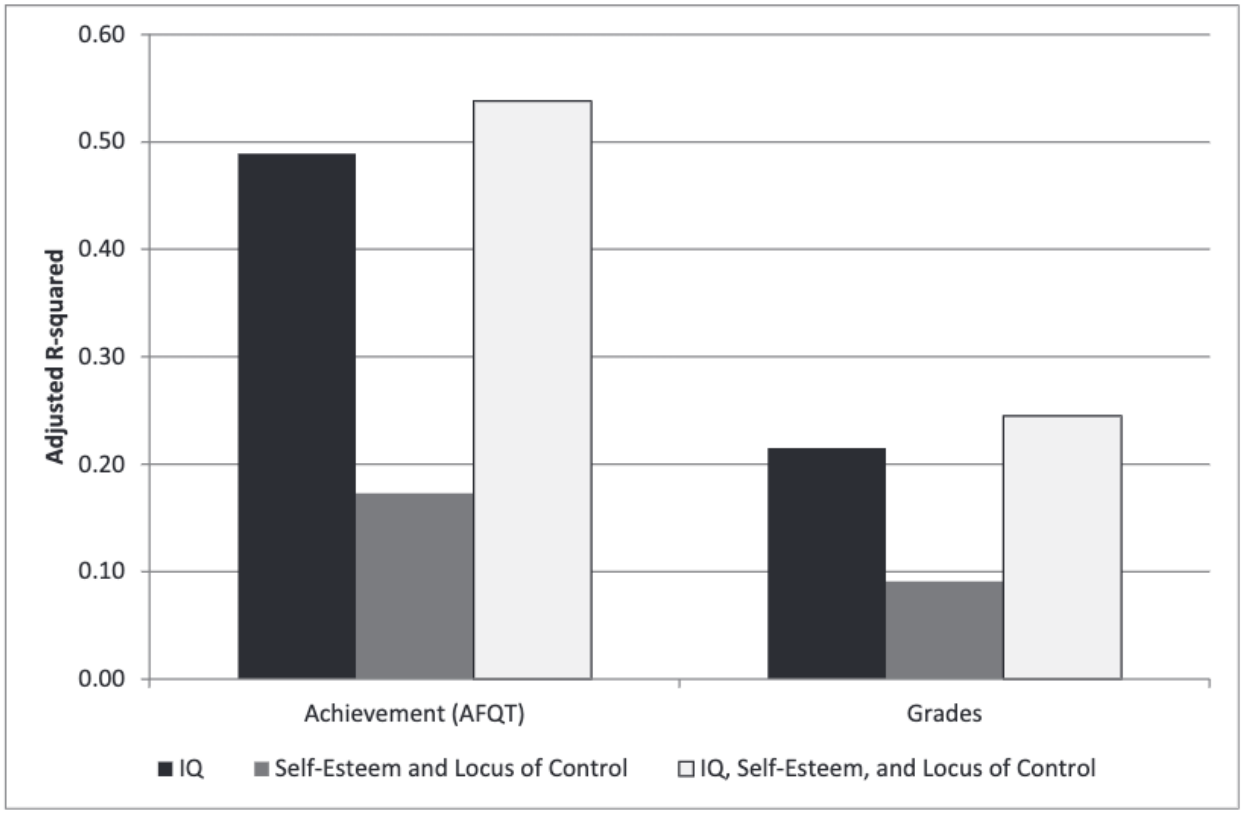


Figure 21: Decomposing achievement tests and grades into IQ and personality (NLSY)

Source: Borghans at al. (2016)

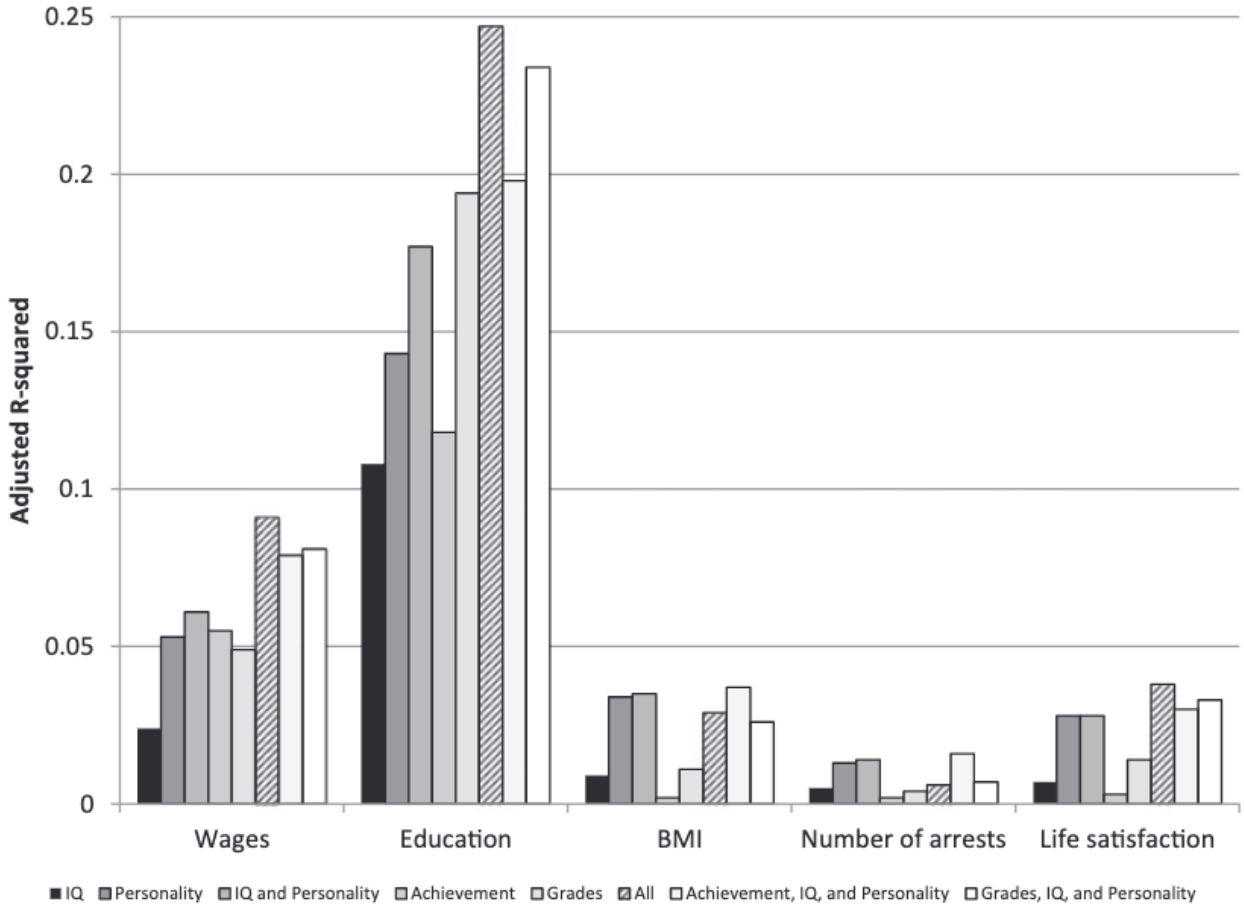


Figure 22: Decomposing life outcomes into IQ and personality (BCS)

Source: Borghans at al. (2016)

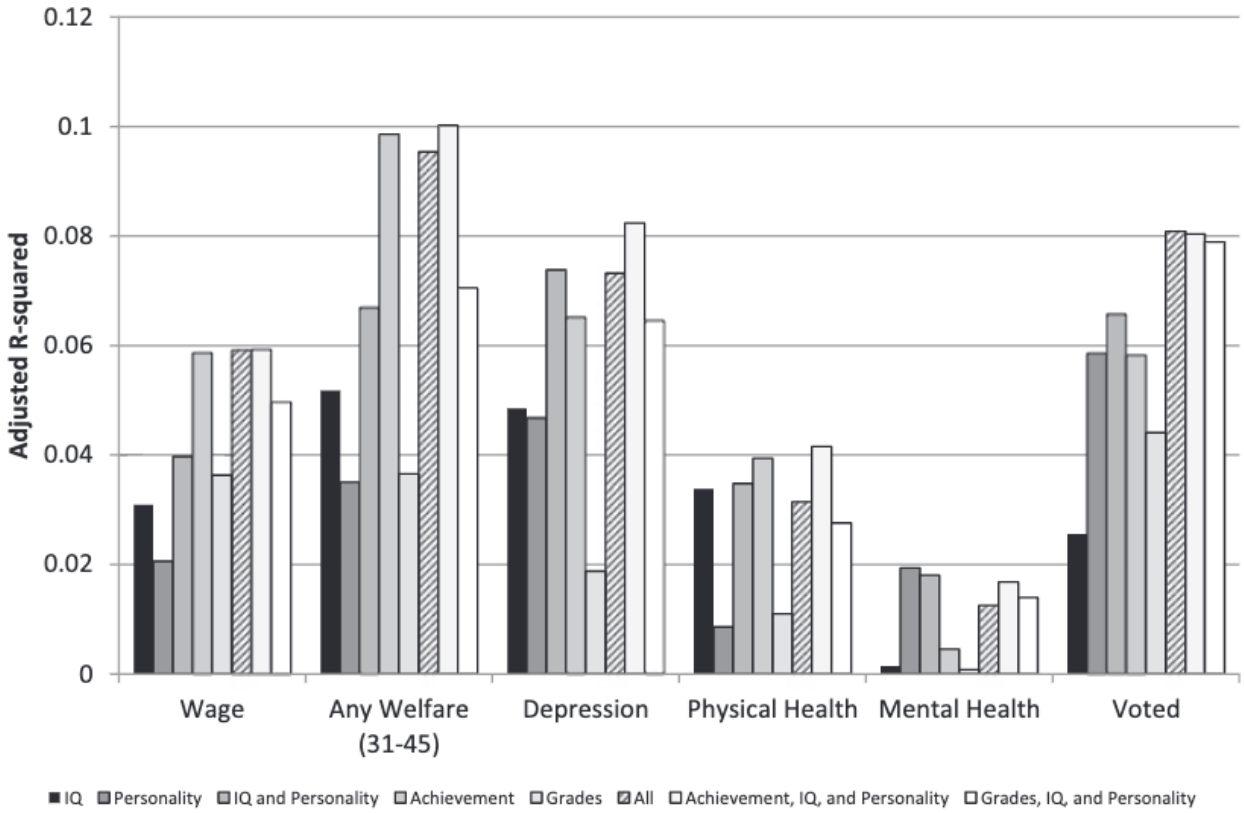


Figure 23: Decomposing life outcomes into IQ and personality (NLSY)

Source: Borghans at al. (2016)

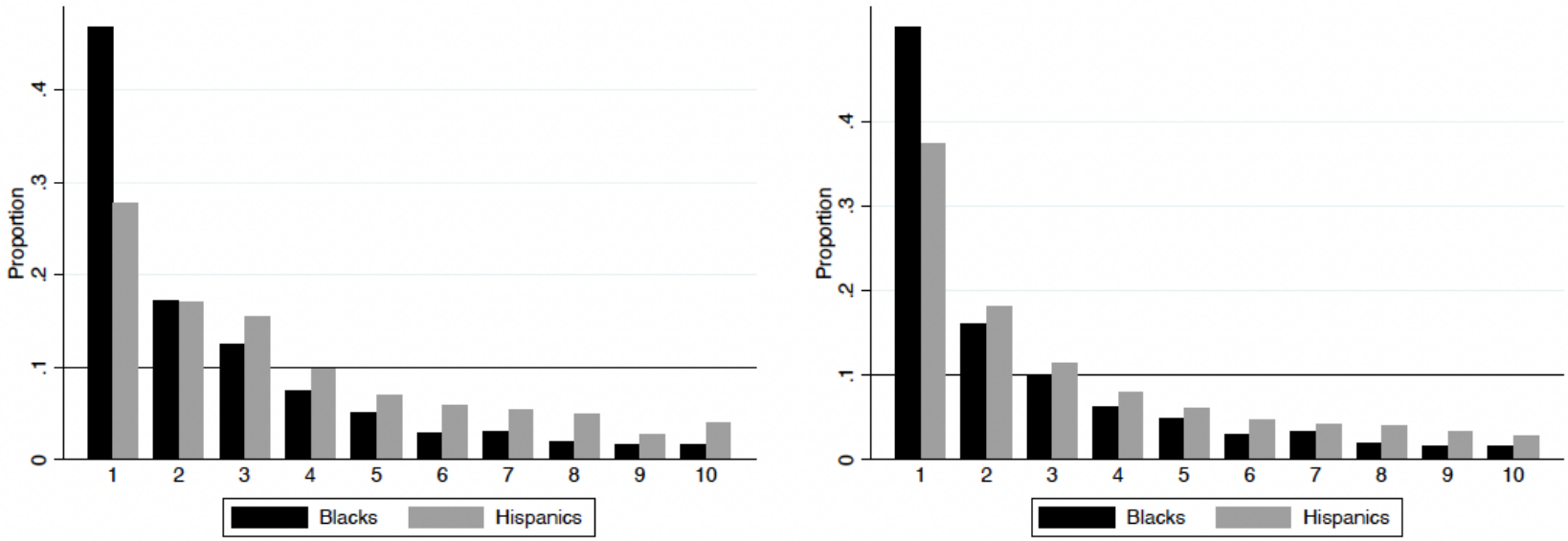


Figure 24: Minority AFQT Scores Placed in the White Distribution - Males (left) and Females (right)

Source: Heckman (2010)

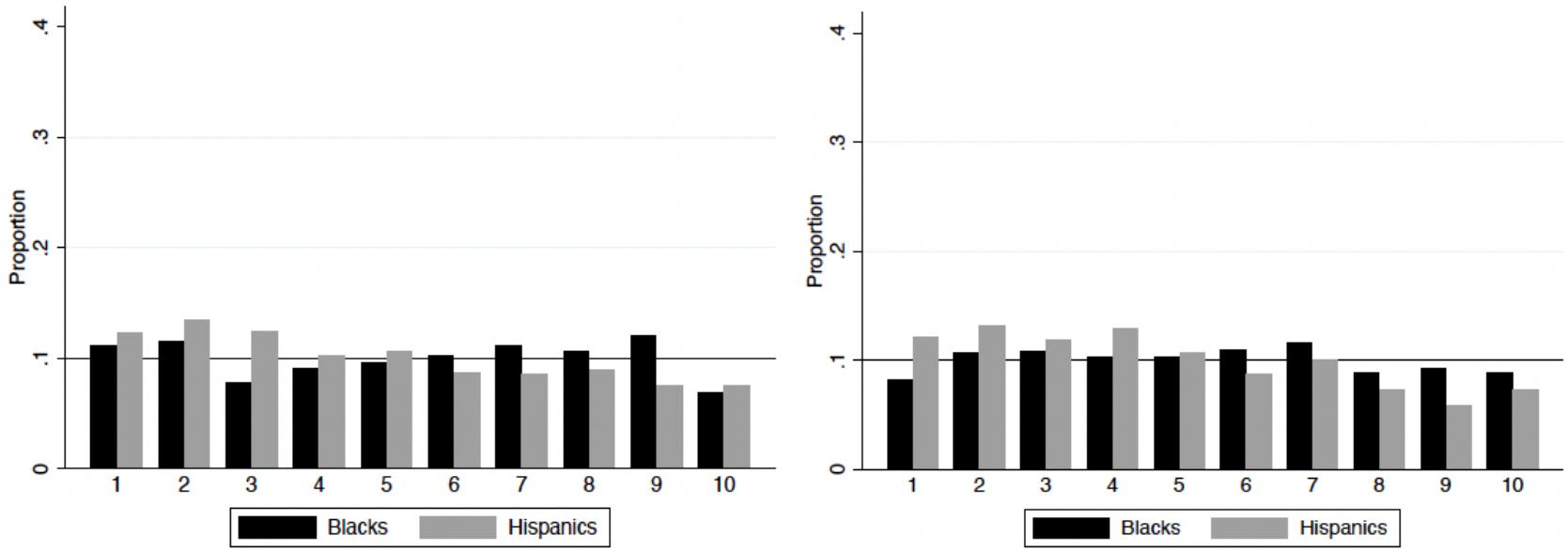


Figure 25: Minority Rosenberg Scores Placed in the White Distribution - Males (left) and Females (right)

Source: Heckman (2010)

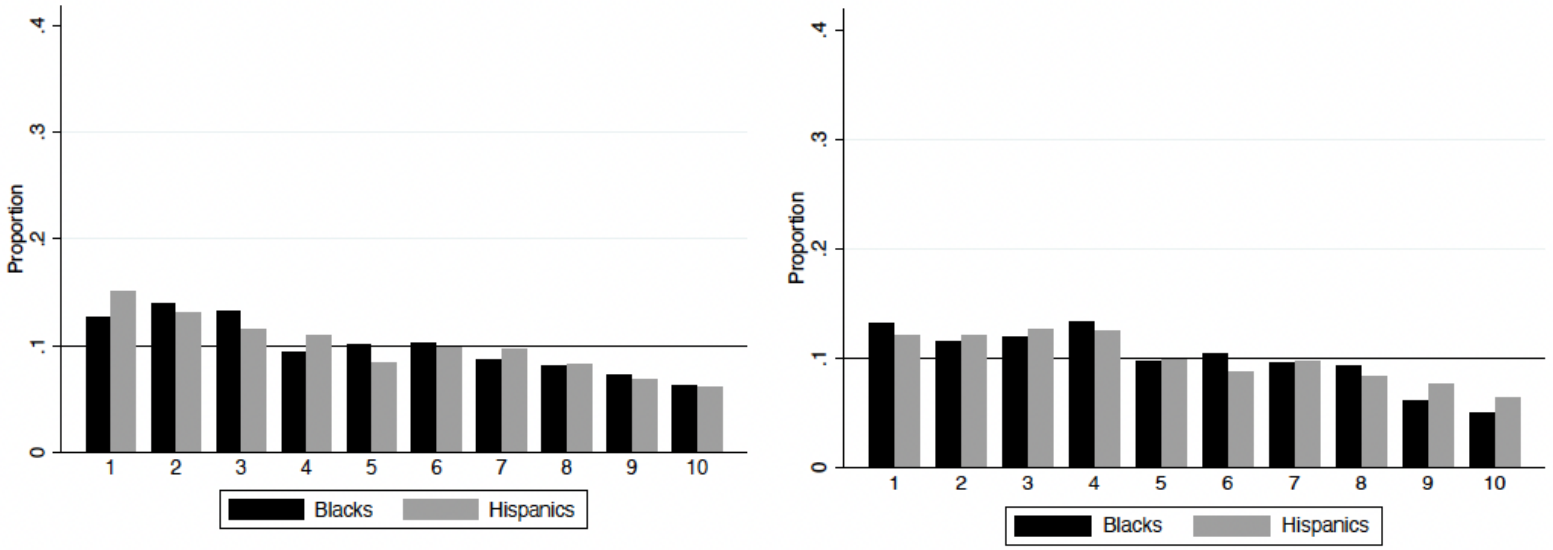


Figure 26: Minority Rotter Scores Placed in the White Distribution - Males (left) and Females (right)

Source: Heckman (2010)

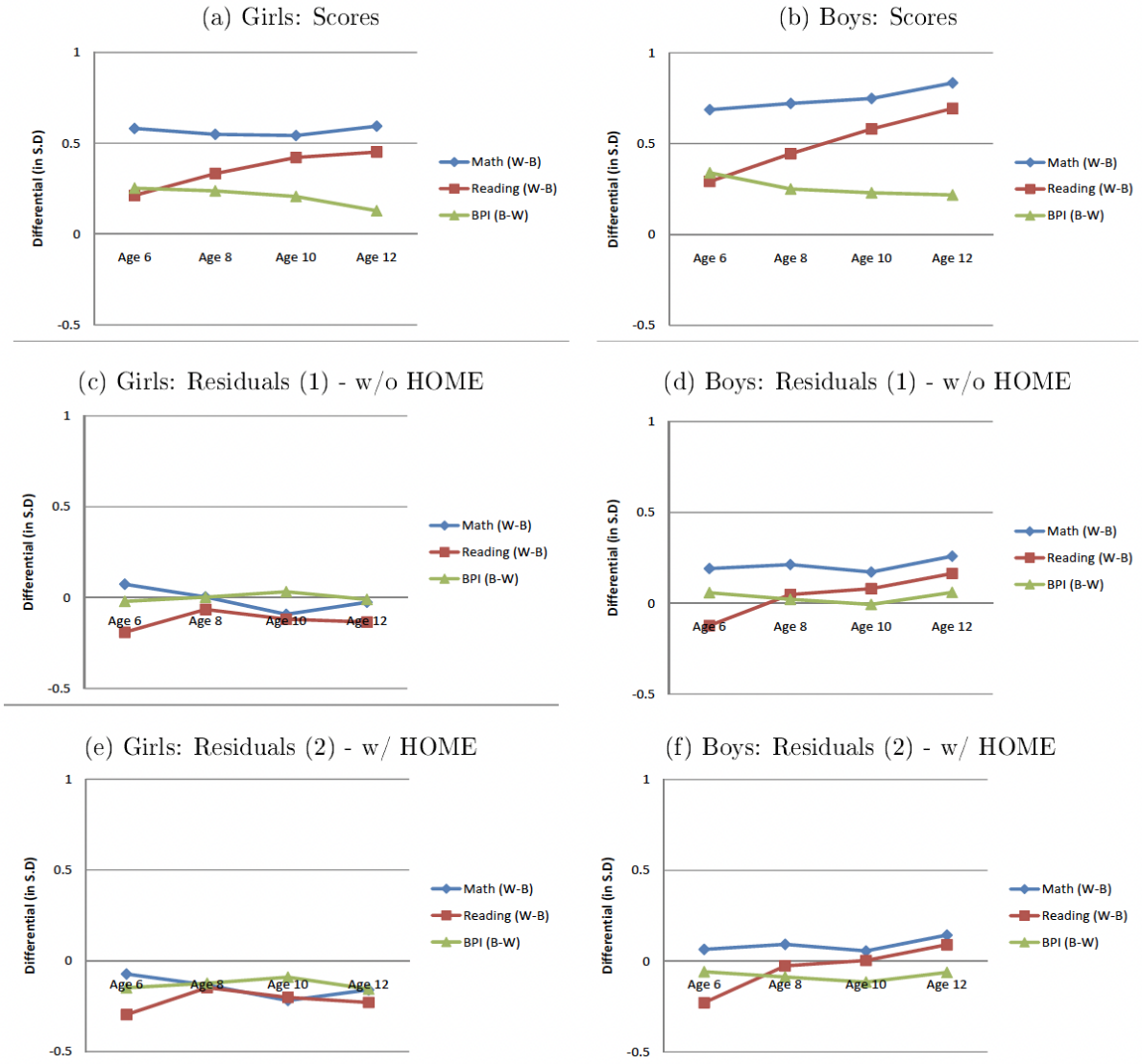


Figure 27: Black-White Gaps in Skill Measures over Ages

Source: Heckman (2010)

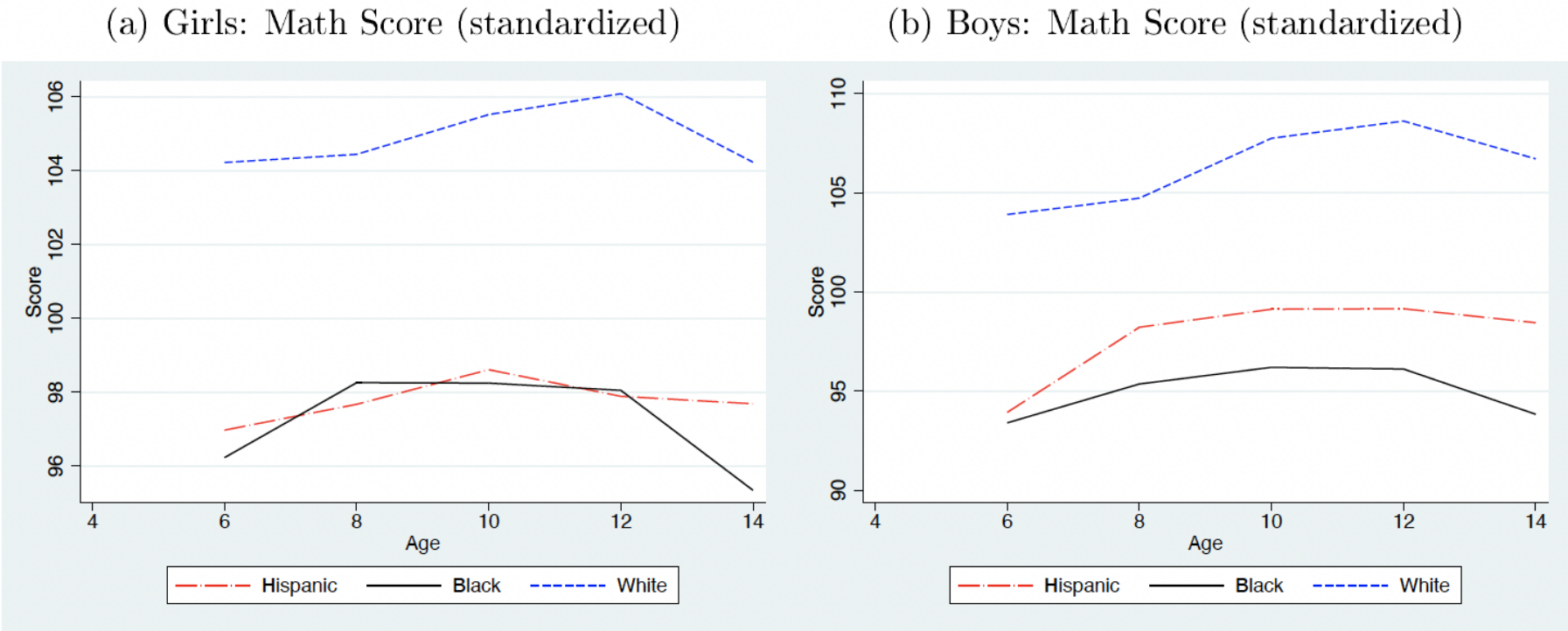


Figure 28: Skill Measures over Childhood across Ethnic Groups

Source: Heckman (2010)

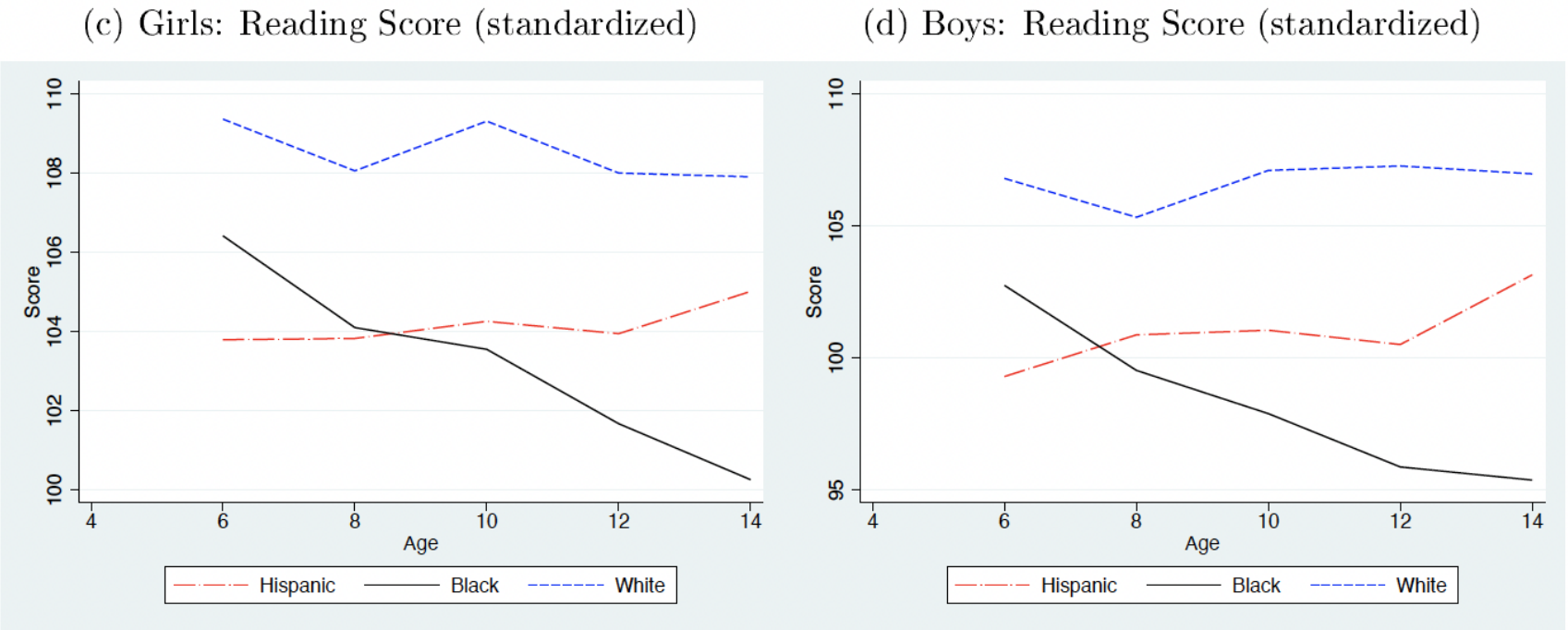


Figure 28: Skill Measures over Childhood across Ethnic Groups (Cont'd)

Source: Heckman (2010)

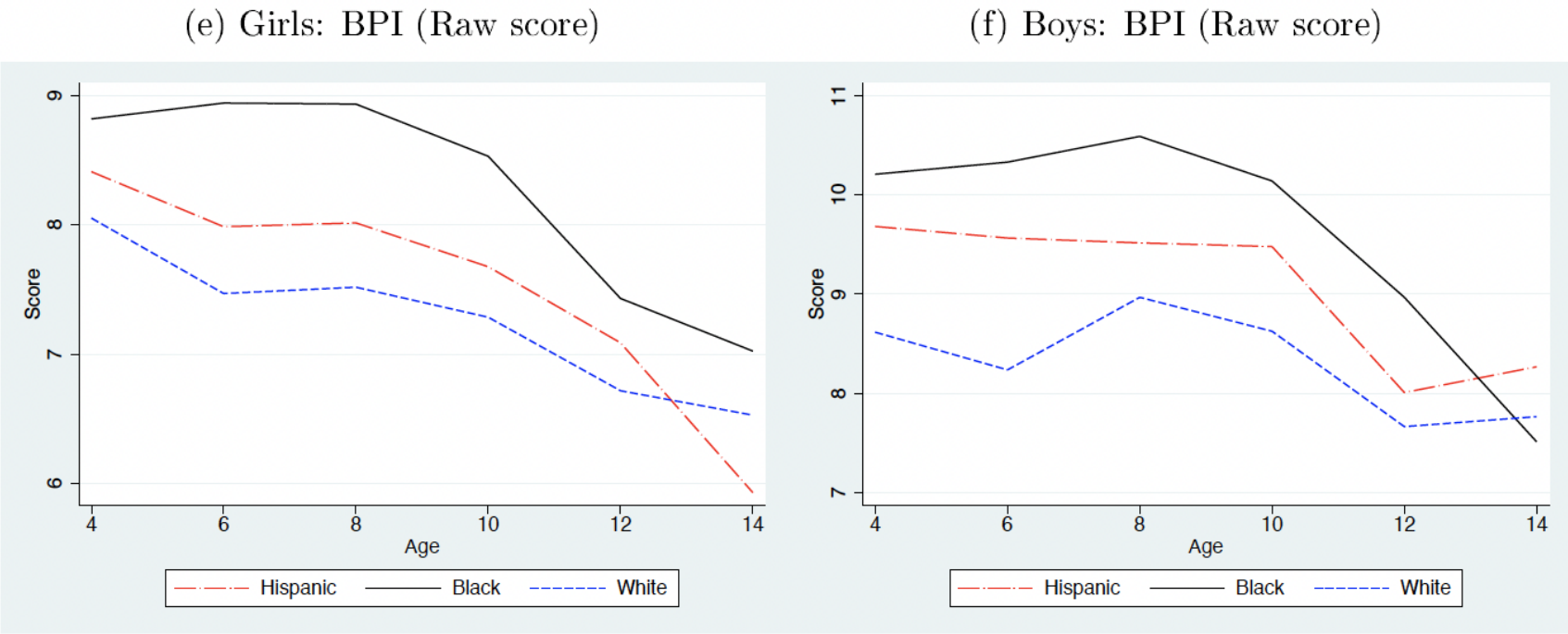
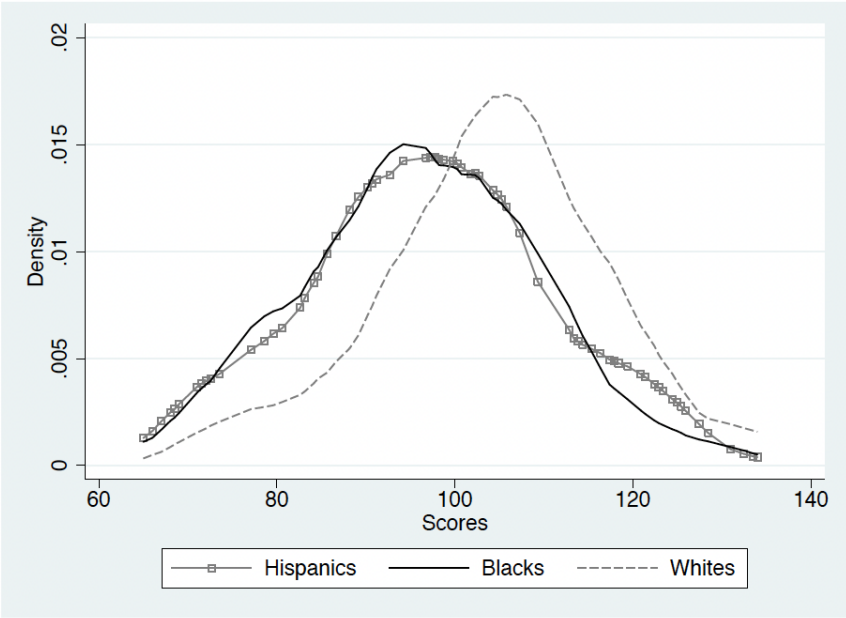


Figure 28: Skill Measures over Childhood across Ethnic Groups (Cont'd)

Source: Heckman (2010)

(a) Girls: Math Score (standardized)



(b) Boys: Math Score (standardized)

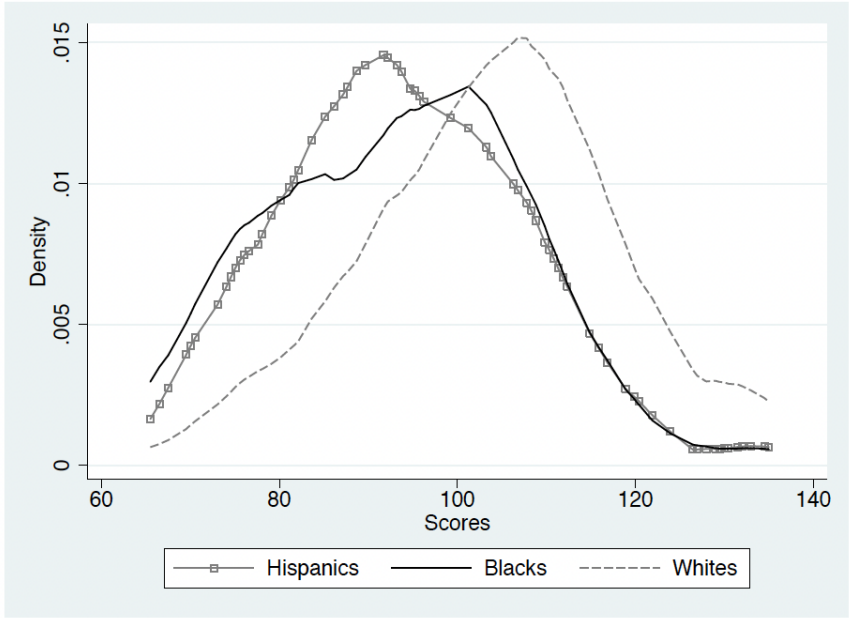
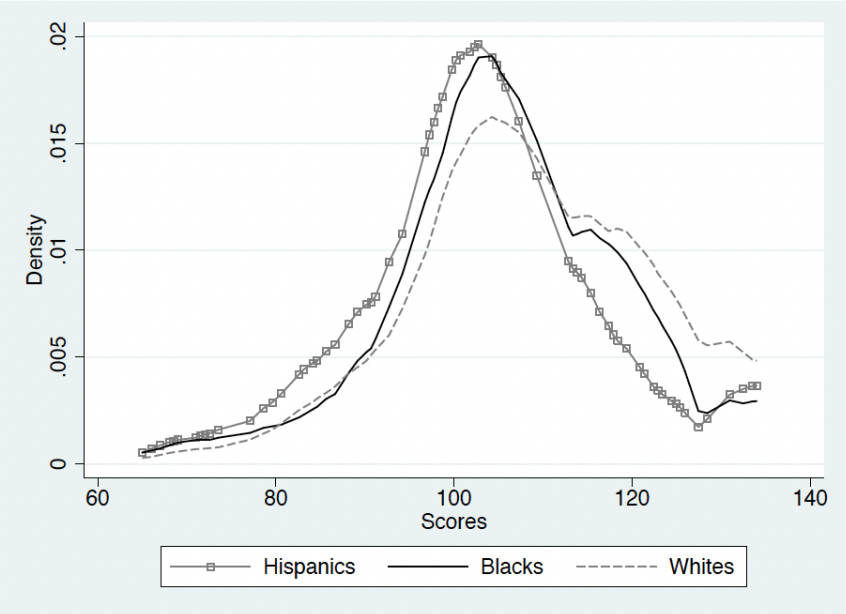


Figure 29: Distribution of Skill Measures across Ethnic Groups: Age 6

Source: Heckman (2010)

(c) Girls: Reading Score (standardized)



(d) Boys: Reading Score (standardized)

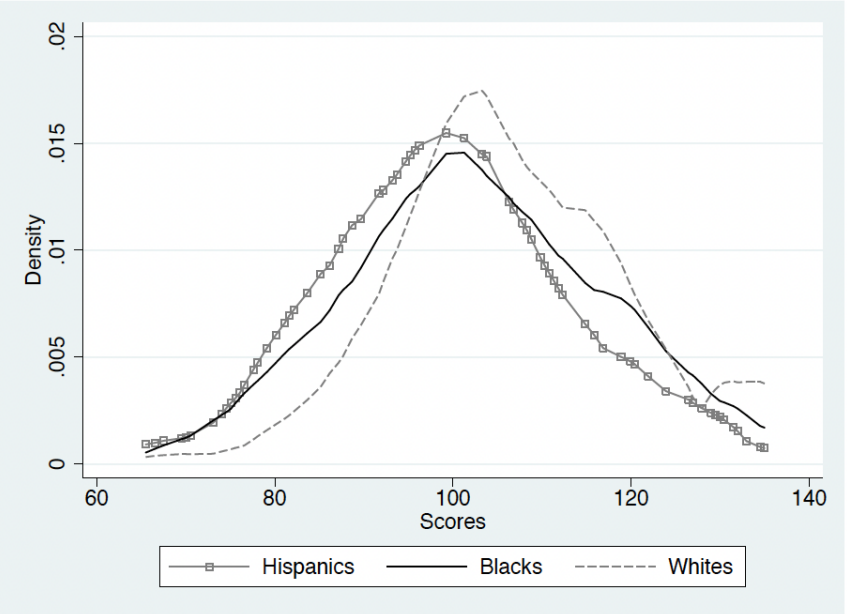
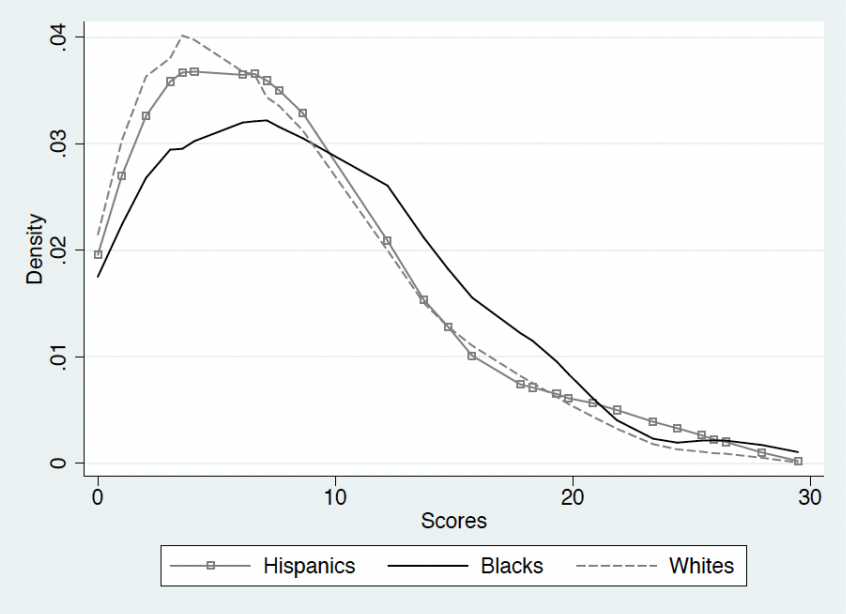


Figure 29: Distribution of Skill Measures across Ethnic Groups: Age 6 (Cont'd)

Source: Heckman (2010)

(e) Girls: BPI (Raw score)



(f) Boys: BPI (Raw score)

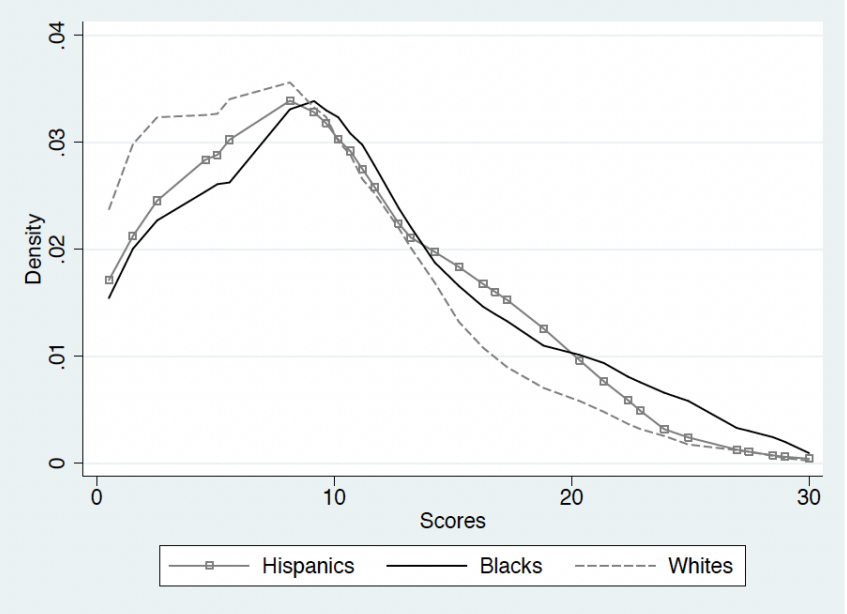


Figure 29: Distribution of Skill Measures across Ethnic Groups: Age 6 (Cont'd)

Source: Heckman (2010)

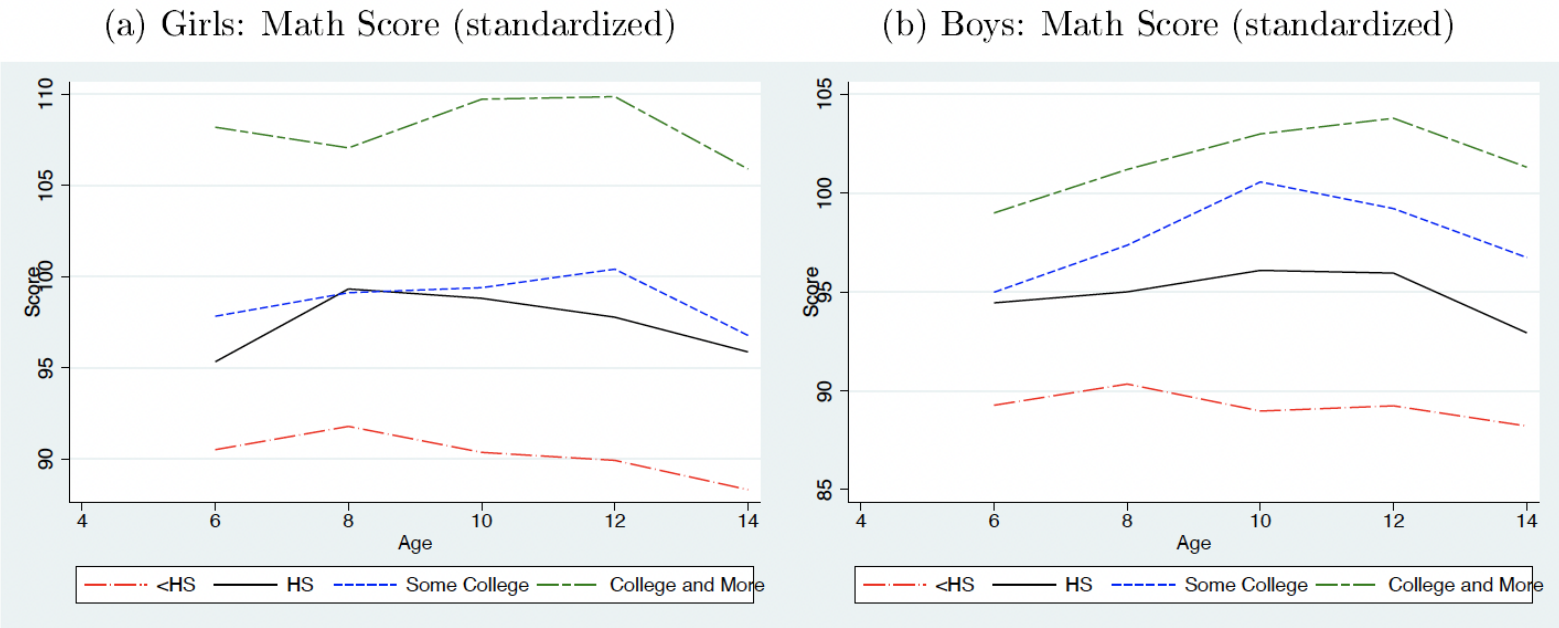


Figure 30: Skill Measures over Childhood by Mother's Education: Black

Source: Heckman (2010)

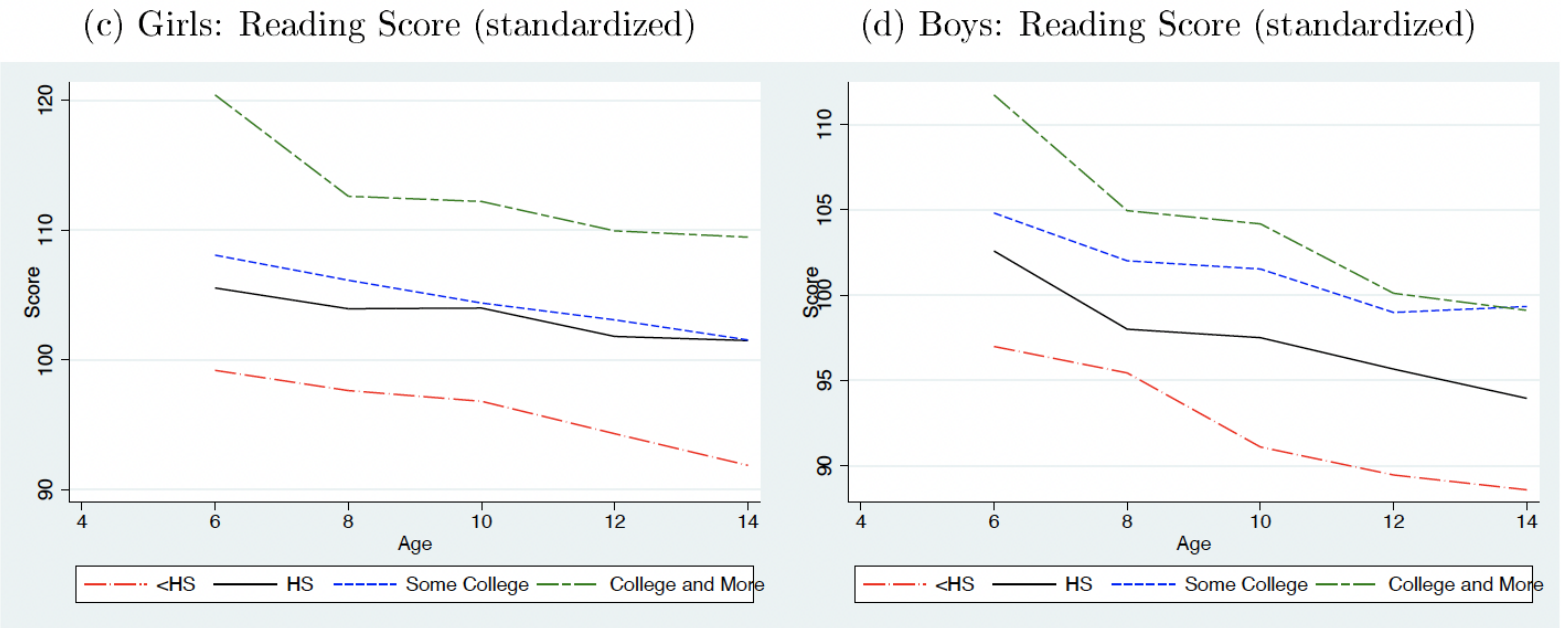


Figure 30: Skill Measures over Childhood by Mother's Education: Black (Cont'd)

Source: Heckman (2010)

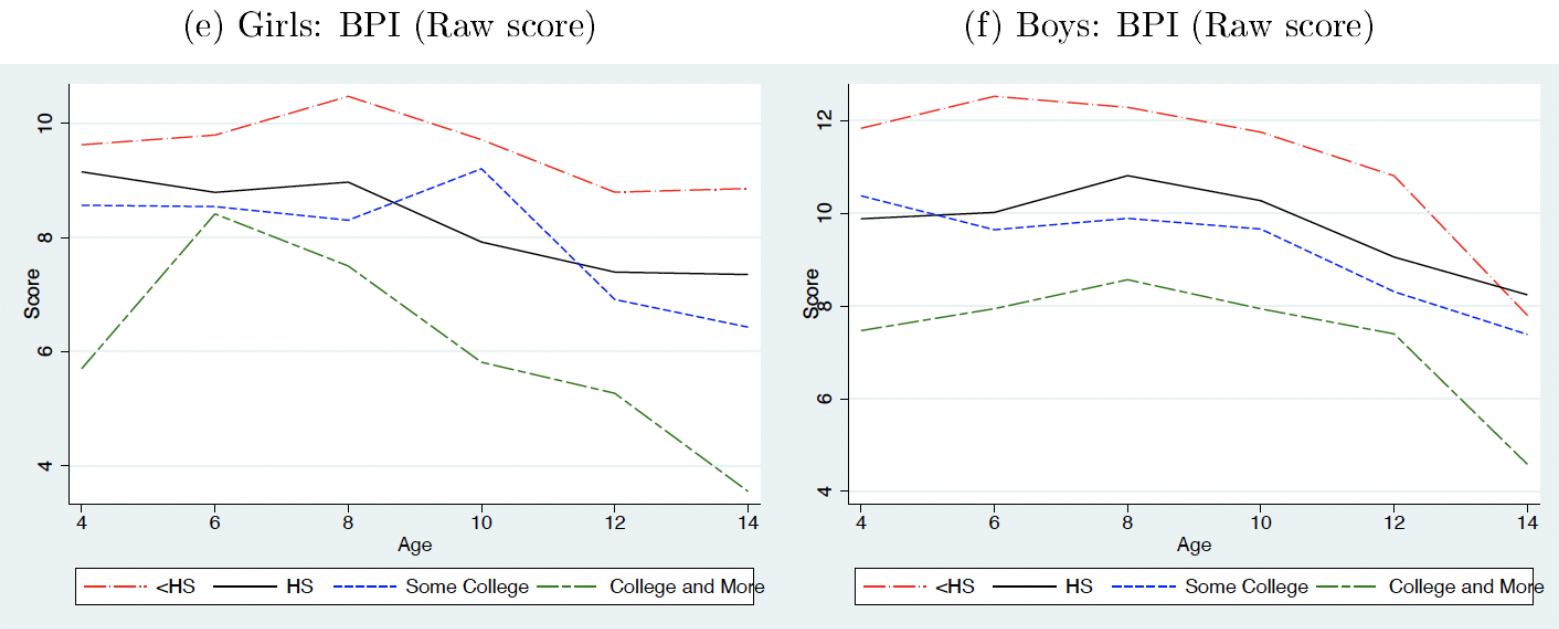


Figure 30: Skill Measures over Childhood by Mother's Education: Black (Cont'd)

Source: Heckman (2010)

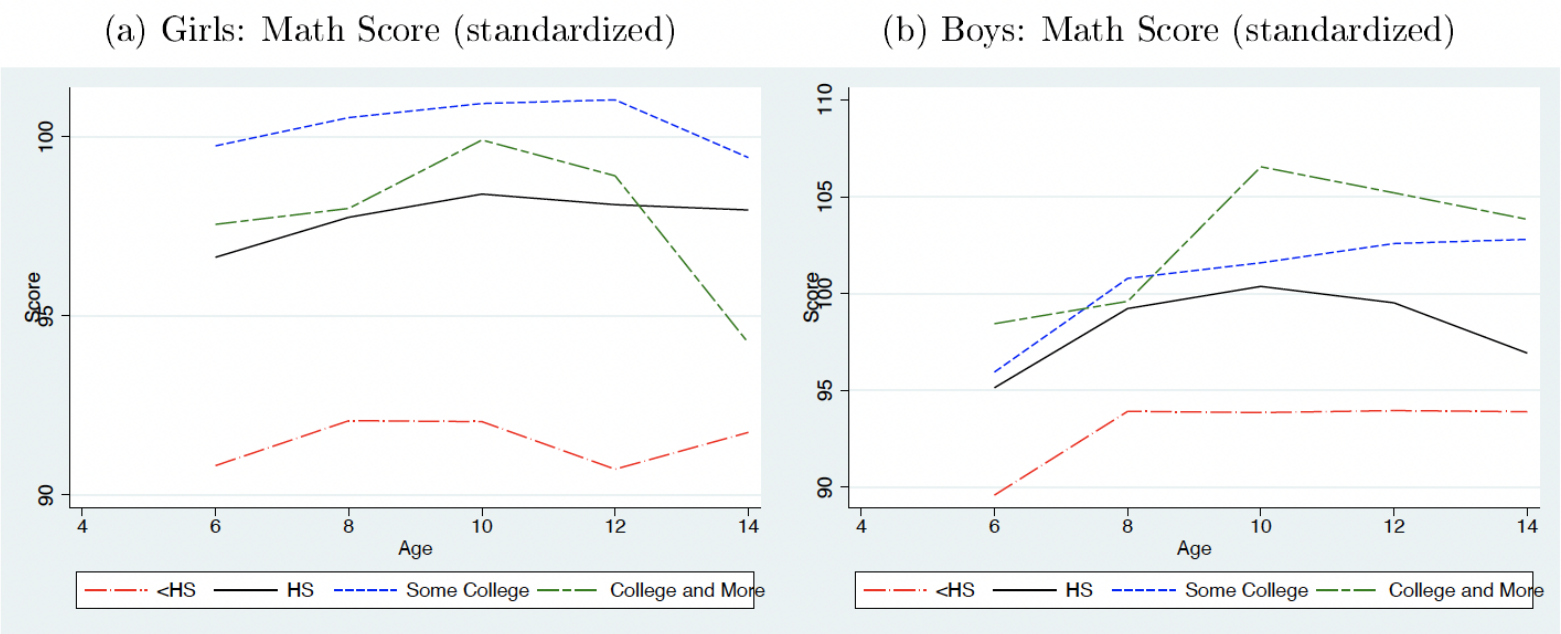


Figure 31: Skill Measures over Childhood by Mother's Education: Hispanic

Source: Heckman (2010)

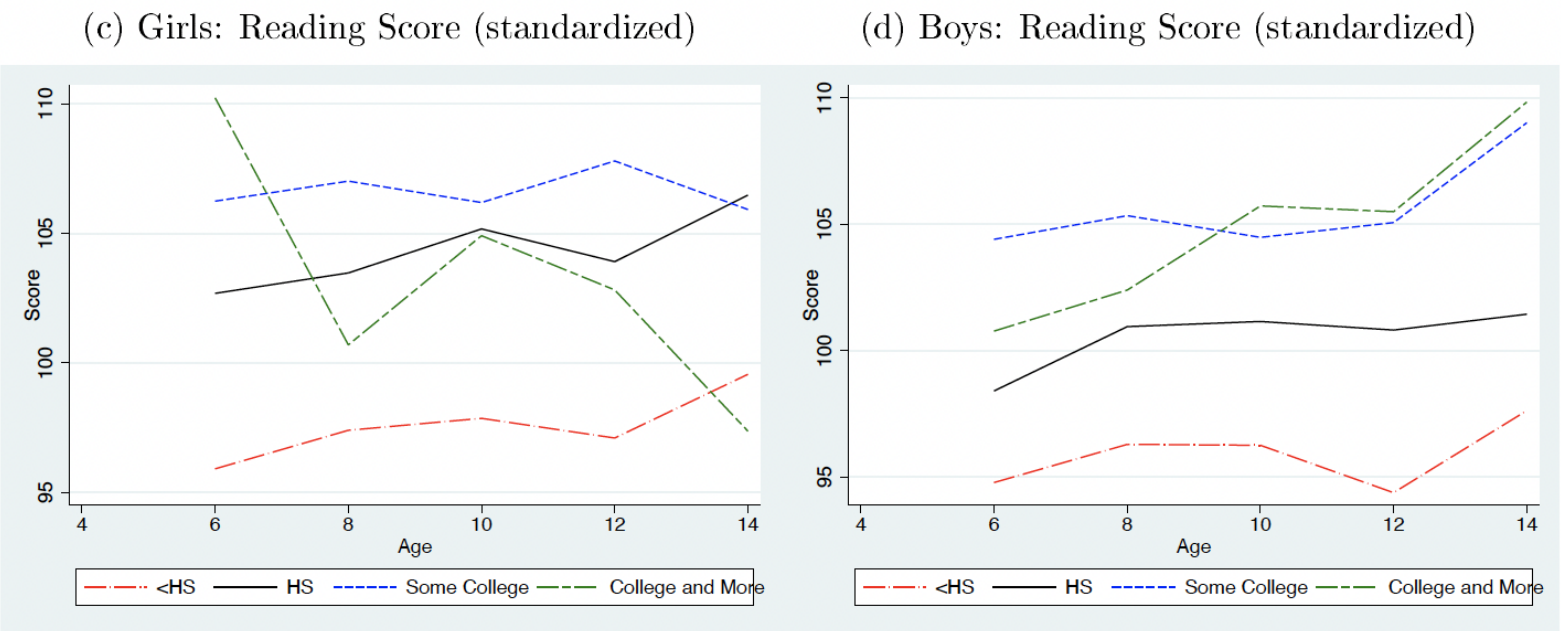


Figure 31: Skill Measures over Childhood by Mother's Education: Hispanic (Cont'd)

Source: Heckman (2010)

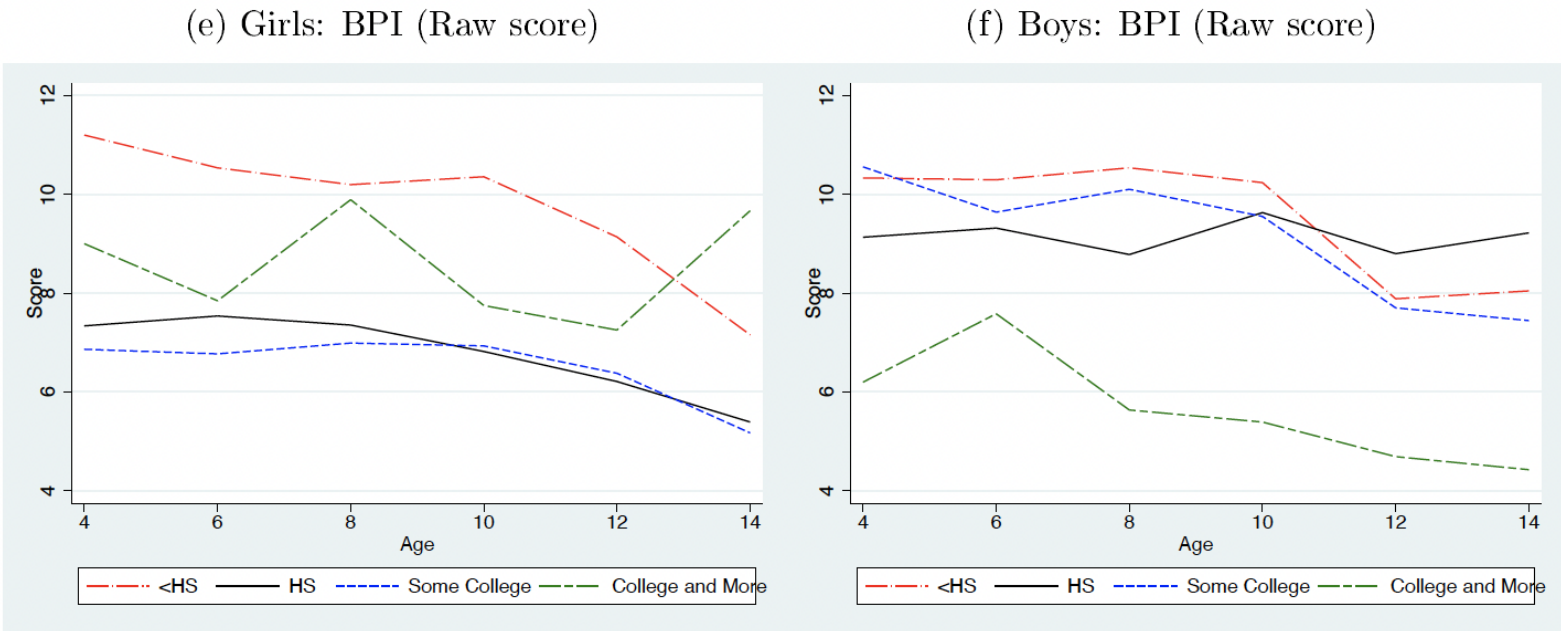


Figure 31: Skill Measures over Childhood by Mother's Education: Hispanic (Cont'd)

Source: Heckman (2010)

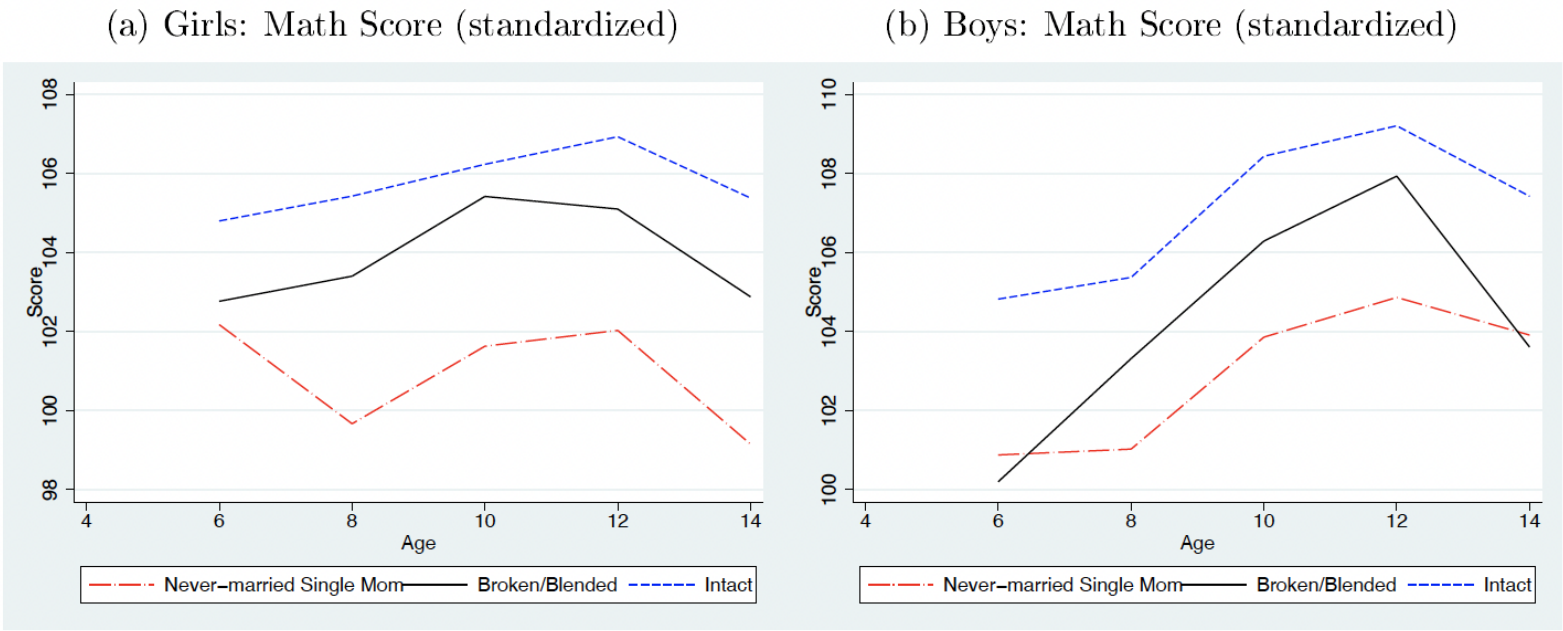


Figure 32: Skill Measures over Childhood among Whites by Family Type

Source: Heckman (2010)

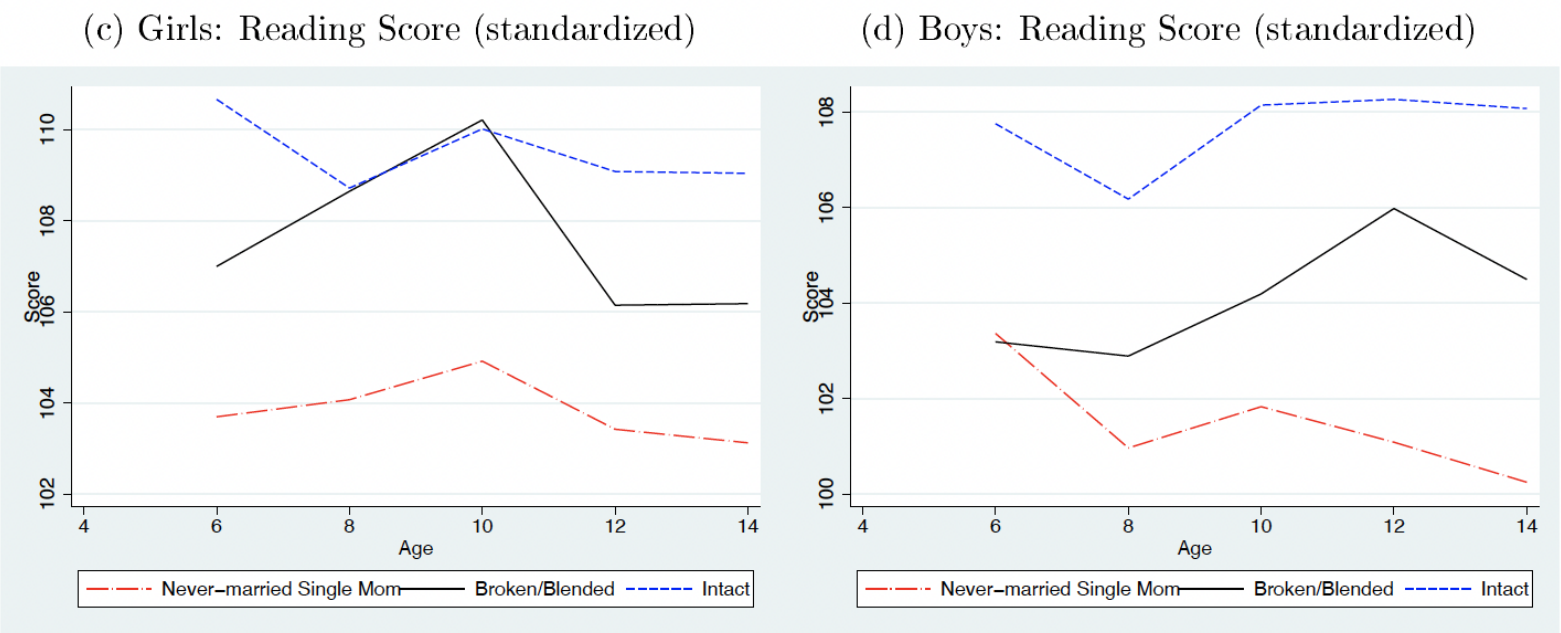


Figure 32: Skill Measures over Childhood among Whites by Family Type (Cont'd)

Source: Heckman (2010)

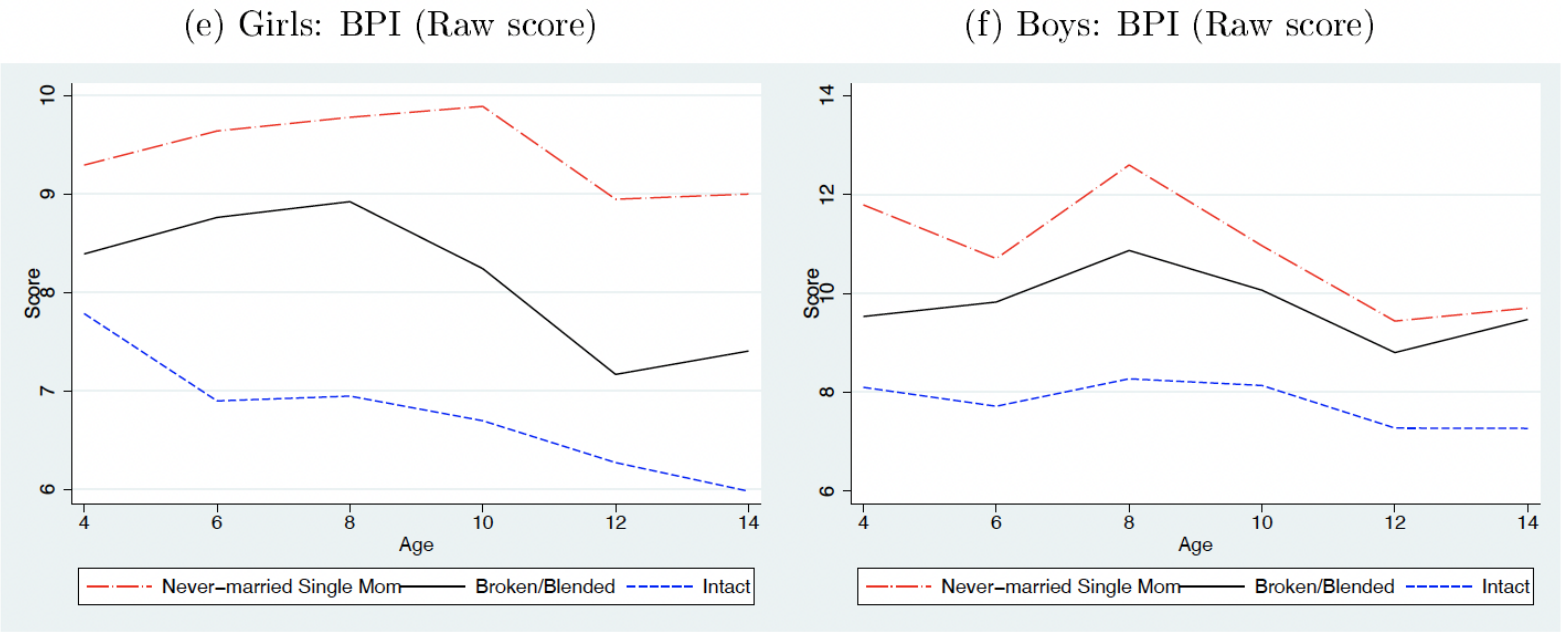


Figure 32: Skill Measures over Childhood among Whites by Family Type (Cont'd)

Source: Heckman (2010)

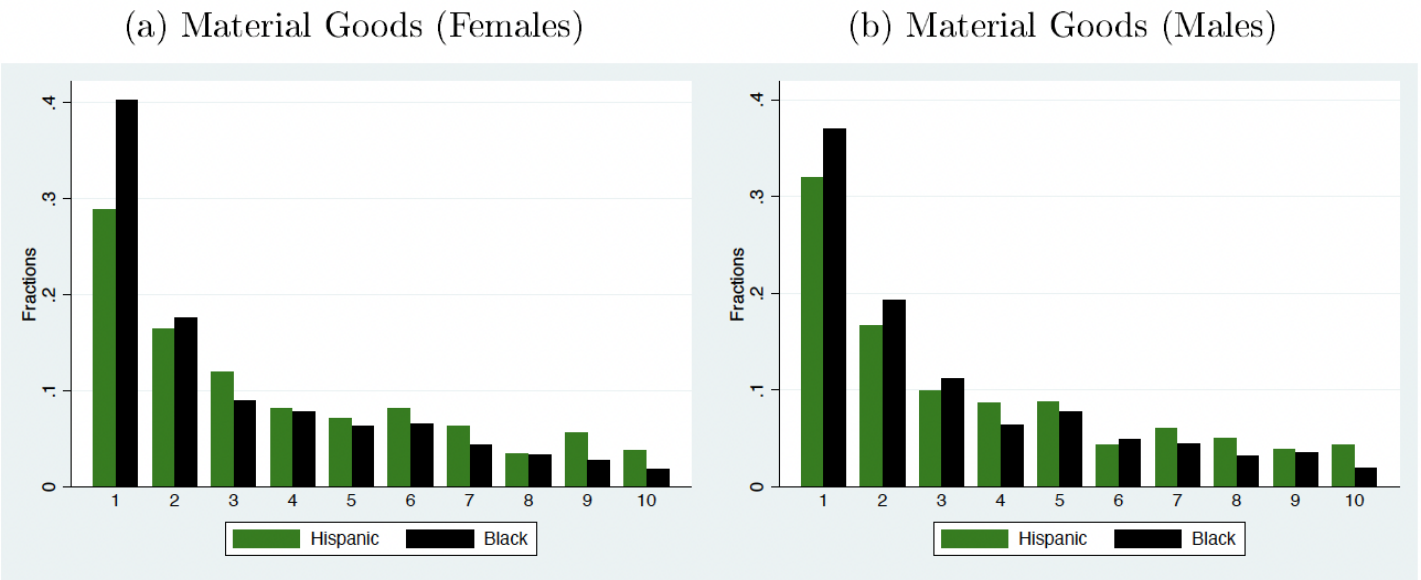


Figure 33: Hispanic and Black Parental Investment in White Distribution: Full Sample, Age 0-3

Source: Heckman (2010)

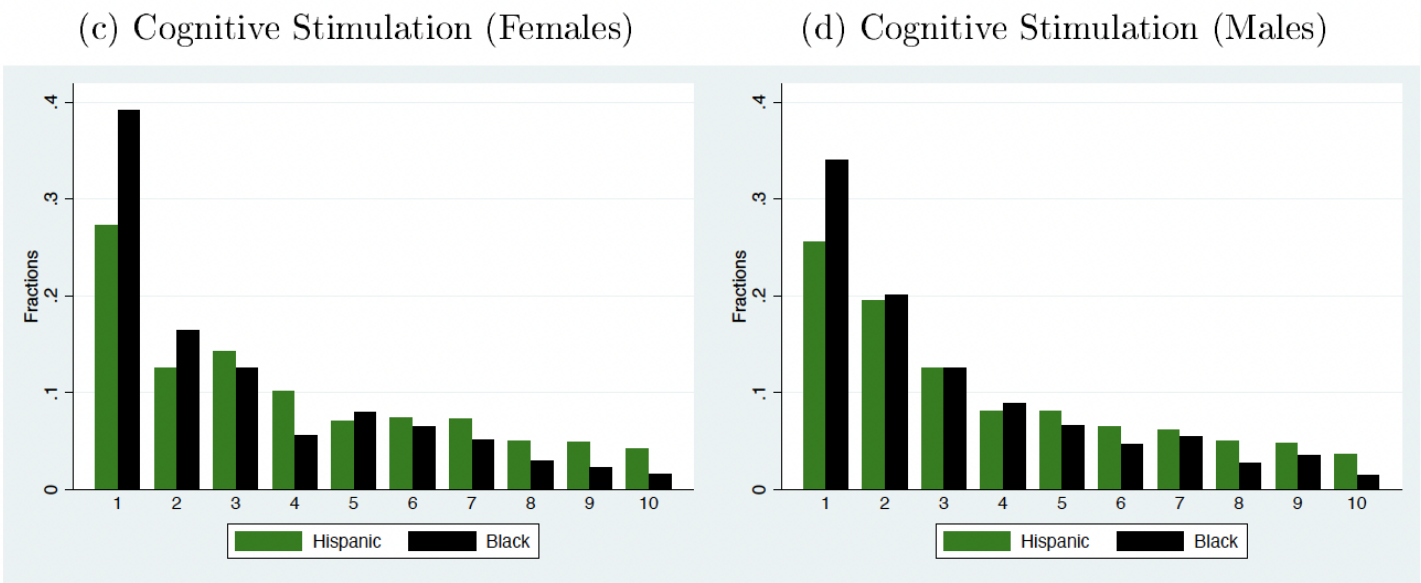


Figure 33: Hispanic and Black Parental Investment in White Distribution: Full Sample, Age 0-3 (Cont'd)

Source: Heckman (2010)

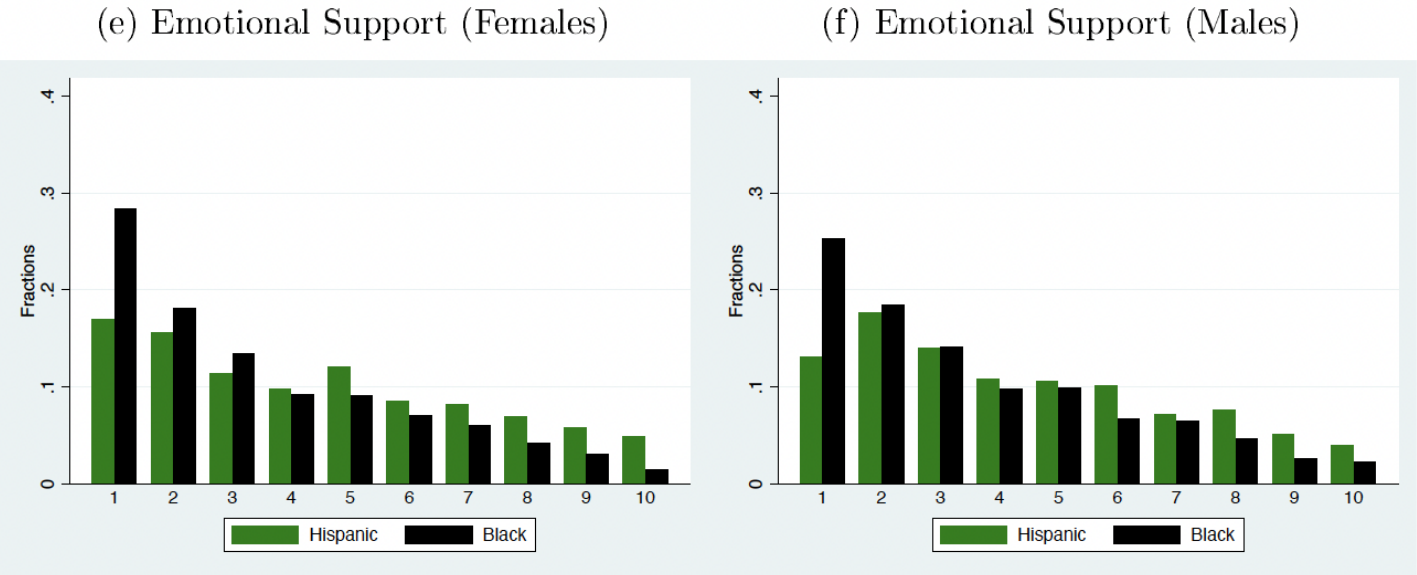


Figure 33: Hispanic and Black Parental Investment in White Distribution: Full Sample, Age 0-3 (Cont'd)

Source: Heckman (2010)

Correlational Evidence

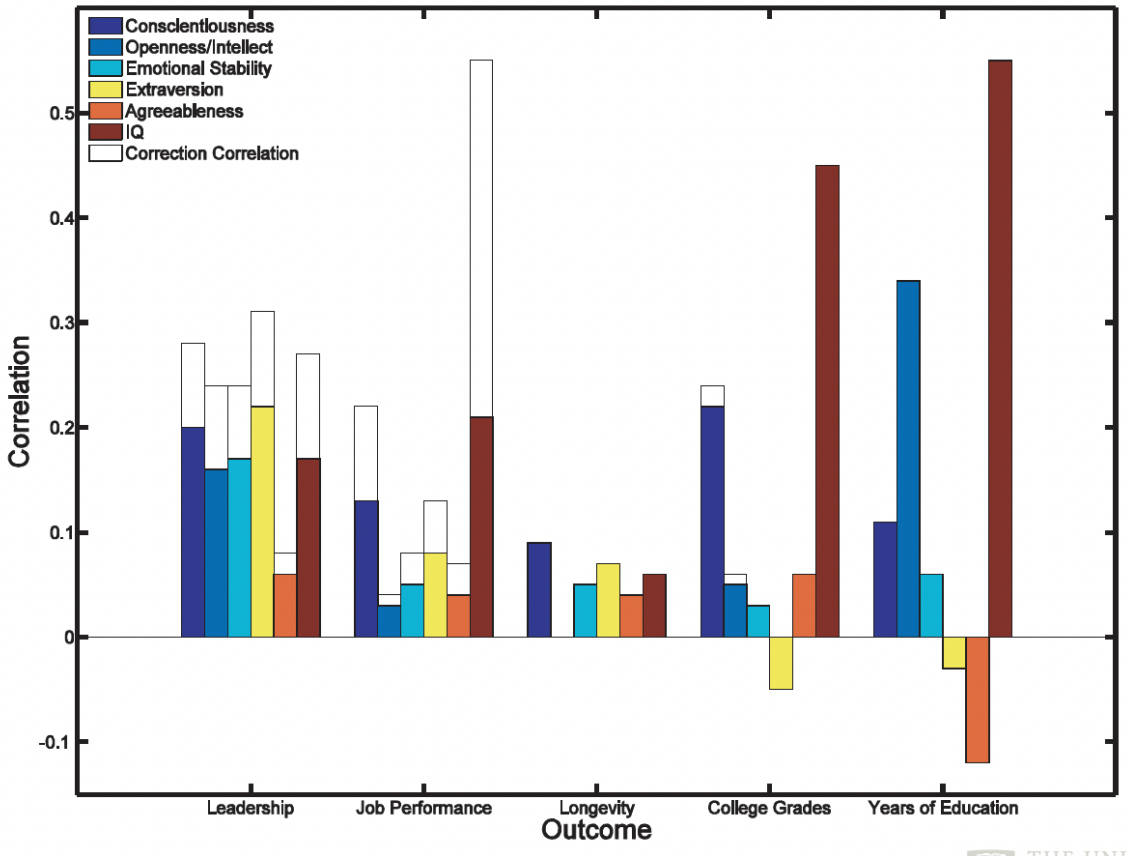


Figure 34: Associations with Lifecycle Outcomes

Source: Borghans at al. (2008)

All Psychological Measurements are on Performance on Some Task; Actions Taken

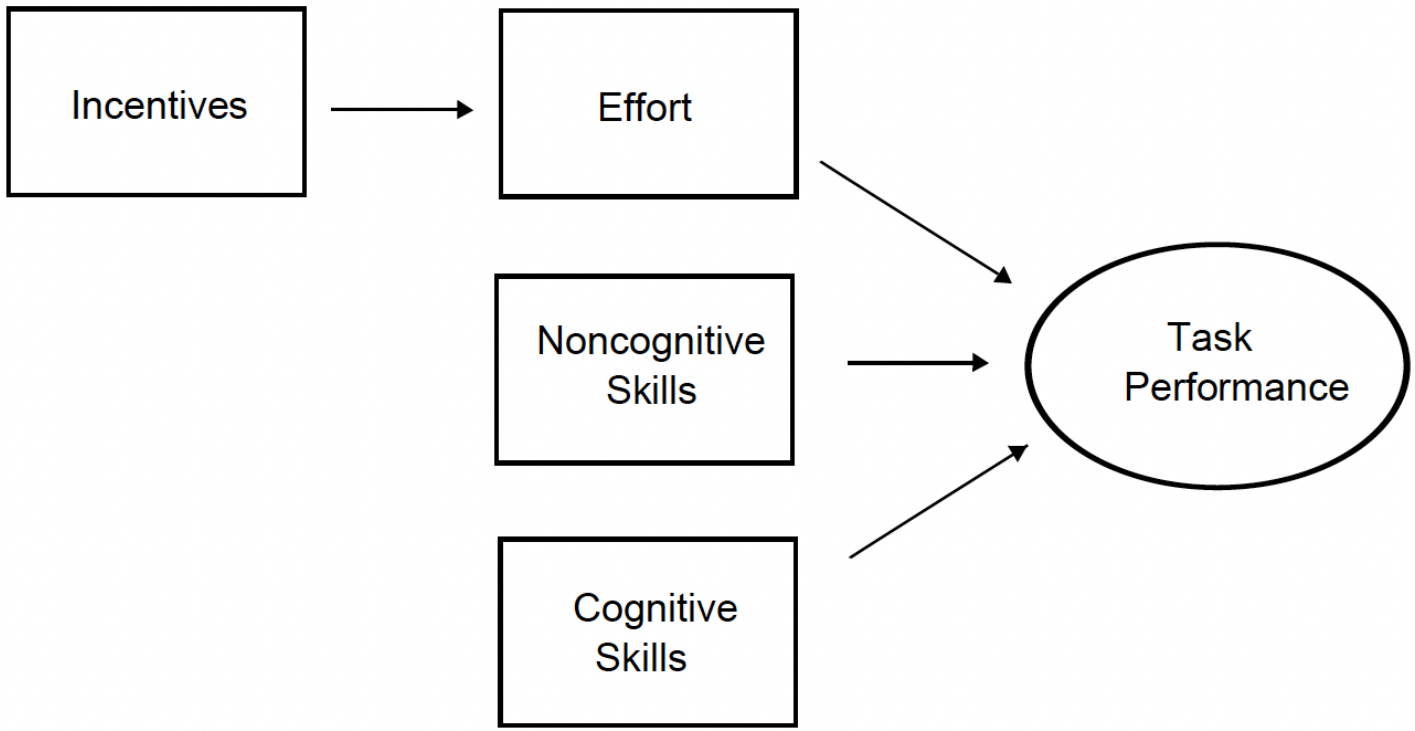


Figure 35: Determinants of Task Performance

IQ Scores Reflect Incentives and Measure Both Cognitive and Personality Traits

THE UNIVERSITY OF CHICAGO

Incentives and Performance on Intelligence Tests

Study	Sample and Study Design	Experimental Group	Effect size of incentive (in standard deviations)	Summary
Edlund (1972)	Between subjects study. 11 matched pairs of low-SES children; children were about one standard deviation below average in IQ at baseline	M&M candies given for each right answer	Experimental group scored 12 points higher than control group during a second testing on an alternative form of the Stanford-Binet (about 0.8 standard deviations)	"…a carefully chosen consequence, candy, given contingent on each occurrence of correct responses to an IQ test, can result in a significantly higher IQ score."(p. 319)
Ayllon & Kelly (1972) Sample 1	Within subjects study. 12 mentally retarded children (average IQ 46.8)	Tokens given in experimental condition for right answers exchangeable for prizes	6.25 points out of a possible 51 points on Metropolitan Readiness Test. $t = 4.03$	"…test scores often reflect poor academic skills, but they may also reflect lack of motivation to do well in the criterion test…These results, obtained from both a population typically limited in skills and ability as well as from a group of normal children (Experiment II), demonstrate that the use of reinforcement procedures applied to a behavior that is tacitly regarded as "at its peak" can significantly alter the level of performance of that behavior." (p. 483)
Ayllon & Kelly (1972) Sample 2	Within subjects study 34 urban fourth graders (average IQ = 92.8)	Tokens given in experimental condition for right answers exchangeable for prizes	$t = 5.9$	

Figure 36: Incentives and Performance on Intelligence Tests

Ayllon & Kelly (1972) Sample 3	Within subjects study of 12 matched pairs of mentally retarded children	Six weeks of token reinforcement for good academic performance	Experimental group scored 3.67 points out of possible 51 points on a post-test given under standard conditions higher than at baseline; control group dropped 2.75 points. On a second post-test with incentives, exp and control groups increased 7.17 and 6.25 points, respectively	
Clingman and Fowler (1976)	Within subjects study of 72 first- and second-graders assigned randomly to contingent reward, noncontingent reward, or no reward conditions.	M&Ms given for right answers in contingent condition; M&Ms given regardless of correctness in noncontingent condition	Only among low-IQ (<100) subjects was there an effect of the incentive. Contingent reward group scored about .33 standard deviations higher on the Peabody Picture Vocabulary test than did no reward group.	"…contingent candy increased the I.Q. scores of only the 'low I.Q.' children. This result suggests that the high and medium I.Q. groups were already functioning at a higher motivational level than children in the low I.Q. group." (p. 22)
Zigler and Butterfield (1968)	Within and between subjects study of 52 low-SES children who did or did not attend nursery school were tested at the beginning and end of the year on Stanford-Binet Intelligence Test under either optimized or standard conditions.	Motivation was optimized without giving test-relevant information. Gentle encouragement, easier items after items were missed, and so on.	At baseline (in the fall), there was a full standard deviation difference (10.6 points and SD was about 9.5 in this sample) between scores of children in the optimized vs standard conditions. The nursery group improved their scores, but only in the standard condition.	"…performance on an intelligence test is best conceptualized as reflecting three distinct factors: (a) formal cognitive processes; (b) informational achievements which reflect the content rather than the formal properties of cognition, and (c) motivational factors which involve a wide range of personality variables. (p. 2) "…the significant difference in improvement in standard IQ performance found between the nursery and nonnursery groups was attributable solely to motivational factors…" (p. 10)

(continued)

Figure 36: Incentives and Performance on Intelligence Tests (Cont'd)

Study	Sample and Study Design	Experimental Group	Effect size of incentive (in standard deviations)	Summary
Breuning and Zella (1978)	Within and between subjects study of 485 *special education* high school students all took IQ tests, then were randomly assigned to control or incentive groups to retake tests. Subjects were below-average in IQ.	Incentives such as record albums, radios (<$25) given for improvement in test performance	Scores increased by about 17 points. Results were consistent across the Otis-Lennon, WISC-R, and Lorge-Thorndike tests.	"In summary, the promise of individualized incentives contingent on an increase in IQ test performance (as compared with pretest performance) resulted in an approximate 17-point increase in IQ test scores. These increases were equally spread across subtests. The incentive condition effects were much less pronounced for students having pretest IQs between 98 and 120 and did not occur for students having pretest IQs between 121 and 140." (p. 225)
Holt and Hobbs (1979)	Between and within subjects study of 80 delinquent boys randomly assigned to three experimental groups and one control group. Each exp group received a standard and modified administration of the WISC-verbal section.	Exp 1-Token reinforcement for correct responses; Exp 2 – Tokens forfeited for incorrect responses (punishment), Exp 3-feedback on correct/incorrect responses	1.06 standard deviation difference between the token reinforcement and control groups (inferred from $t = 3.31$ for 39 degrees of freedom)	"Knowledge of results does not appear to be a sufficient incentive to significantly improve test performance among below-average I.Q. subjects…Immediate rewards or response cost may be more effective with below-average I.Q. subjects while other conditions may be more effective with average or above-average subjects." (p. 83)

Figure 36: Incentives and Performance on Intelligence Tests (Cont'd)

Larson, Saccuzzo, and Brown (1994)	Between subjects study of 109 San Diego State University psychology students	Up to $20 for improvement over baseline performance on cognitive speed tests	"While both groups improved with practice, the incentive group improved slightly more." (p. 34) $F(1,93) = 2.76, p < .05$	2 reasons why incentive did not produce dramatic increase: 1) few or no unmotivated subjects among college volunteers, 2) information processing tasks are too simple for 'trying harder' to matter
Duckworth (2007)	Within subjects study of 61 urban low-achieving high school students tested with a group-administered Otis-Lennon IQ test during their freshman year, then again 2 years later with a one-on-one (WASI) test	Standard directions for encouraging effort were followed for the WASI brief test. Performance was expected to be higher because of the one-on-one environment.	Performance on the WASI as juniors was about 16 points higher than on the group-administered test as freshmen. Notably, on the WASI, this population looks almost "average" in IQ, whereas by Otis-Lennon standards they are low IQ. $t(60) = 10.67, p < .001$	The increase in IQ scores could be attributed to any combination of the following 1) an increase in "g" due to schooling at an intensive charter school, 2) an increase in knowledge or crystallized intelligence, 3) an increase in motivation due to the change in IQ test format, and/or 4) an increase in motivation due to experience at high performing school

Figure 36: Incentives and Performance on Intelligence Tests (Cont'd)

Source: Borghans at al. (2008)

Sensitivity of self-reported noncognitive skills to survey administration conditions

- Self-reported measures of noncognitive skills might capture other dimensions aside from the skill, such as aspects of a respondent's situation, which could include incentives and the conditions in which they complete the questionnaire.
- Research has been done to estimate the extent to which survey administration conditions can affect student responses on noncognitive skill questionnaires.

Sensitivity of self-reported noncognitive skills to survey administration conditions (Cont'd)

- The first experiment tested whether providing information about the importance of noncognitive skills to students directly affects their responses. In treatment classrooms, the survey administrator provided instructions for completing the survey and read a description of the Big Five and their importance for life outcomes.
- The treatment condition was designed to mimic aspects of noncognitive skill development interventions that define and explain the importance of various skills.

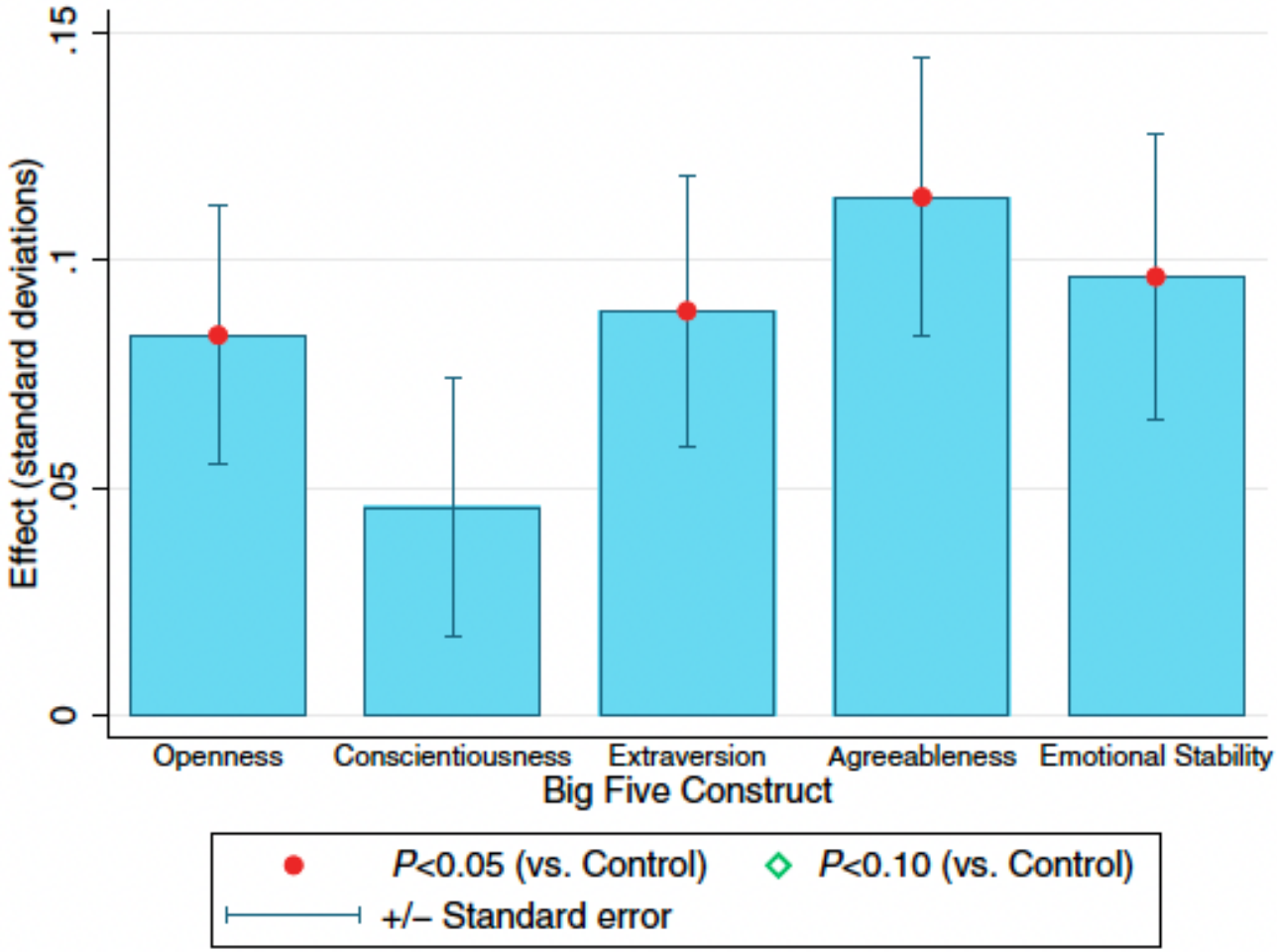


Figure 37: Effect of the explanation condition on students' self-reported Big Five

Source: Chen at al. (2019)

Sensitivity of self-reported noncognitive skills to survey administration conditions (Cont'd)

- The second experiment tested whether incentives tied to performance on another task indirectly affect responses. In this experiment, immediately before taking a math test and completing the BFI, students were randomly assigned to 1 of 3 groups: 1) a control group, 2) a treatment group that could receive a certificate of recognition if they performed well on the math test (honor incentive), or 3) a treatment group that could receive financial rewards if they performed well on the math test (financial incentive).

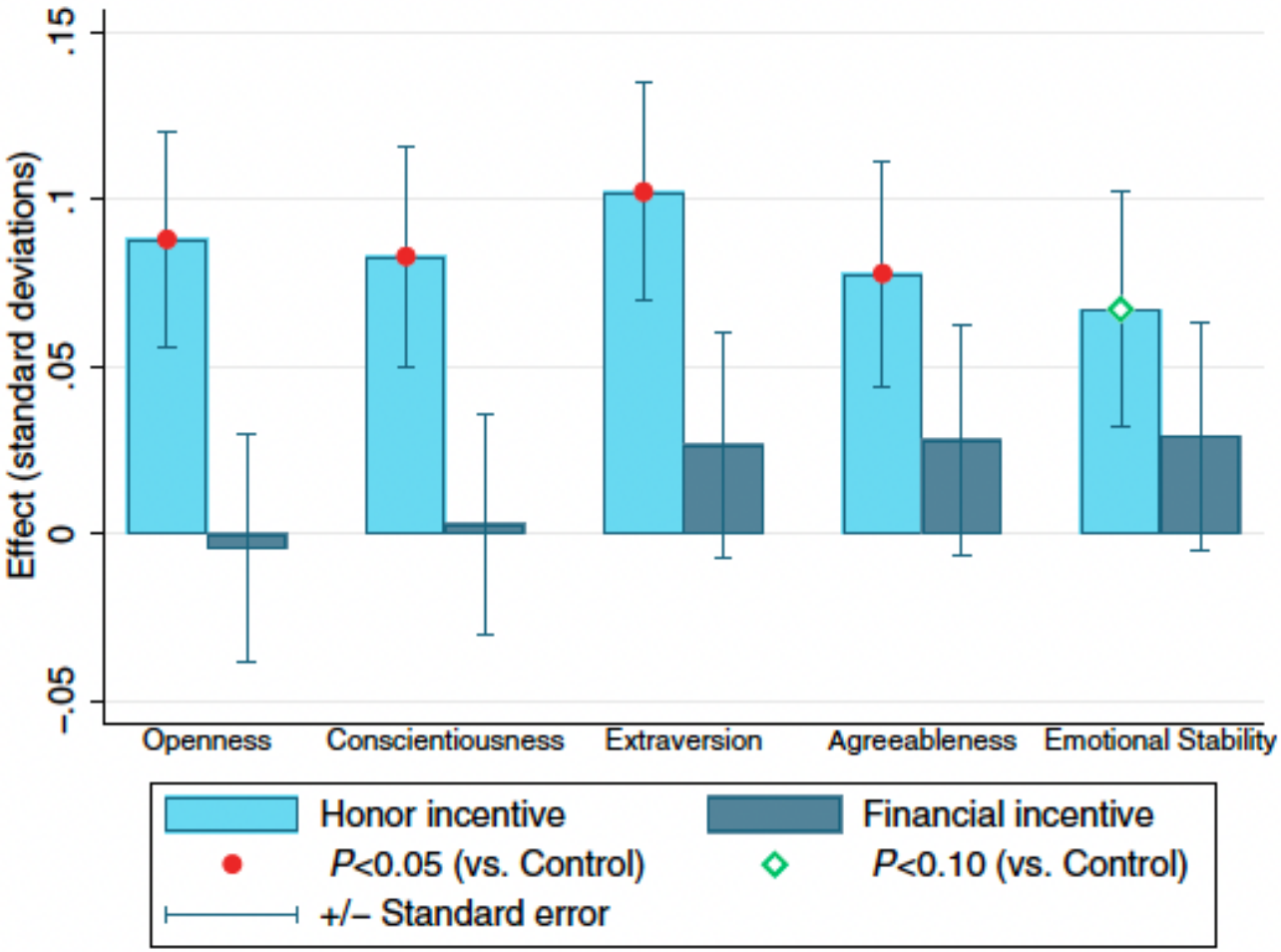


Figure 38: Effects of the honor and financial incentives on students' self-reported Big Five

Source: Chen at al. (2019)

Externalizing Behavior

- One outcome of interest in child development is the rate of externalizing behaviors.
- Externalizing behaviors include aggressive, antisocial, and rule-breaking behaviors.
- High rates of externalizing behaviors in adolescence are correlated with worse labor market outcomes and health behaviors into adulthood.
- Improvements in social and emotional skills have been shown to correlate with a reduction in these externalizing behaviors.

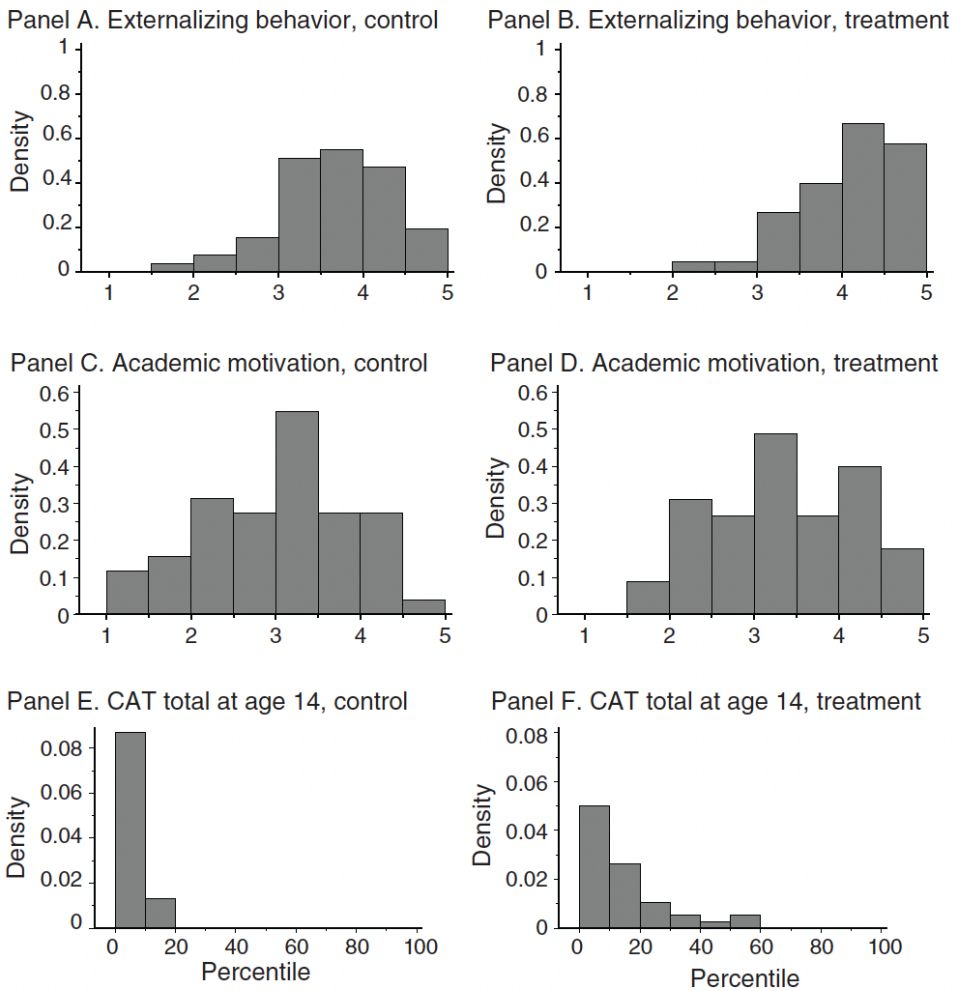


Figure 39: Histograms of Indices of Personality Skills and CAT Scores

Source: Heckman at al. (2013)

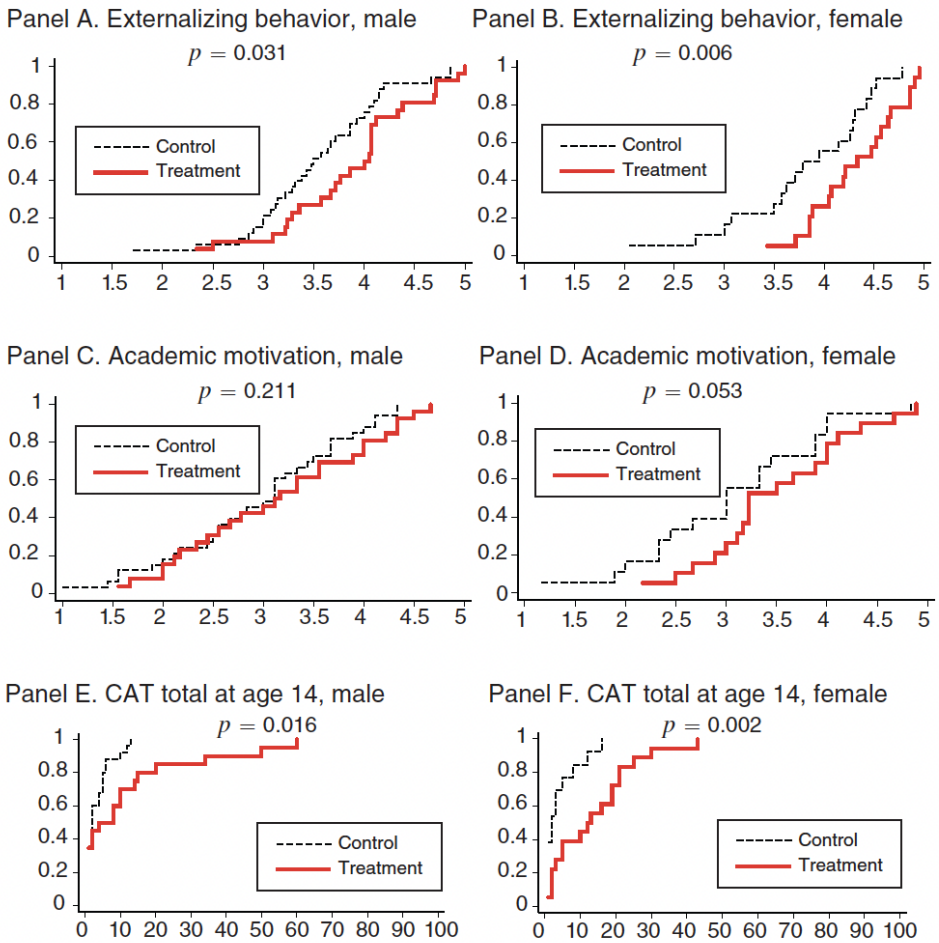


Figure 40: Cumulative Distribution Functions of Indices of Personality Skills and CAT Scores by Gender

Source: Heckman at al. (2013)

TABLE 2—COGNITIVE AND PERSONALITY FACTORS AND THEIR MEASURES

Cognition		Externalizing behavior		Academic motivation	
Measures[a]	Age	Measures[a]	Age[b]	Measures[a]	Age[b]
Stanford-Binet IQ	7	Disrupts classroom procedures	7–9	Shows initiative	7–9
Stanford-Binet IQ	8	Swears or uses obscene words	7–9	Alert and interested in school work	7–9
Stanford-Binet IQ	9	Steals	7–9	Hesitant to try, or gives up easily	7–9
		Lying or cheating	7–9		
		Influences others toward troublemaking	7–9		
		Aggressive toward peers	7–9		
		Teases or provokes students	7–9		
Cronbach's alpha,[c] males	0.838	Cronbach's alpha, males	0.906	Cronbach's alpha, males	0.901
Cronbach's alpha, females	0.913	Cronbach's alpha, females	0.916	Cronbach's alpha, females	0.896

Figure 41: Cognitive and Personality Factors and Their Measures

Source: Heckman at al. (2013)

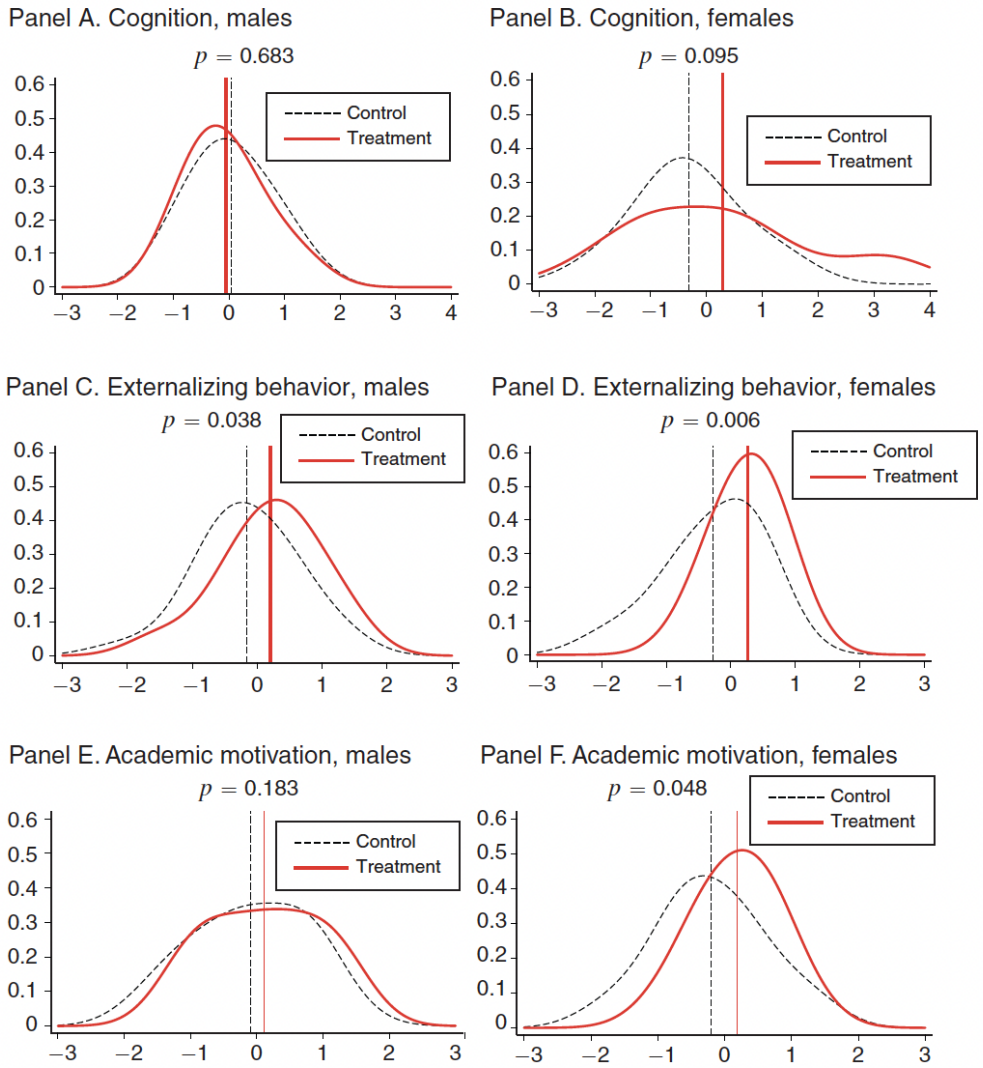


Figure 42: Kernel Densities of Factor Scores

Source: Heckman at al. (2013)

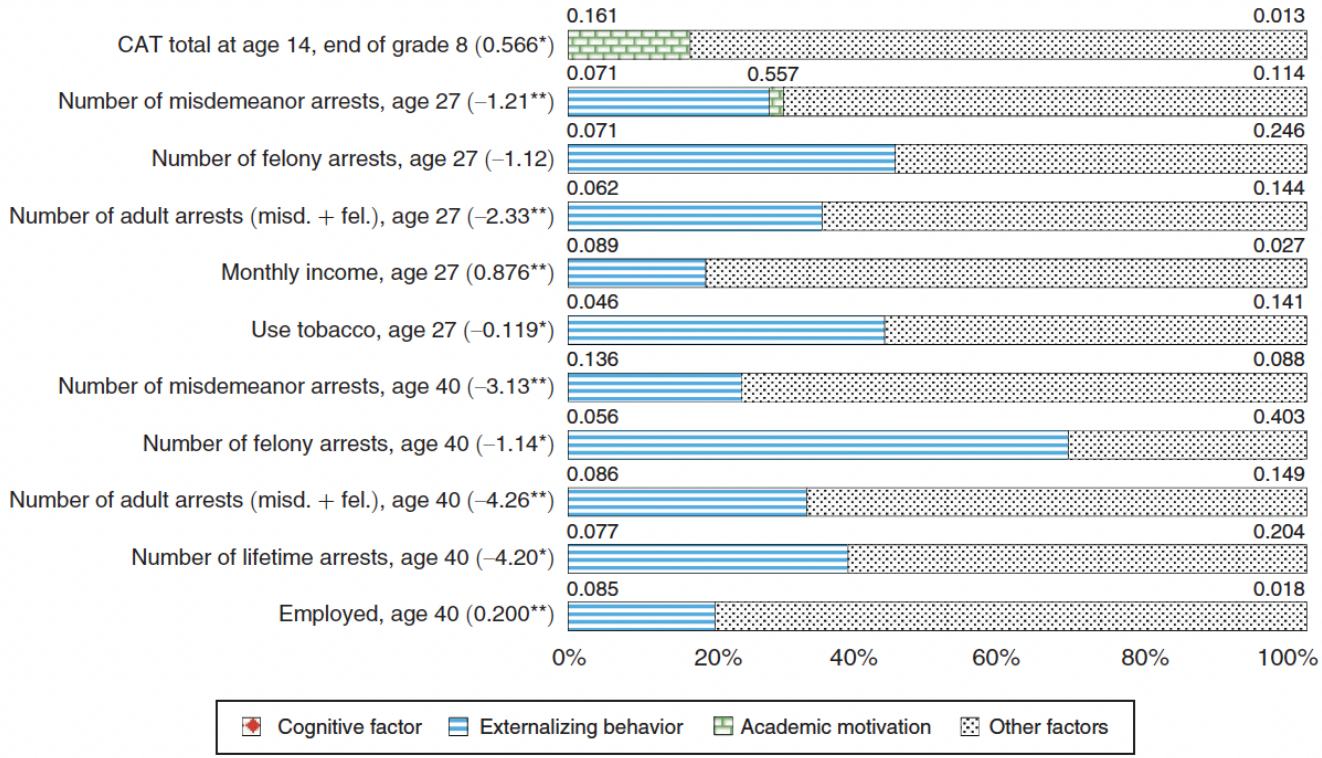


Figure 43: Decompositions of Treatment Effects on Outcomes, Males

Source: Heckman at al. (2013)

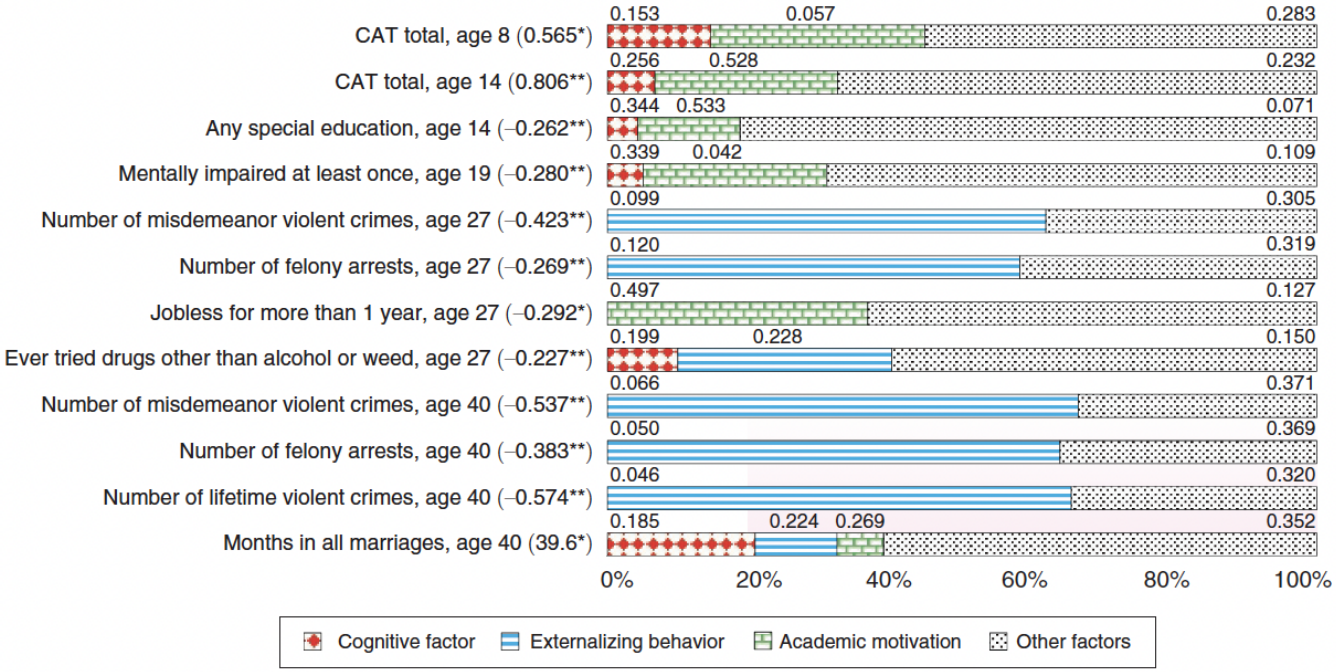


Figure 44: Decompositions of Treatment Effects on Outcomes, Females

Source: Heckman at al. (2013)

- As with most studies in personality psychology, the evidence presented in most of the literature do not address the question of causality; that is, do measured skills cause (rather than just predict) outcomes?
- Empirical associations are not a reliable basis for policy analysis.
- As previously noted (see Figure 35), multiple skills and effort all generate performance in a given task.
- Many studies in personality psychology do not control for all of the factors that produce performance on measured tasks.
- They equate measures of outcomes with the skill being measured.
- This practice can lead to a substantial bias in inference about the importance of any particular skill.

Conceptualizing Personality and Personality Traits within Economic Models

- Personality psychologists rarely use formal models to define or measure their constructs.
- In order to introduce their knowledge to economists, we formalize their frameworks.
- Doing so makes the concepts of personality psychology more precise and provides a basis for measurement and policy analysis.
- Roberts' framework (Figure 45) captures the main features of the influential models used in personality psychology.
- We use it as a point of departure.
- Psychology adds new and often more nuanced descriptions of human behavior to the standard descriptions used in economics.

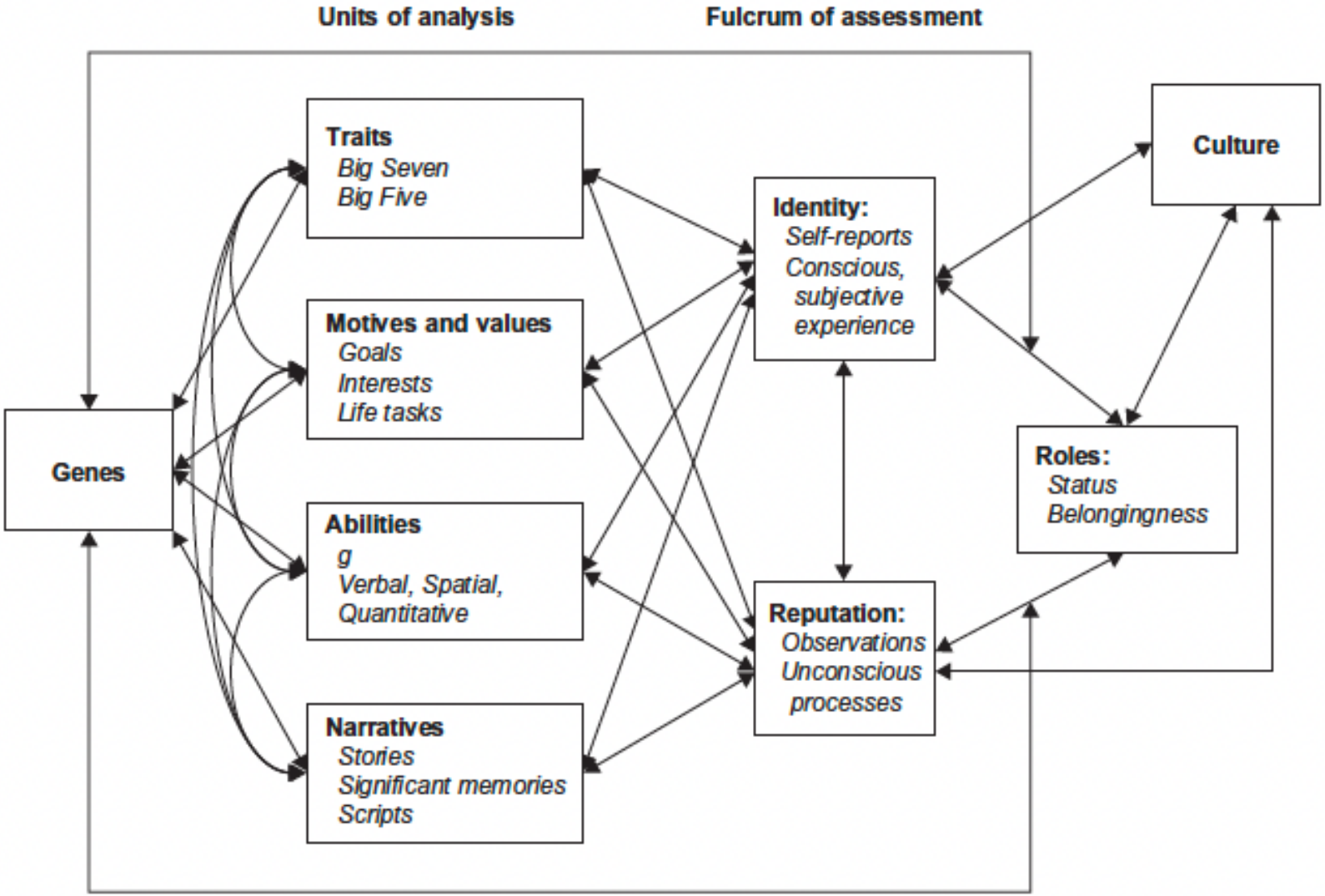


Figure 45: Roberts' Model of Personality as the Output of a System

- Preferences, constraints, and expectations provide the most direct way to introduce psychological variables into economic models.
- We begin our analysis with a barebones approach that focuses on constraints.
- For example, cognitive and personality traits affect earnings capacity because they enhance productivity.

Identifying Personality Traits

THE UNIVERSITY OF CHICAGO

Understanding Traits and the Challenge of Measurement

THE UNIVERSITY OF CHICAGO

- Let θ be a vector of traits at age a (we drop individual subscripts).
- They may change with age. θ_a cannot be directly observed.
- Instead, we observe behaviors B_a (e.g., studying, expressing empathy, taking an IQ test, lying, etc.).
- Behaviors depend on θ_a, but also incentives and rewards, R_a, to act in a certain way in a certain situation, S_a facing the agent.

$$B_a = \phi_a(\theta_a, R_a, S_a), a \in \{1, \ldots, A\}. \tag{1}$$

- B_a is a high dimensional vector and ϕ_a may depend on age (e.g., IQ manifests itself differently at different ages).
- For Mischel: ϕ_a depends only on R_a and S_a and there is no θ_a.

- θ_a as it relates to interventions (I_a), including parenting and schooling, and the effort (e_a) exerted by the person, measured in various ways.
- The technology of skill formation is:

$$\theta_{a+1} = \Gamma_a(\theta_a, I_a, e_a). \tag{2}$$

- Effort, e_a, depends on incentives to change (C_a).

- At the current level of generality, all traits can potentially affect productivity in all tasks. However, some tasks may require only a single trait or a subset of all of the traits.
- Following a traditional dichotomy in psychology that is explicit in Roberts' Figure 45, divide θ into "mental" (μ) and "personality" (π), traits: θ_μ and θ_π, each of which may in turn be a vector.

- Psychological measurement systems sometimes use productivity measured in different tasks to identify θ_μ and θ_π.
- This is the way Carroll (1993) defines mental ability where the task is performance on "mental" tests.
- To use performance on a task (or on multiple measures of the task) to identify a trait requires that performance on certain tasks (performance on a test, performance in an interpersonal situation, etc.) depends exclusively on one component of θ, say $\theta_{1,j}$.
- In that case,

$$P_j = \phi_j(\theta_{1,j}, e_j).$$

- Even if we can measure productivity P_j in task j, and only one component of θ affects P_j, to identify the level of a trait one must control for the level of effort applied to j in order to use P_j to infer the level of $\theta_{1,j}$.
- That is, one must standardize for the effort at a benchmark level, say e^* to use P_j to identify a measure of the trait that is uniform across different situations that elicit different levels of effort.

- The activity of picking a task (or a collection of tasks) to measure a particular trait ($\theta_{1,j}$ in our example) is called *operationalization* in psychology.
- *Construct validity* refers to whether or not a purported measure of the trait constructed in the stage of operationalization correlates with measures deemed to represent the trait.
- Considerable judgment is required to operationalize a trait and independently validate it.
- There is clear danger of circularity.
- Economists should carefully scrutinize how the measures they borrow from psychology are operationalized and validated in that literature.
- We should not necessarily assume that the measures created in that field have been rigorously established.

- Assuming that construct validity has been established, if effort is involved in the performance of a task used to uniquely define a trait, the measurement of performance must be standardized in order to use measured productivity, P_j, to identify the trait.
- Otherwise, the endowment of effort, and all of the factors that contribute to the exertion of effort, including the reward to the task, R_j, will contaminate the estimate of the trait.
- Failure to adjust for effort produces the kind of variability across situations with different rewards that was much discussed in the person-situation debate.

- Operationalization and construct validation clearly require heroic assumptions.
- Even if one adjusts for effort in a task, and thus adjusts for situational specificity, productivity in a task may depend on multiple traits.
- Thus two components of θ (say $\theta_{1,\mu}$, $\theta_{1,\pi}$) may determine productivity in task j.
- Without further information, one cannot infer which of the two traits produces the productivity in j. But in general, even having two (or more) measures of productivity that depend on $(\theta_{1,\mu}, \theta_{1,\pi})$ is not enough to identify the separate components.

- Consider the following case of two productivity measurements on tasks j and k ($j \neq k$):

$$P_j = \phi_j(\theta_{1,\mu}, \theta_{1,\pi}, e_j)$$

$$P_k = \phi_k(\theta_{1,\mu}, \theta_{1,\pi}, e_j)$$

  One might have such measurements if data are available on the productivity of the same person performing two different tasks.
- Standardize measurements at a common level of effort $e_j = e_k = e^*$.
- If the functional forms of the $\phi_j(\cdot)$ and $\phi_k(\cdot)$ are known, and the system of equations satisfies a local rank condition, then one can solve for the pair $(\theta_{1,\mu}, \theta_{1,\pi})$ at e^*.

- The rank condition might not be satisfied, and the functional forms ϕ_j and ϕ_k might not be known.
- The productivity functions need not be monotone in $\theta_{1,\mu}$ or $\theta_{1,\pi}$.
- Interacting systems might produce multiple equilibria so that the same values of θ produce different values of (P_j, P_k). Interacting systems might also have no solution.

Appendix: Demographic Differences

Big Five factor	*N*	*N* White	*N* Black	*K*	Mean *d*	SD_{obs}	SD_{res}	Lower 90% CV	Upper 90% CV	Direction
Emotional Stability	151,523	102,716	49,719	143	**−.09**	.25	.24	−.39	.22	W>B
Global measures[1]	135,409	88,450	47,871	128	**−.12**	.24	.23	−.40	.18	W>B
Self-esteem	13,212	12,070	1,142	9	**.17**	.22	.22	−.11	.45	B>W
Low anxiety	1,880	1,521	359	3	**−.23**	.06	.00	−.23	−.23	W>B
Even tempered	3,491	2,685	806	6	**.06**	.16	.14	−.12	.24	B>W
Extraversion	109,922	90,772	19,330	55	**−.16**	.33	.33	−.55	.22	W>B
Global measures[1]	39,989	29,789	10,380	28	**−.21**	.33	.32	−.62	.21	W>B
Dominance	39,552	34,338	5,214	15	**−.03**	.18	.18	−.20	.25	W>B
Sociability	27,592	24,127	3,465	11	**−.39**	.23	.23	−.67	.10	W>B
Openness to Experience	24,957	21,749	3,208	9	**−.10**	.07	.07	−.18	−.02	W>B
Agreeableness	25,247	21,590	3,297	9	**−.03**	.09	.09	−.14	.08	W>B
Conscientiousness	180,478	161,283	19,195	67	**.07**	.15	.15	−.12	.27	B>W
Global measures[1]	21,001	18,506	2,495	3	**.17**	.12	.12	.13	.43	B>W
Achievement	49,730	44,791	4,939	19	**−.03**	.10	.10	−.10	.15	W>B
Dependability	23,839	21,015	2,824	4	**−.05**	.08	.08	−.15	.06	W>B
Cautiousness	1,943	1,301	642	4	**.16**	.09	.03	.12	.19	B>W
Order	26,886	23,678	3,208	9	**.01**	.17	.17	−.21	.23	B>W

Note. [1]Global measures includes only those scales that assess traits at the broad factor level. *K* = number of independent effect sizes, mean *d* = mean *d* corrected for sampling error, SD_{obs} = standard deviation of observed *d*, SD_{res} = residual standard deviation, CV = 90% credibility value. Negative *d*-values indicate that Whites scored higher on the factor of interest.

Figure 46: Meta-Analysis of Black–White Differences in the Big Five Personality Factor and Facet Measures

Source: Ones (2008)

Big Five factor	*N*	*N* White	*N* Asian	*K*	Mean *d*	SD_{obs}	SD_{res}	Lower 90% CV	Upper 90% CV	Direction
Emotional Stability	85,585	82,187	3,398	38	**−.12**	.37	.35	−.57	.33	W>A
Global measures[1]	61,529	59,313	2,216	26	**−.16**	.34	.34	−.59	.27	W>A
Self-esteem	12,294	12,074	220	5	**.30**	.21	.20	.03	.56	A>W
Low anxiety	808	728	80	3	**.27**	.29	.26	−.06	.61	A>W
Even tempered	10,954	10,072	882	4	**−.38**	.10	.09	−.49	−.26	W>A
Extraversion	56,267	53,254	3,013	30	**−.14**	.21	.20	−.40	.12	W>A
Global measures[1]	28,284	27,164	1,117	12	**−.07**	.22	.22	−.35	.20	W>A
Dominance	16,103	15,142	961	11	**−.19**	.16	.15	−.39	.04	W>A
Sociability	6,137	5,665	472	4	**−.09**	.12	.10	−.22	.04	W>A
Openness to Experience	1,596	1,464	132	6	**.11**	.12	.07	.03	.21	A>W
Agreeableness	1,309	1,216	93	5	**.63**	1.05	1.04	−.70	1.96	A>W
Conscientiousness	107,711	104,257	3,454	42	**.11**	.21	.21	−.15	.37	A>W
Global measures[1]	22,067	20,868	1,199	5	**.04**	.28	.28	−.33	.40	A>W
Achievement	49,304	48,437	864	20	**.14**	.12	.11	.00	.29	A>W
Dependability	532	488	44	2	**−.01**	.18	.12	−.17	.15	W>A
Order	5,743	5,280	463	3	**.50**	.10	.09	.39	.62	A>W

Note. [1]Global measures includes only those scales that assess traits at the broad factor level. *K* = number of independent effect sizes, mean *d* = mean *d* corrected for sampling error, SD_{obs} = standard deviation of observed *d*, SD_{res} = residual standard deviation, CV = 90% credibility value. Negative *d*-values indicate that Whites scored higher on the factor of interest.

Figure 47: Meta-Analysis of Asian–White Differences in the Big Five Personality Factor and Facet Measures

Source: Ones (2008)

Big Five factor	*N*	*N* Hispanic	*N* White	*K*	Mean *d*	SD_{obs}	SD_{res}	Lower 90% CV	Upper 90% CV	Direction
Emotional Stability	124,081	28,327	95,754	72	**.03**	.27	.26	−.31	.36	H>W
Global measures[1]	96,012	15,639	80,373	58	**−.04**	.26	.25	−.36	.29	W>H
Self-esteem	24,698	12,183	12,515	8	**.25**	.11	.11	.12	.39	H>W
Low anxiety	1,012	206	806	2	**.25**	.45	.45	−.32	.83	H>W
Even tempered	2,359	299	2,060	4	**.09**	.12	.09	−.03	.20	H>W
Extraversion	94,520	20,449	74,071	29	**−.02**	.17	.16	−.22	.19	W>H
Global measures[1]	33,459	13,000	20,459	10	**.12**	.05	.04	.07	.17	H>W
Dominance	34,991	4,376	30,615	11	**−.04**	.19	.19	−.29	.19	W>H
Sociability	24,112	3,036	21,076	6	**−.16**	.03	.01	−.17	−.14	W>H
Openness to Experience	24,993	3,082	21,911	10	**−.02**	.06	.05	−.09	.05	W>H
Agreeableness	24,640	3,052	21,588	7	**−.05**	.27	.27	−.30	.39	W>H
Conscientiousness	232,771	81,564	151,207	53	**.08**	.12	.12	−.07	.24	H>W
Global measures[1]	17,499	1,791	15,708	3	**.20**	.15	.15	.01	.40	H>W
Achievement	119,182	50,134	69,048	28	**.10**	.11	.10	−.03	.23	H>W
Dependability	23,854	2,994	20,860	4	**.00**	.08	.07	−.09	.10	H=W
Order	23,688	2,997	20,691	5	**.00**	.06	.05	−.06	.06	H=W

Note. [1]Global measures includes only those scales that assess traits at the broad factor level. *K* = number of independent effect sizes, mean *d* = mean *d* corrected for sampling error, SD_{obs} = standard deviation of observed *d*, SD_{res} = residual standard deviation, CV = 90% credibility value. Negative *d*-values indicate that Whites scored higher on the factor of interest.

Figure 48: Meta-Analysis of Hispanic–White Differences in the Big Five Personality Factor and Facet Measures

Source: Ones (2008)

Big Five factor	*N*	*N* White	*N* American Indian	*K*	Mean *d*	SD_{obs}	SD_{res}	Lower 90% CV	Upper 90% CV	Direction
Emotional Stability	22,840	22,097	743	16	**−.21**	.28	.27	−.56	.14	W>AI
Extraversion	8,602	8,090	512	18	**−.33**	.50	.49	−.95	.30	W>AI
Openness to Experience	*Insufficient*									
Agreeableness	407	337	70	3	**−.28**	.18	.04	−.32	−.23	W>AI
Conscientiousness	13,139	12,921	218	11	**.25**	.27	.26	−.09	.59	AI>W

Note. K = number of independent effect sizes, mean d = mean d corrected for sampling error, SD_{obs} = standard deviation of observed d, SD_{res} = residual standard deviation, CV = 90% credibility value. Negative d-values indicate that Whites scored higher on the factor of interest.

Figure 49: Meta-Analysis of American Indian–White Differences in the Big Five Personality Measures

Source: Ones (2008)

Big Five factor	*N*	*N* Black	*N* Hispanic	*K*	Mean *d*	SD_{obs}	SD_{res}	Lower 90% CV	Upper 90% CV	Direction
Emotional Stability	8,333	4,244	4,039	19	**−.02**	.17	.14	−.20	.16	H>B
Global measures[1]	7,509	3,736	3,723	14	**.04**	.16	.13	−.13	.21	B>H
Even tempered	806	507	299	4	**.02**	.05	.00	.02	.02	B>H
Extraversion	11,494	5,349	6,135	11	**−.11**	.19	.18	−.34	.12	H>B
Global measures[1]	418	245	163	3	**.01**	.05	.00	.01	.01	B>H
Dominance	5,491	2,515	2,976	3	**.08**	.01	.00	.08	.08	B>H
Sociability	5,567	2,588	2,979	4	**−.30**	.05	.00	−.30	−.30	H>B
Openness to experience	5,657	2,593	3,064	9	**−.10**	.06	.00	−.10	−.10	H>B
Agreeableness	5,527	2,517	3,010	5	**−.09**	.05	.00	−.09	−.09	H>B
Conscientiousness	16,950	7,849	9,101	17	**−.05**	.09	.06	−.12	−.02	H>B
Achievement	5,603	2,590	3,013	6	**−.06**	.10	.08	−.16	.03	H>B
Dependability	5,491	2,515	2,976	3	**−.05**	.07	.05	−.11	.01	H>B
Order	5,567	2,588	2,979	4	**−.05**	.04	.00	−.05	−.05	H>B

Note. [1]Global measures includes only those scales that assess traits at the broad factor level. K = number of independent effect sizes, mean d = mean d corrected for sampling error, SD_{obs} = standard deviation of observed d, SD_{res} = residual standard deviation, CV = 90% credibility value. Negative d-values indicate that Hispanics scored higher on the factor of interest.

Figure 50: Meta-Analysis of Black–Hispanic Differences in the Big Five Personality Factor and Facet Measures

Source: Ones (2008)

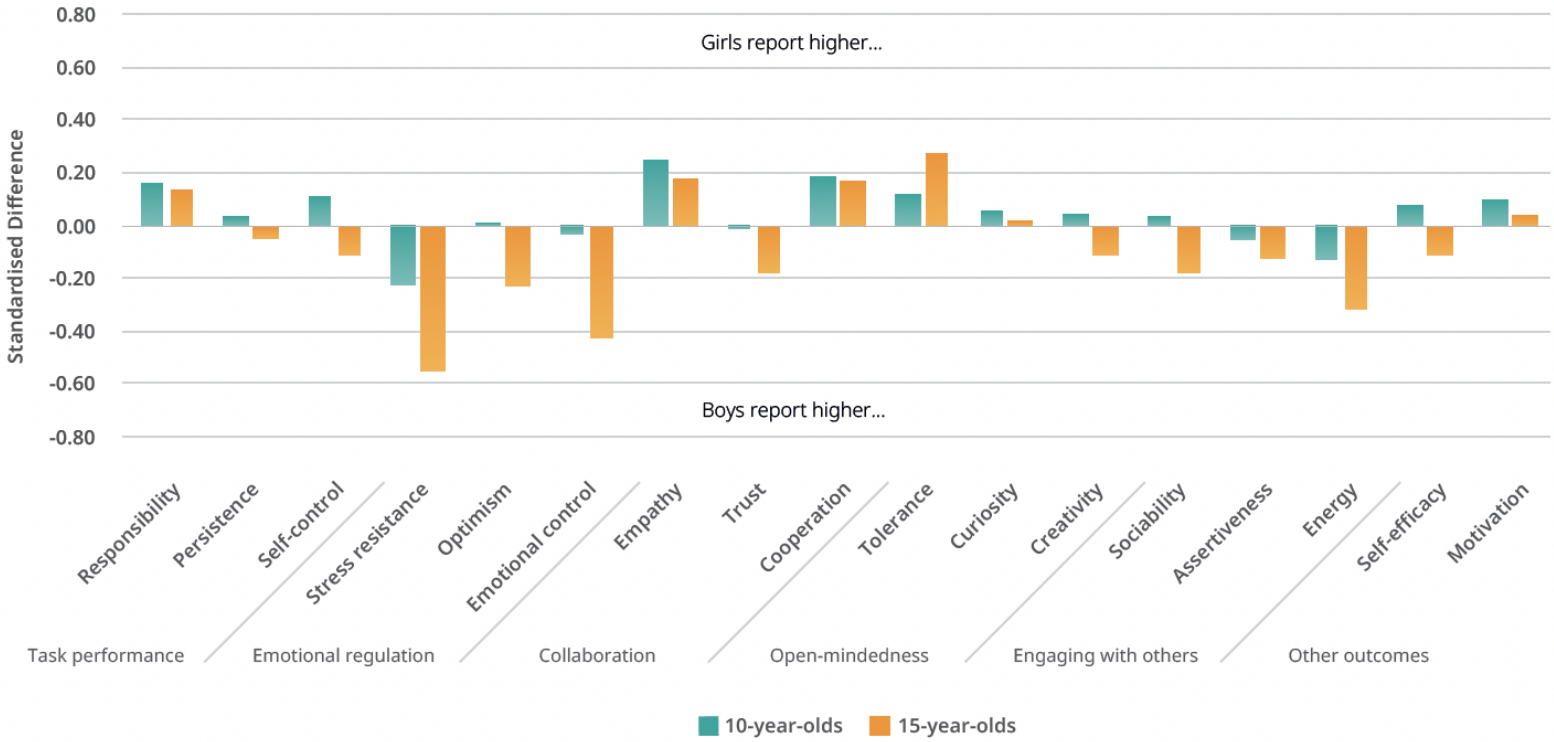


Figure 51: Gender differences in social and emotional skills

Source: OECD (2019)

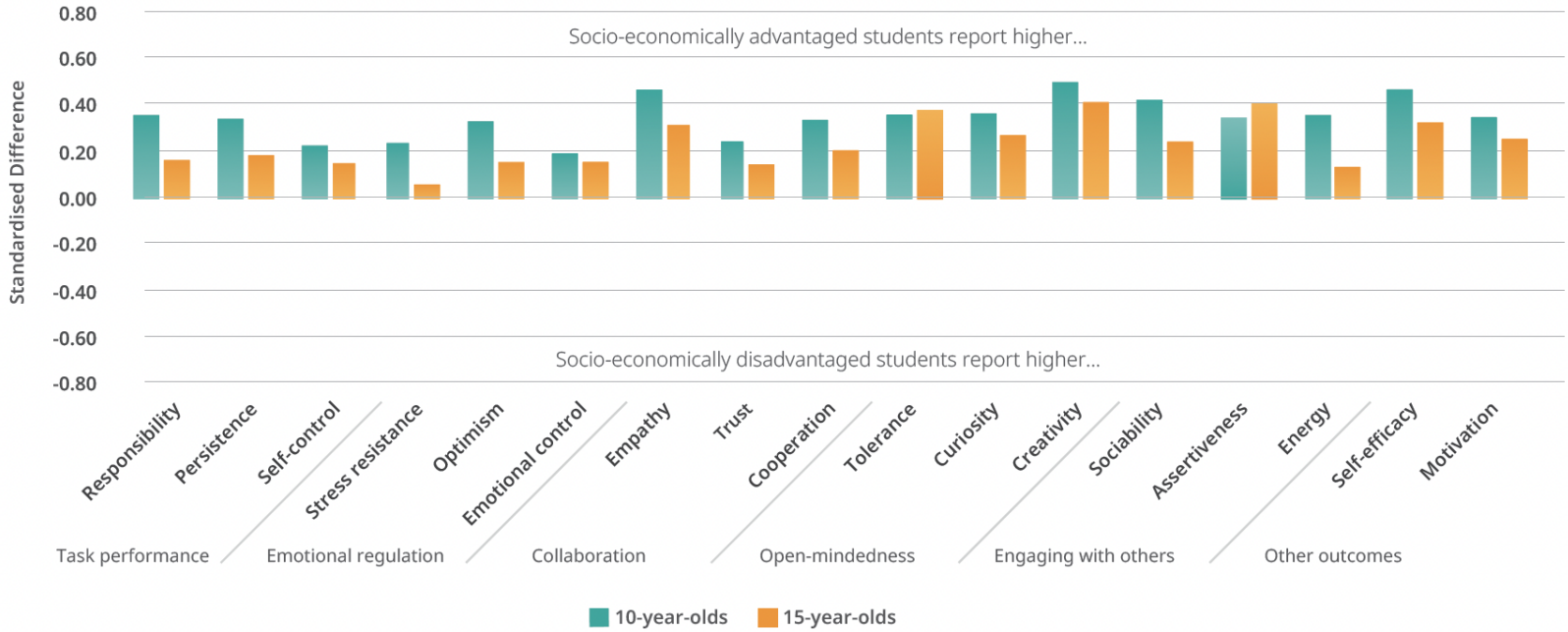


Figure 52: Socio-economic status differences in social and emotional skills, by age

Source: OECD (2019)

	12th Grade GPA	SSAT: V	SSAT: Q	SSAT: R	Race: W/NW	MF
12th Grade GPA	1.00	.50**	.38**	-.07	-.01	-.01
SSAT: V	.39**	1.00	.42**	.60**	.04	-.06
SSAT: Q	.51**	.42**	1.00	.39**	-.14**	-.13*
SSAT: R	.38**	.60**	.39**	1.00	.08	-.09
Race: W/NW	-.99	.04	-.14**	.08	1.00	-.04
MF	-.01	-.06	1.00	-.09	-.04	1.00
HEXACO: Honesty-Humility	-.02	-.02	.03	.04	.07	-.13
HEXACO: Emotionality	-.03	-.02	-.05	.02	.10*	.01
HEXACO: Extraversion	-.03	-.05	-.02	.01	.06	-.13*
HEXACO: Agreeableness	-.02	.03	.02	.07	.08	-.11*
HEXACO: Conscientiousness	-.09	.03	.02	.07	.08	-.11*
HEXACO: Openness to Exp.	-.06	-.06	-.04	-.09	.01	.01

Note. * $p < .05$, ** $p < .01$

Figure 53: Pearson Correlation Matrix of Personality, Demographics, and Achievement Scores

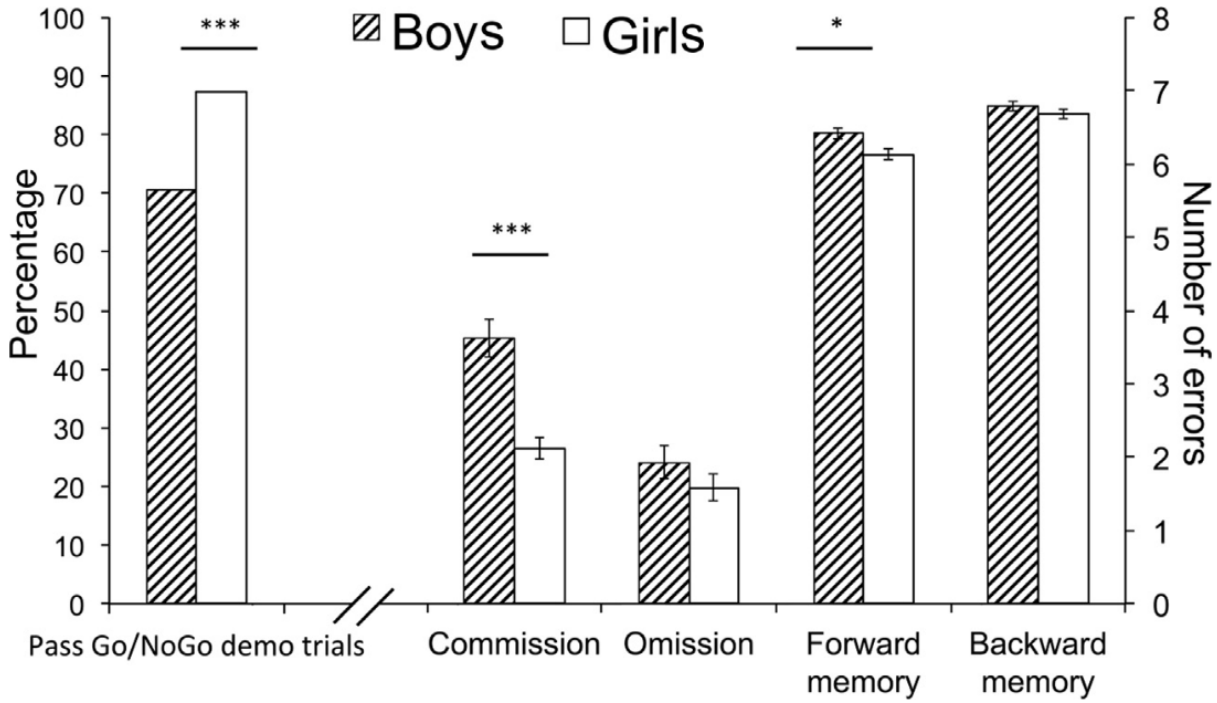


Figure 54: Executive Function and Gender

Source: Mileva-Seitz et al. (2014)

	Numbers Reversed				DCCS			
	Fall K	Spring K	Spring First	Spring Second	Fall K	Spring K	Spring First	Spring Second
Race								
Black	–0.54*	–0.54*	–0.42*	–0.41*	–0.42*	–0.44*	–0.45*	–0.48*
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)
Hispanic	–0.59*	–0.55*	–0.39*	–0.27*	–0.40*	–0.31*	–0.30*	–0.20*
	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Asian	0.03	0.06	0.12*	0.13*	–0.17*	–0.12*	–0.02	0.00
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)
Other race	–0.10*	–0.13*	–0.10*	–0.14*	–0.14*	–0.09*	–0.10*	0.02
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
SES quintile								
Second quintile	0.31*	0.32*	0.34*	0.20*	0.23*	0.16*	0.18*	0.23*
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Third quintile	0.52*	0.52*	0.45*	0.34*	0.35*	0.23*	0.27*	0.29*
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Fourth quintile	0.81*	0.74*	0.61*	0.47*	0.46*	0.35*	0.38*	0.44*
	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
Fifth quintile	1.01*	0.93*	0.75*	0.66*	0.59*	0.52*	0.52*	0.52*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)

Figure 55: Coefficients and Standard Errors for Racial and Socioeconomic Gaps in Executive Functions

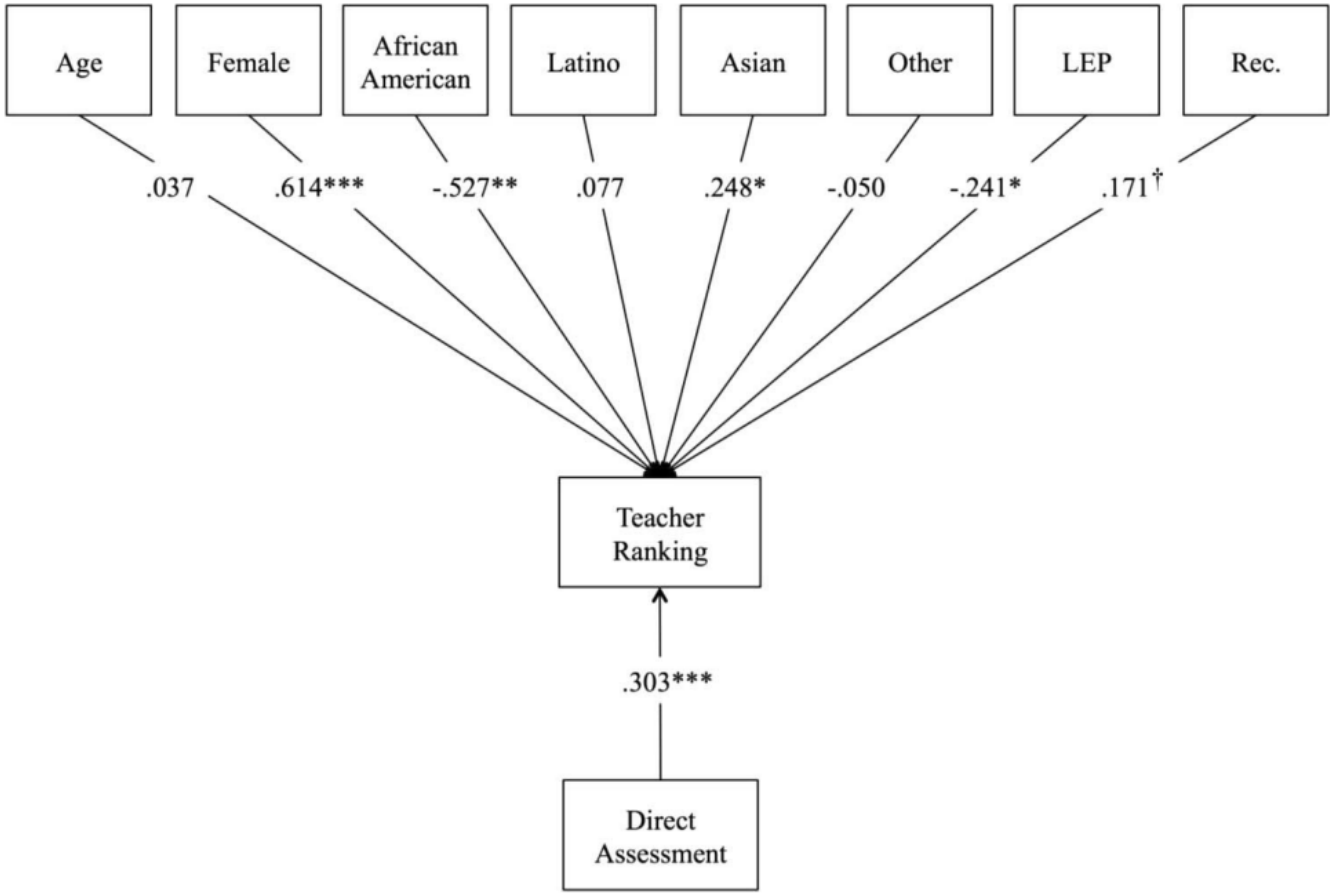


Figure 56: Longitudinally constrained, unstandardized path model results for Executive Functioning Skills

Source: Garcia et al. (2019)

	Unconstrained										Constrained		
	Fall teacher rank					Spring teacher rank							
Variable	*b*	*SE*	β	*SE*	*p*	*b*	*SE*	β	*SE*	*p*	*b*	*SE*	*p*
DA[a]	.35	.05	.31	.04	.000	.27	.05	.25	.05	.000	.30	.04	.000
Age, years[a]	.04	.04	.04	.03	.217	.03	.04	.03	.04	.438	.04	.03	.263
Female	.62	.08	.32	.04	.000	.61	.08	.31	.04	.000	.61	.07	.000
African American	−.41	.20	−.10	.05	.038	−.63	.16	−.15	.04	.000	−.53	.17	.002
Hispanic	.17	.13	.08	.06	.182	−.01	.10	−.01	.05	.896	.08	.11	.474
Asian	.25	.10	.12	.05	.011	.24	.11	.12	.05	.026	.25	.10	.011
Other	−.02	.16	−.01	.04	.882	−.07	.18	−.02	.04	.676	−.05	.15	.738
LEP	−.24	.12	−.11	.06	.039	−.24	.12	−.11	.06	.043	−.24	.11	.026
Reclassified	.16	.11	.06	.04	.141	.19	.12	.08	.05	.107	.17	.10	.080

Note. $N = 558$. DA = direct assessment; TR = teacher ranking; LEP = limited English proficient. Ethnicity coded relative to non-Hispanic White. English proficiency coded relative to English dominant.

[a] Variable was centered within-classroom cluster. Unconstrained coefficients differ between the fall and spring. Constrained coefficients are set to be equal in the fall and spring. Standard errors clustered at classroom level.

Figure 57: Path Analysis Results for Fall and Spring Teacher Executive Function Rankings by Fall and Spring Direct Assessments and Child Demographics

Source: Garcia et al. (2019)

SES Group	Digit Span (number correct)		Letter Working Memory (number correct)		Spatial Working Memory (span length)		Stroop (proportion correct)		Tower of London (milliseconds)		Verbal Fluency (total number of words)	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Low SES	8.02	2.16	6.86	1.35	5.26	1.69	0.88	0.16	5367.11	4029.36	79.10	24.75
High SES	8.83	2.36	7.30	0.90	5.48	1.67	0.89	0.17	8895.64	10041.47	87.41	24.26

Figure 58: Descriptive statistics for each EF measure by SES group

Source: Last et al. (2018)

Variable	Mean grit score (SD)	*t*	*P*	Effect size[a]
Gender				
Males (*n*=71)	3.95 (0.45)	1.70	0.09	0.31
Females (*n*=59)	4.08 (0.38)			
Race				
White (*n*=98)	4.05 (0.42)	1.99	0.05*	0.41
Nonwhite (*n*=32)	3.88 (0.41)			

Figure 59: Grit and participant characteristics in medical students

Source: Miller-Matero et al. (2018)

Table 4
Female. Intercorrelations (Spearman's) between passion, grit and mindset (growth) ($N = 80$).

	Passion	Grit	Mindset (growth)
Passion	1	**.382**[a]	**.299**[a]
Grit		1	**.356**[a]
Mindset (growth)			1

[a] Correlation is significant at the 0.01 level (2-tailed).

Table 5
Male. Intercorrelations (Spearman's) between passion, grit and mindset (growth) ($N = 66$).

	Passion	Grit	Mindset (growth)
Passion	1	**.500**[b]	**.260**[a]
Grit		1	**.215**
Mindset (growth)			1

[a] Correlation is significant at the 0.05 level (2-tailed).
[b] Correlation is significant at the 0.01 level (2-tailed).

Figure 60: Intercorrelations (Spearman's) between passion, grit and mindset (growth)

Source: Sigmundsson et al. (2020)

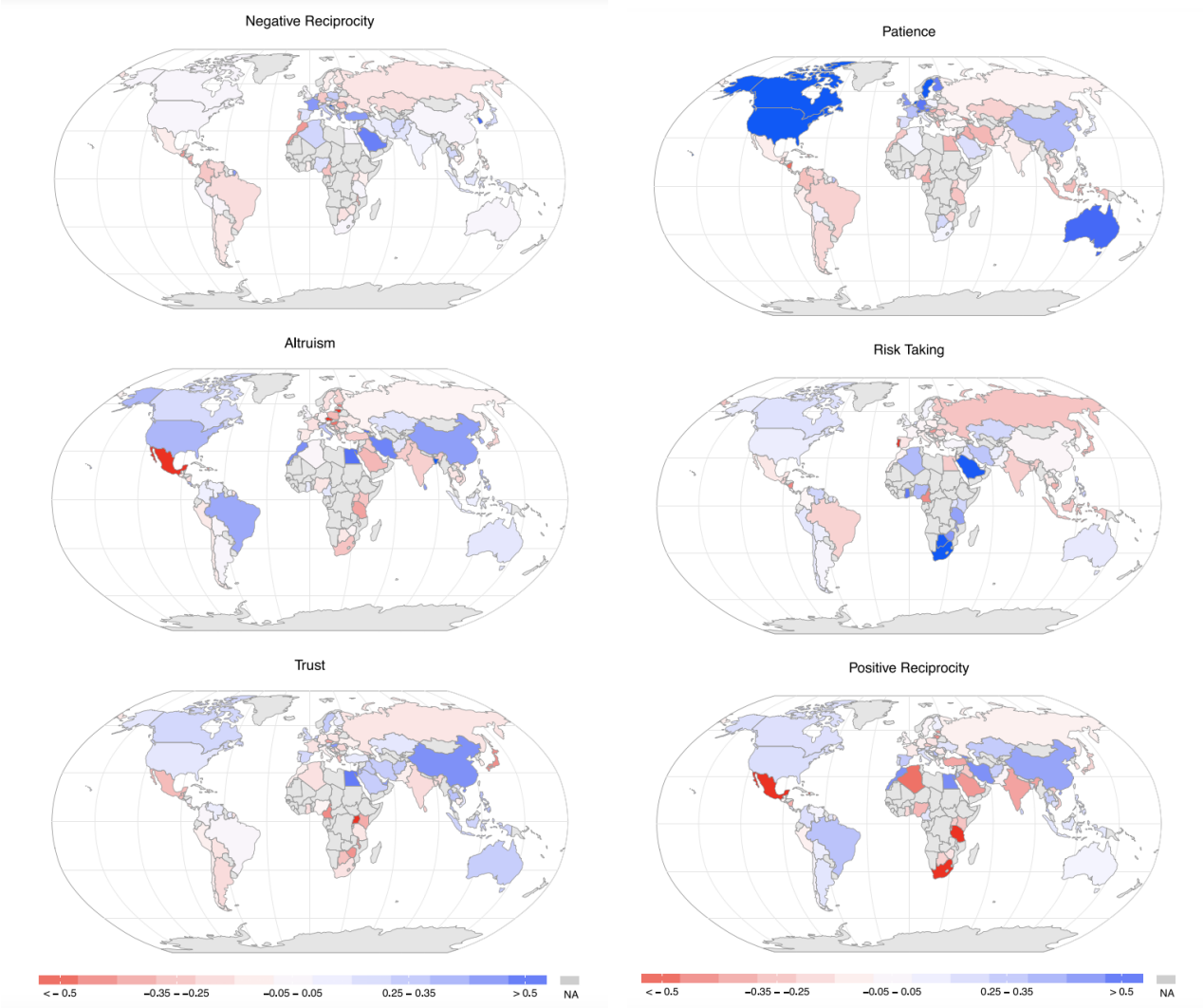


Figure 61: World Maps of Economic Preferences

Source: Falk et al. (2018)

REGIONAL AVERAGES AND VARIANCE DECOMPOSITION

	Patience	Risk taking	Pos. recip.	Neg. recip.	Altruism	Trust	# Obs.
Western Europe	0.49	−0.11	0.06	0.04	−0.04	0.10	11
Eastern Europe	−0.12	−0.12	−0.02	0.10	−0.22	−0.07	16
Neo-Europe	0.73	0.15	0.16	0.02	0.26	0.23	3
South and East Asia	−0.00	−0.10	0.07	0.11	0.13	0.04	13
North Africa and ME	−0.14	0.16	0.07	0.08	0.13	0.23	9
Sub-Saharan Africa	−0.16	0.34	−0.34	−0.11	−0.15	−0.33	11
South America	−0.21	−0.03	−0.08	−0.16	−0.05	−0.10	13
% between-country variation	13.5	9.0	12.0	7.0	12.3	8.2	

Notes. Neo-Europe includes the United States, Canada, and Australia. Regional averages of each preference are expressed in terms of standard deviations from the world individual mean. The variance decomposition in the bottom row decomposes the individual-level variation into the variance of the average preference across countries and the average of the within-country variance. Formally, the between-country variation corresponds to the R^2 of an OLS regression of all individual-level observations on a set of country dummies in which all observations are weighted by the sampling weights provided by Gallup to achieve (ex post) representativeness. ME = Middle East.

Figure 62: Regional Averages in Economic Preferences

Source: Falk et al. (2018)

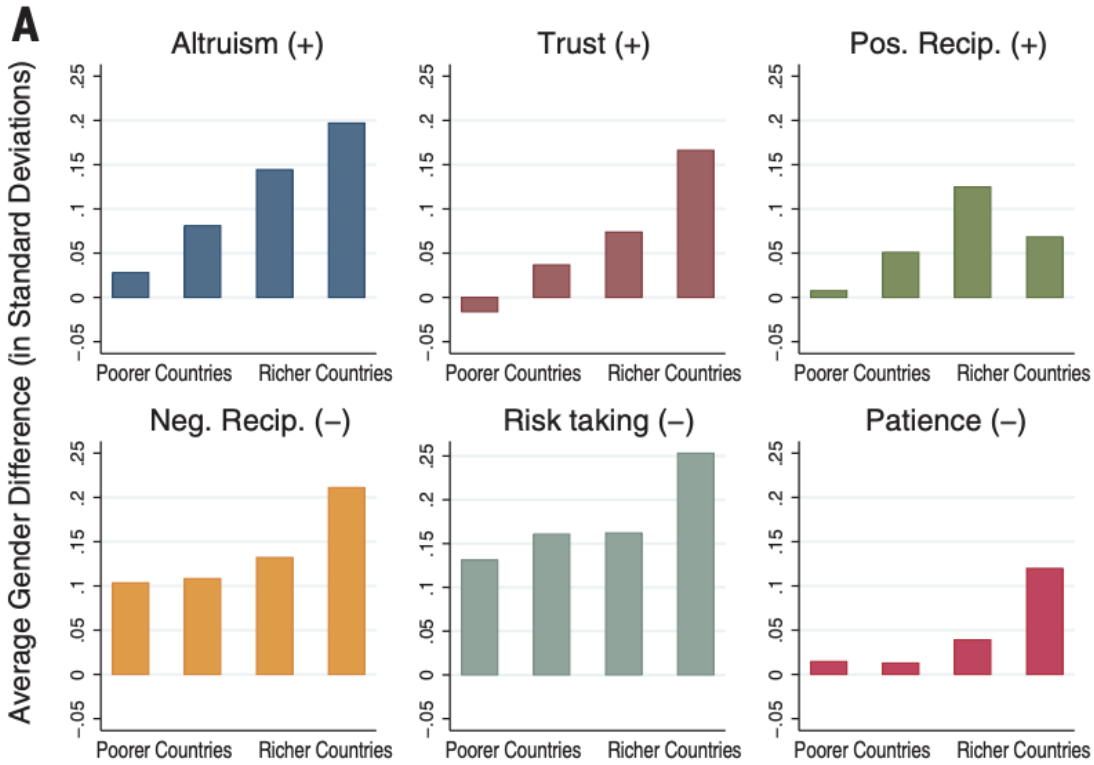


Figure 63: Analysis of gender differences in preferences in relation to economic development and gender equality

Source: Falk & Hermle (2018)

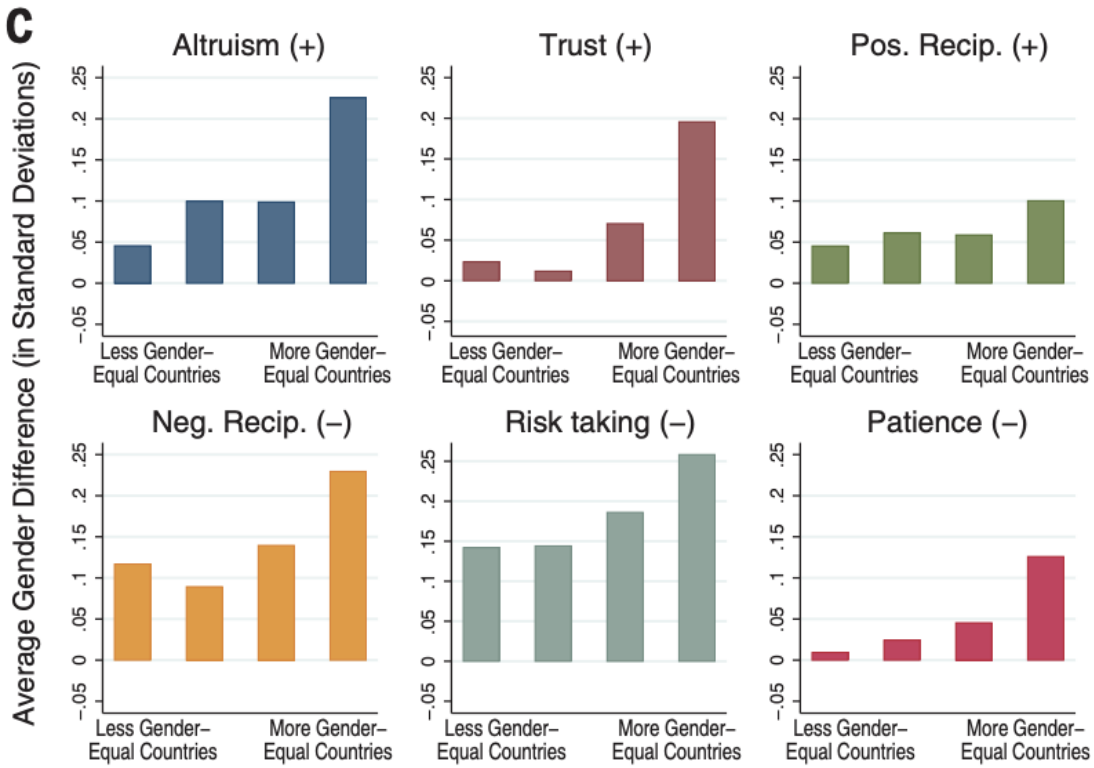


Figure 63: Analysis of gender differences in preferences in relation to economic development and gender equality (Cont'd)

Source: Falk & Hermle (2018)