Personality Psychology and Economics[1]

Mathilde Almlund
University of Chicago
Department of Economics

Angela Lee Duckworth
University of Pennsylvania

James Heckman
University of Chicago
University College Dublin
American Bar Foundation
Cowles Foundation, Yale University

Tim Kautz
University of Chicago
Department of Economics

[1] This research was supported by grants from NIH R01-HD054702, R01-HD065072, and K01-AG033182; the University of Chicago; A New Science of Virtue project sponsored by the John Templeton Foundation; the American Bar Foundation; a conference series from the Spencer Foundation; the JB & MK Pritzker Family Foundation; the Buffett Early Childhood Fund; and the Geary Institute, University College Dublin, Ireland. The opinions expressed in this report are those of the authors and do not necessarily reflect the views of the University of Chicago, the John Templeton Foundation, nor any of the other funders mentioned here. Amanda Agan and Pietro Biroli are major contributors to this essay through their surveys of the effect of personality on crime (presented in Web Appendix A7.B) and health (presented in Web Appendix A7.A), respectively. We are grateful to Pia Pinger for her analyses of the German Socio-Economic Panel (GSOEP) survey data. We have benefited from comments received from Pietro Biroli, Dan Black, Dan Cervone, Flavio Cunha, Kathleen Danna, Moshe Hoffman, John Eric Humphries, Miriam Gensowski, Dan McAdams, Lawrence Pervin, Pia Pinger, Armin Rick, and Brent Roberts. We also benefited from a workshop at the University of Illinois, Department of Psychology, on an early draft of this paper. Additional material that supplements the text is presented in a Web Appendix (http://jenni.uchicago.edu/personality_economics/).

Abstract

This paper explores the power of personality both as a predictor and as a cause of academic and economic success, health, and criminal activity. Personality is interpreted as a construct derived from an economic model of preferences, constraints, and information. Evidence is reviewed about the "situational specificity" of personality and preferences. An extreme version of that view claims that there are no stable personality traits or preference parameters that persons carry across different situations. It further claims that personality psychology has little relevance for economics. The biological and evolutionary origins of personality are explored. Personality measurement systems and relationships among the measures used by psychologists are examined. The predictive power of personality measures is compared with the predictive power of measures of cognition captured by IQ and achievement tests. For many outcomes, personality measures are just as predictive as cognitive measures, even after controlling for family background and cognition. Moreover, standard measures of cognition are heavily influenced by personality traits and incentives. Personality traits are positively correlated over the life cycle. However, they are not fixed and can be altered by experience and investment. Intervention studies, along with studies in biology and neuroscience, establish a causal basis for the observed effect of personality on economic and social outcomes. There is greater malleability of personality traits over the life cycle compared to cognition, which becomes highly rank stable around age 10. Interventions that change personality are promising avenues for addressing the problems of poverty and disadvantage.

JEL Codes: I2, J24

Key words: Personality, behavioral economics, cognitive traits, wages, economic success, human development, person-situation debate

Mathilde Almlund
University of Chicago
Department of Economics
1126 East 59th Street
Chicago, Illinois 60637
almlund@uchicago.edu

Angela Lee Duckworth
University of Pennsylvania
3701 Market St., Suite 215
Philadelphia, PA 19104
(215) 898-1339
duckworth@psych.upenn.edu

James Heckman
University of Chicago,
American Bar Foundation,
University College Dublin,
Cowles Foundation, Yale University
Department of Economics
1126 East 59th Street
Chicago, Illinois 60637
(773) 702-0634
jjh@uchicago.edu

Tim Kautz
University of Chicago
Department of Economics
1126 East 59th Street
Chicago, Illinois 60637
tkautz@uchicago.edu

Contents

1. Introduction

This paper examines the evidence on the predictive and causal power of personality for a variety of life outcomes. It develops an economic model of personality and shows how it can be used to interpret the evidence presented by personality psychologists. Mechanisms for changing personality are investigated.

It is well established that measures of intelligence and academic achievement predict a variety of social and economic outcomes.[2] Table 1 displays the correlations of three widely used measures of cognition recorded in the adolescent years—IQ, an achievement test (the Armed Forces Qualifying Test or AFQT), and report card grades (in tenth grade)—with a variety of adult labor market and social outcomes.[3,4]

All of the reported correlations are 0.36 or less and most are below 0.25. However, most are statistically significant. The AFQT is more highly correlated with outcomes than conventional IQ tests, suggesting that standardized achievement tests capture traits valued in economic and social life other than measured intelligence. The correlation of grades with outcomes is usually intermediate between IQ and AFQT.[5] None of the measures of cognition predicts a great deal of the variance in the listed outcomes—at most 14% and for most measures less than 7%—leaving a lot of room for the operation of other factors. Adjusting for family

[2] Cognitive traits include fluid intelligence, acquired skills and knowledge, processing speed, memory, etc. These are discussed in detail in Section 5. For evidence on their predictive power, see, for example, Herrnstein and Murray [1994], Gottfredson [2008], Cawley, Heckman and Vytlacil [1999], Heckman, Stixrud and Urzua [2006], Taubman and Wales [1973], Noyes [1945], Jencks, Smith, Acland et al. [1972], and Bowles, Gintis and Osborne [2001a].
[3] The AFQT consists of four subtests: word knowledge, paragraph comprehension, arithmetic reasoning, and mathematics knowledge (Roberts, Goff, Anjoul et al. [2000, p. 19]).
[4] Many interpret the AFQT as an IQ test. For discussion of the contrast between achievement and IQ tests see the collection of papers in Green [1974]. Many of the contributors to that book do not think any distinction is meaningful.
[5] Grades are not adjusted for schooling quality.

background, most correlations remain statistically significant and the predictive ordering of IQ, grades, and AFQT is unchanged. See Table A1 in the Web Appendix.[6]

Why do grades and achievement test scores predict adult outcomes better than IQ? We show that up to 35% of the variance in the scores on achievement tests can be explained by measures of personality.[7] This may explain the greater predictive power of AFQT than of IQ shown in Table 1. Grades are also associated with measures of personality which may explain their generally higher predictive validity than of IQ as revealed in Table 1, especially for the outcomes of women.[8,9] Another interpretation of this evidence is that acquired knowledge is more predictive than fluid intelligence as measured by IQ. We demonstrate the role of personality in promoting the acquisition of knowledge. Personality has both direct and indirect effects on many economic and social outcomes.

---

[6] http://jenni.uchicago.edu/personality_economics/. In that table, we report the partial correlations between the measures in Table 1 adjusting both variables for the effect of family backgrounds (i.e. we report the correlations between the residuals of the variables after removing the influence of family background variables).
[7] Borghans, Golsteyn, Heckman et al. [2010] and Duckworth, Quinn and Tsukayama [2010]. We discuss this evidence in Section 5.
[8] Bowen, Chingos and McPherson [2009b], Willingham, Pollack and Lewis [2002] and Duckworth and Seligman [2005].
[9] The predictive validity of grades would likely increase if data were available to condition on schooling quality and grading standards.

Table 1. Validities from the National Longitudinal Survey of Youth, 1979

	NLSY79 Correlation Table (tests and school performance)					
	Males			Females		
Outcomes	IQ	GPA (10th grade)	AFQT	IQ	GPA (10th grade)	AFQT
Hourly Wage Age 25	0.17***	0.17***	0.22***	0.14***	0.20***	0.25***
Hours Worked Age 25	0.08***	0.02	0.08***	0.14***	0.19***	0.28***
Wage Income Age 25	0.19***	0.17***	0.25***	0.21***	0.25***	0.36***
Weeks Worked Age 25	0.08***	0.04**	0.09***	0.16***	0.20***	0.30***
Weeks Unemployed Age 25	-0.14***	-0.11***	-0.18***	-0.12***	-0.11***	-0.12***
Weeks Out of Labor Force Age 25	-0.02	0.03	0.02	-0.11***	-0.15***	-0.26***
Total Jobs by Age 25	0.04	-0.08***	-0.04***	0.16***	0.03*	0.19***
Num. of Spouses/Partners by Age 25	-0.06**	-0.08***	-0.06***	0	-0.06***	-0.02
Any Welfare Age 25	-0.09***	-0.12***	-0.16***	-0.19***	-0.21***	-0.36***
Hourly Wage Age 35	0.03	0.05***	0.05***	0.11***	0.10***	0.13***
Hours Worked Age 35	0.10***	0.12***	0.21***	0.02	0.10***	0.17***
Wage Income Age 35	0.21***	0.21***	0.26***	0.08***	0.15***	0.19***
Weeks Worked Age 35	0.10***	0.15***	0.23***	0.11***	0.13***	0.23***
Weeks Unemployed Age 35	-0.10***	-0.11***	-0.15***	-0.17***	-0.11***	-0.14***
Weeks Out of Labor Force Age 35	-0.09**	-0.14***	-0.22***	-0.04	-0.11***	-0.18***
Total Jobs by Age 35	-0.02	-0.13***	-0.06***	0.09***	-0.02	0.18***
Num. of Spouses/Partners by Age 35	-0.05*	-0.10***	-0.05***	0.04	-0.05***	-0.01
Any Welfare Age 35	-0.09***	-0.11***	-0.23***	-0.20***	-0.23***	-0.36***

Notes: AFQT was administered in 1979. IQ is a percentile score obtained by equating IQ across different IQ tests from NLSY79 transcript data following the procedure in Borghans, Golsteyn, Heckman et al. [2010]. Tenth grade GPA is reported because after this grade attrition losses are substantial. (* p<0.10, ** p<0.05, *** p<0.01)
Source: National Longitudinal Survey of Youth 1979 (NLSY79). These estimates are taken from Heckman and Humphries [2010].

Associations are useful for predicting outcomes. Effective policy is based on causal relationships that establish if interventions work and how they work. This paper discusses causal evidence from a variety of interventions.

The causal power of traits other than those measured by IQ and achievement tests in determining life outcomes is demonstrated by an analysis of the Perry Preschool Program.[10] This experimental intervention enriched the early social and emotional environments of disadvantaged children ages 3 and 4 with subnormal IQs. It primarily focused on fostering the

[10] See Weikart, Epstein, Schweinhart et al. [1978], Schweinhart, Montie, Xiang et al. [2005]; Sylva [1997] and Heckman, Moon, Pinto et al. [2010a] for descriptions of the Perry program. We discuss this evidence further below.

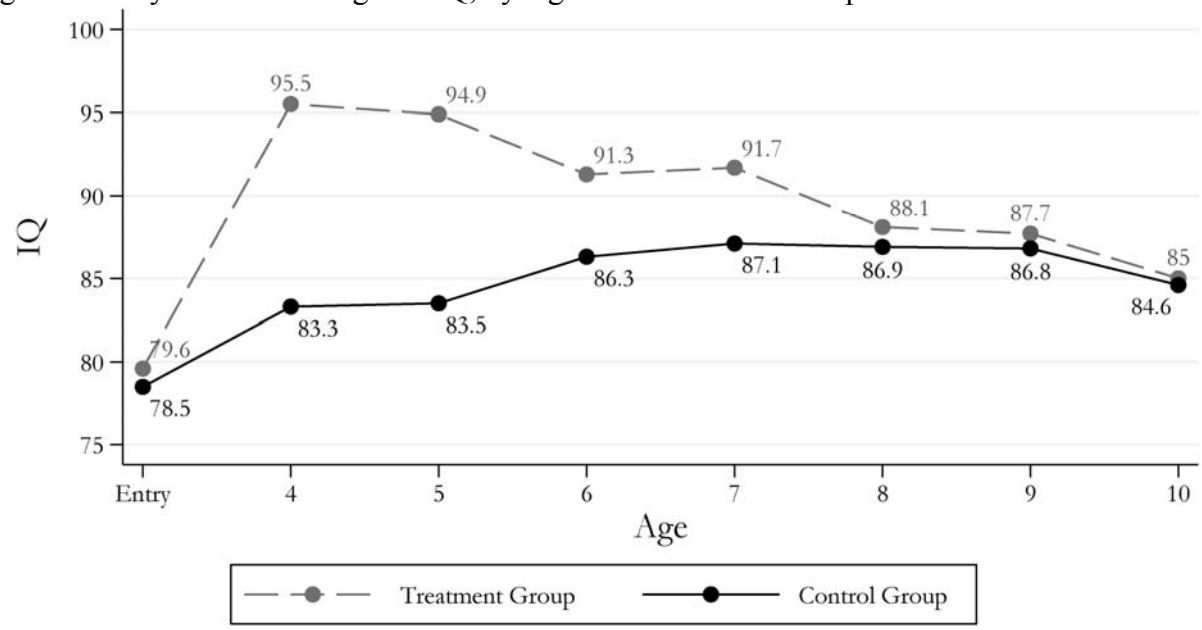
ability of participants to plan tasks, to execute their plans, and to review their work in social groups.[11] In addition, it taught reading and math skills, although this was not its focus. Both treatment and control group members were followed into their 40s.[12]

Figure 1 shows that, by age ten, treatment group mean IQs were the same as control group mean IQs. Many critics of early childhood programs seize on this and related evidence to dismiss the value of early intervention studies.[13] Yet on a variety of measures of socioeconomic achievement, over their life cycles, the treatment group was far more successful than the control group.[14] The annual rate of return to the Perry Program is in the range 6-10% for boys and girls separately.[15] These rates of return are statistically significant and above the returns to the US stock market over the post-war period.[16] Something important beyond IQ was changed by the intervention. Heckman, Malofeeva, Pinto et al. [2010] show that the personality traits of the participants were beneficially improved in a lasting way.[17]

---

[11] Sylva [1997] shows that the Perry Program has important features that are shared with programs designed to foster self-control in children, for example, Tools of the Mind (Bodrova and Leong [2001]).
[12] Plans are underway to follow the Perry sample through age 50.
[13] See the Westinghouse study of Head Start (Project Head Start [1969]).
[14] See Heckman, Malofeeva, Pinto et al. [2010], and Heckman, Moon, Pinto et al. [2010a].
[15] See Heckman, Moon, Pinto et al. [2010b].
[16] See DeLong and Magin [2009] for estimates of the return on equity.
[17] We discuss this evidence in Section 8. The traits changed were related to self-control and social behavior. Participants of both genders had better "externalizing behavior" while girls also had improved "internalizing behavior." See Heckman, Malofeeva, Pinto et al. [2010]. Duncan and Magnuson [2010] offer a different interpretation of the traits changed by the Perry experiment, but both analyses agree that it was not a boost in IQ that improved the life outcomes of Perry treatment group members.

Figure 1. Perry Preschool Program: IQ, by Age and Treatment Group



Notes: IQ measured on the Stanford-Binet Intelligence Scale (Terman and Merrill [1960]). Test was administered at program entry and each of the ages indicated.
Source: Cunha, Heckman, Lochner et al. [2006] and Heckman and Masterov [2007] based on data provided by the High Scope Foundation.

Economists, psychologists, and sociologists are now beginning to analyze the determinants of social and economic success beyond those captured by standard measures of cognitive skills.[18] Social policy designed to remediate deficits in achievement can be effective by operating outside of purely cognitive channels. There is a substantial imbalance in the scholarly and policy literatures in the emphasis placed on cognitive skills compared to other traits which this chapter aims to correct. Understanding these other factors deepens

[18] See Bowles, Gintis and Osborne [2001a] and Borghans, Duckworth, Heckman et al. [2008] for reviews of the literature in economics. Marxist economists and sociologists (e.g., Bowles and Gintis [1976] and Mueser [1979], respectively) pioneered the analysis of the impact of personality on earnings. Mueller and Plug [2006] estimate empirical relationships between personality traits and earnings, schooling and occupational attainment. Hartog [1980; [2001] relates the Big Five personality factors to earnings. van Praag [1985] draws on the psychology literature to analyze economic preferences. van Praag and van Weeren [1988] and Borghans, Duckworth, Heckman et al. [2008] link economics with psychology.

understanding of the sources of inequality and suggests new avenues of intervention to promote the well-being of disadvantaged populations.

We present evidence on the role of personality in predicting and *causing* a variety of life outcomes. We invest the mechanisms of causality. We relate definition of personality and personality traits by psychologists to economic models. We interpret personality as a *strategy* for responding to life situations. Personality traits, along with other influences, generate measured personality.

Both cognitive and personality traits are implicitly defined and measured by the performance of people on tasks or actions that are designated by *a priori* considerations to be "personality" or "cognitive" activities. The choice of which tasks or actions are used to define and measure these traits is delicate, and no hard and fast rules apply. Different analysts disagree about these choices. We discuss fundamental identification problems that arise in using measured behavior, tests, and observer reports to identify personality traits.

Many economists, especially behavioral economists, remain to be convinced of the predictive validity, stability, or causal status of economic preference parameters or personality traits, believing instead that behavior is almost entirely determined by the constraints and incentives in each situation. We review the history of the "person-situation" debate in psychology sparked in 1968 by Mischel's critique of the personality literature of that era. The extreme "situationist" point of view is often cited in behavioral economics and has helped delay the widespread use of findings from personality psychology in economics.[19] We suggest that most psychologists, including Mischel and his critics, accept the notion of a stable personality as

[19] For an example of this point of view see Thaler [2008].

specified in this chapter.[20] Measured personality exhibits both stability and variation across situations.[21]

Although personality is not merely a situation-driven ephemera, it is also not set in stone. We present evidence that both cognitive and personality traits evolve over the life cycle—but at different rates at different stages. Cognitive processing speed, for example, tends to rise sharply during childhood, peak in late adolescence, and then slowly decline. Crystallized intelligence—available knowledge—rises over most of the life cycle, then slowly declines in later years. Some personality traits, such as Conscientiousness, increase monotonically from childhood to late adulthood.[22] Rank-order stability for many personality measures peaks between the ages of 50 to 70, whereas IQ reaches these same levels of rank stability by middle childhood.[23] Recently developed economic models of parental and environmental investment in children help to explain the evolution of these traits.

This chapter addresses the following specific questions:

(1) *How can we fit the psychological construct of personality into an economic framework? Can conventional models of preferences in economics explain the main theories in personality psychology?*

(2) *What are the main measurement systems used in psychology for personality, and how are they validated? How are different systems related to each other? What is the relationship between measures of personality and measures of psychopathology and measures of child temperament?*

---

[20] See, e.g., Mischel and Shoda [1995; [2008].
[21] McAdams [2006, p. XVIII], Funder [2009], Mischel [2009], Roberts [2007; [2009], and Revelle, Wilt and Condon [2010] discuss the stability question.
[22] Roberts, Walton and Viechtbauer [2006], Roberts [2007], and Jackson, Bogg, Walton et al. [2009].
[23] Roberts and DelVecchio [2000].

(3) *What is the relationship between economic preference parameters and psychological measurements?*

(4) *How stable across situations and over the life cycle are preference parameters and personality traits?*

(5) *What is the evidence on the predictive power of cognitive and personality traits?*

(6) *What is the evidence on the causal power of personality on behavioral outcomes?*

(7) *Can personality be altered across the life cycle? Are interventions that change personality traits likely fruitful avenues for policy?*

(8) *Do the findings from psychology suggest that conventional economic theory should be enriched?*

The plan of the paper is as follows. Section 2 presents an initial definition of personality traits that is widely used in the literature on personality psychology. It also presents a brief history of personality psychology and the person-situation debate. Section 3 defines personality as a strategy response function using an economic model of preferences, expectations, and constraints. Our model distinguishes measured personality from personality traits. Our definition formalizes various definitions of personality used in the literature on personality psychology and facilitates the analysis of personality using the tools of economics. Many psychologists use measurements to *define* psychological traits.[24] Economists take a different approach by defining the traits abstractly and then seeking ways to identify them from data.

Section 4 discusses alternative criteria that psychologists use to define traits and examines the strengths and limitations of each criterion. It links our abstract definition to linear factor models that are widely used by psychologists to measure traits.

[24] Thus personality psychology is operationalist in the sense of Bridgman [1959].

Section 5 presents the main systems used to *measure* personality and cognition and discusses the relationship among the systems. We illustrate the nonidentification result discussed in Section 3 by showing how scores on IQ tests can be greatly affected by incentives and context. We present additional evidence showing that the scores on achievement tests depend on cognition and personality, with a substantial predictive role for personality.

Section 6 discusses economic preferences and relates economic preference parameters to psychological parameters. We examine the predictive validity of preference parameters.

Section 7 surveys the predictive validity of personality measures for education, crime, health, and labor market outcomes. The main text briefly summarizes a vast empirical literature. The Web Appendix presents a detailed survey of the literature relating cognition and personality in each of these areas of economic and social life.[25]

Section 8 presents evidence on the *causal* impact of personality on outcomes, as well as the evidence on the stability and malleability of personality and preferences. It discusses the evidence from intervention studies. Section 9 concludes with provisional answers to the eight questions.

[25] The web appendix can be found online at http://jenni.uchicago.edu/personality_economics/. Amanda Agan and Pietro Biroli are authors of some of these surveys as noted in the appendix.

2. Personality and Personality Traits: Definitions and a Brief History of Personality Psychology

Personality psychology addresses the whole person. It considers human universals, individual differences and how and why people are unique.[26] We find it helpful to distinguish **personality traits** from **measured personality**.

Personality traits as defined by one leading personality psychologist defines personality traits in the following way:

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

This definition, or closely related versions, are used throughout personality psychology.[27]

Roberts' definition of personality traits refers to the stability of certain patterns of behavior—actions or responses to situations that people take. His definition includes patterns of thoughts or feelings. Perceptions and expectations of future events may shape behavior, feelings and thoughts. These are products of cognitive activities that help to determine measured personality. Many, if not all, personality traits specify tendencies to perceive or process information in particular ways.

There are many different models of personality. Roberts [2006] presents the schematic displayed in Figure 2 to relate personality traits to behavior.[29] He distinguishes mental abilities

---

[26] Cervone and Pervin [2009] provide a clear introduction to the literature on personality psychology.

[27] However, some personality psychologists use this or a very similar definition to define **personality**. Thus Cervone and Pervin [2009] define personality as

> *"psychological qualities that contribute to an individual's enduring and distinctive patterns of thinking, feeling and behaving"* (p. 8).

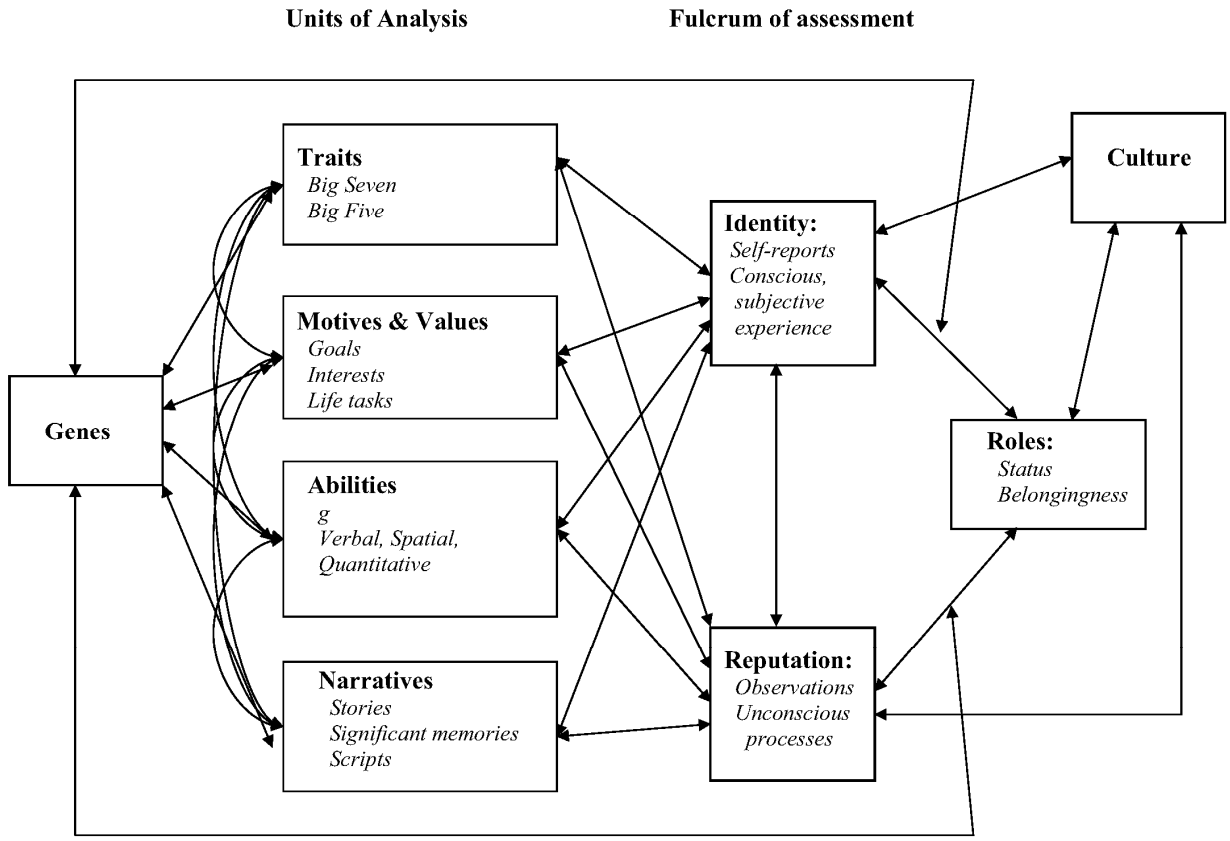
Another definition in a graduate text on personality by McAdams emphasizes context more strongly:

> *"Personality is a patterning of dispositional traits, characteristic adaptations, and integrative life stories set in culture and shaped by human nature"* McAdams [2006].

from personality traits (the items in the boxes will be discussed in later sections). These along with preferences (motives and values) and narratives (the stories people tell themselves in organizing their lives and making meanings of them) shape one's identity and reputation (the views of the person by others). These in turn shape the roles of individuals in the economy and the society and the larger culture to which they belong. In Roberts' vision of personality, feedback processes operate among all components of the individual. So his conception of personality is broad enough to include the possibility that identity shapes traits and abilities, perhaps through a mechanism such as epigenetics, in which environment affects gene expression.[30]

[29] Graphical models like Figure 2 are the rule in personality psychology. Explicit formal models are rare. Section 3 presents a formal model.
[30] See, e.g., Rutter [2006a].

Figure 2. Roberts's Model of Personality Psychology



Source: Roberts [2006].

In Section 3, we formalize aspects of Roberts' model within a standard economic framework of production, choice, and information. Figure 2 presages our discussion of a central identification problem discussed in Section 3. Measurements and behaviors that arise from responses to incentives and interactions with the culture are used to infer traits. It is the "traits" along with the other "units of analysis" in Figure 2 that produce the observed behaviors and measurements that are used to infer the traits. To infer traits from behaviors requires that one parse out all other contributing factors that produce the behavior—a challenging task that requires detailed measurements on other influences of behavior besides the traits. The inability to parse and localize behaviors that depend on a single trait produces a fundamental

identification problem discussed in depth in Section 3. Behavior changes in response to incentives. One has to standardize for the incentive effects of situations to accurately define personality traits. We discuss this topic in depth in Section 3.

Using the language of game theory, we define **measured personality** as a *strategy* for coping with life under different situations. Cognitive performance is just one of many possible types of responses to situations and is an aspect of personality. The observed behavior is generated by the traits, but more than the traits generate observed behavior.[32]

2.A. A Brief History of Personality Psychology

Interest in how individuals can differ radically from one another in their reaction to a common situation is as old as human history. The importance of personality for predicting educational outcomes was recognized by the creators of the first IQ tests. Alfred Binet, architect of the first modern intelligence test that became the Stanford-Binet IQ test, noted that performance in school

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

Lewis Terman, the psychologist who created the Stanford-Binet test in its modern form, wrote along similar lines. Comparing more successful high-IQ people to less successful ones, he wrote contrasting the A's (the high achievers) with the C's (the low achievers):

> *"The subjects, their wives, and their parents showed remarkable agreement in rating the A's far higher than the C's on Perseverance, Self-confidence, and Integration*

[32] Some psychologists deny that any traits exist. See Borsboom, Mellenbergh and Van Heerden [2003].

toward goals." (Terman and Oden [1947, p. 351])

David Wechsler [1943], who helped usher intelligence testing into widespread practice, made a similar observation about the unfortunate neglect of "non-intellective" factors that, in conjunction with general intelligence, determine intelligent behavior.

At about the same time that Binet was writing, Charles Spearman, best known for his work on "*g*"—a unitary factor that is claimed to capture the structure of intelligence—along with his student, Edward Webb, undertook studies of "character" because of "the urgency of its practical application to all the business of life" (Webb [1915, p. 1]). Spearman and Webb concluded that many positive aspects of character shared a relation to what modern personality psychologists term "Conscientiousness."[33] This general factor, which Spearman and Webb chose to call "persistence of motives," meaning "consistency of action resulting from deliberate volition, or will," was distinct from a general intelligence factor (Webb [1915, p. 60]).

Arthur Jensen, an intellectual heir of Spearman, writes

"What are the chief personality traits which, interacting with g, relate to individual differences in achievement and vocational success? The most universal personality trait is conscientiousness, that is, being responsible, dependable, caring, organized and persistent" Jensen [1998, p. 575].

One reason why traits related to Conscientiousness are so important to academic success is that, according to William James [1899], in "schoolroom work" there is inevitably "a large mass of material that must be dull and unexciting." In a series of essays entitled *Talks to Teachers*, James observed:

[33] Here and elsewhere through this essay, we capitalize personality traits.

> *"There is unquestionably a great native variety among individuals in the type of their attention. Some of us are naturally scatter-brained, and others follow easily a train of connected thoughts without temptation to swerve aside to other subjects."* (James [1899, p. 112])

James notes that while classroom teachers should do their utmost to engage students in learning, a dispositional advantage in the capacity for sustaining attention in spite of diversions and distractions puts some students at a tremendous advantage. The importance of Conscientiousness for predicting success both inside and outside of the classroom is a recurrent finding which we summarize in Section 7.

Progress in Personality Psychology

Over the past century, interest in personality among psychologists has fluctuated dramatically. During the first half of the twentieth century, many of the most prominent psychologists (e.g., Gordon Allport, Raymond Cattell, Hans Eysenck, Charles Spearman, Lewis Terman) were vigorously engaged in the study of individual differences, in personality traits as well as in intelligence, interests, and motivation. In 1968, Walter Mischel published a monograph entitled *Personality and Assessment*, challenging the most important theoretical assumptions and empirical findings of the field. An acrimonious "person-situation" debate ensued, which pitted those who favored situational factors as explaining behavior against those who considered person variables as more consequential. During this time, considered by many to be a fallow period in the history of personality psychology, the general Zeitgeist favored experimental social psychology research which focused on the importance of the situation compared to the individual traits featured in personality psychology. Arguably, the past three decades have witnessed a

revival of interest in personality, though it would be an overstatement to say that personality psychology is as fashionable a discipline as it was a century ago.[34]

A more systematic approach to the study of personality conceived, in particular by psychologists in the lexical tradition who believed that the most important dimensions on which human beings differed would be captured in natural language. These personality pioneers extracted words from the (English) dictionary that characterized individual differences between people (e.g., irritable, proud) and, after eliminating synonyms and non-trait words (e.g., laughing), administered these trait inventories to large samples and applied the same statistical methods developed by Galton, Spearman, Binet, Pearson, Cattell, and Thorndike to identify the structure of cognitive abilities.[35]

The fruits of several decades of research in this tradition beginning in the 1970s have produced a widely (but not universally) shared consensual taxonomy of traits, known as the Five-Factor model or, more colloquially, as the Big Five, that is arrived at through factor analysis of observer and self-reports of behaviors. The Big Five posits a hierarchical organization for personality traits, with five factors at the highest level and progressively more narrowly defined traits (or facets) at lower and lower levels. The Big Five factors are Openness to Experience (also called Intellect or Culture), Conscientiousness, Extraversion, Agreeableness, and Neuroticism (also called Emotional Stability).[36] The Big Five factors represent personality at the broadest level of abstraction. They summarize a large number of distinct, more specific, personality facets. See Table 2 and the extended definition in Section 5.

[34] See Revelle, Wilt and Condon [2010] for an informative history of personality psychology.
[35] Barenbaum and Winter [2008], John and Srivastava [1999a], Krueger and Johnson [2008].
[36] The acronym OCEAN is sometimes used to summarize these traits.

Table 2. The Big Five Traits

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].

The Big Five are defined without reference to any context (i.e., situation). John [1990] and Costa and McCrae [1992a] present evidence that most of the variables used to assess personality in academic research in the field of personality psychology can be mapped into one or more of the dimensions of the Big Five. They argue that the Big Five are the longitude and latitude of personality, by which all more narrowly defined traits may be categorized (see also Costa and McCrae [1992a]). We discuss the Big Five further in Section 5, where we also consider alternative measurement systems.

Renewed interest in personality belies the disinterest and even suspicion with which personality research is regarded by social psychologists. To understand the origins of this stance, we must return to Mischel's [1968] monograph. Among other observations, Mischel commented that correlations between behavioral task measures of personality and questionnaire measures seldom if ever exceeded .30.[37,38] The implication of such within-individual behavioral heterogeneity suggested to Mischel that "the behaviors which are often construed as stable personality trait indicators are highly specific and depend on the details of the evoking situations and the response mode employed to measure them" (p. 37). Mischel's most emphatic statement at that time was

> "*…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, p. 146]

Mischel suggested that global (i.e., domain-general) traits (e.g., "impulsive", "confident") measured in one situation did not predict future behavior and outcomes in other situations. In sum, Mischel's view was that global traits, in attempting to summarize behavioral dispositions without regard to situational contingencies, were "excessively crude, gross units to encompass adequately the extraordinary complexity and subtlety of the discriminations that people constantly make" (p. 301).

More recently, Mischel [2009] writes that the loudest voices in the person-situation debate were those of extremists arguing, quite unhelpfully, either that personality traits dictated

---

[37] There is some irony in the fact that none of the correlations of cognitive measures with outcomes that are reported in Table 1 are as high as .3, but no one questions the power of cognition in social life.

[38] Psychologists often work with standardized variables (variables normalized by standard deviations). They report correlations between standardized variables as "effect sizes."

behavior or that people completely conformed to situation influences. In his own defense, Mischel has argued that his message was much more nuanced:

> *"The main message of my 1968 monograph was that the situation has to be incorporated into the conception and assessment of personality"* Mischel [2004]

In particular, Mischel suggests that there are consistencies in behavior across time, but that the locus of consistency is to be found in highly contextualized if-situation/then-behavior contingencies (e.g., "*If* I feel threatened, *then* I am aggressive"). Variance across situations was, in Mischel's view, improperly treated by most personality psychologists as "error."[39] Indeed, in Mischel's [2004] view, the systematic variation of behavior across situations pointed to underlying motivations, beliefs, schemas, strategies, and other factors that collectively and interactively give rise to coherence in any individual's personality.

We formalize the "if-then" relationship within a standard economic model in Section 3. We establish that the person-situation debate boils down to an empirical question about the relative importance of person, situation, and their interaction in explaining behaviors. Although Mischel may have intended otherwise, proponents of the situationist view have used his monograph as ammunition in the battle against personality psychology. As is characteristic of many heated debates in social science, this one occurred in the absence of much data. We discuss the body of evidence that has emerged over the past four decades on the existence of stable personality traits in Section 5.

The critiques of personality psychology were so polemical in part because they were based on weak data and were rooted in deeper political or ideological views. Personality studies

[39] I.e. unobserved heterogeneity.

were criticized as being "elitist, racist, or exclusionary."[40] Behaviorism, a theory that posited, among other things, that all aspects of behavior can be explained by experience, was still influential, and the notion that situation and experience were all powerful—that people were born as blank slates—was very appealing.[41] Ross and Nisbett [1991] summarize the position of many social psychologists:

> *"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" (p. xiv).*

A similar view is held by many behavioral economists. Richard Thaler, in a recent round table discussion, noted that

> *"The great contribution to psychology by Walter Mischel [...] is to show that there is no such thing as a stable personality trait"* (Thaler [2008]).[42]

Many concepts in behavioral economics are attempts to explain why people behave seemingly inconsistently across situations, in violation of standard assumptions in economics. For instance, several studies find very low correlations in risk behavior across situations (e.g., Slovic [1962], Kogan and Wallach [1967], Slovic [1972], Blais and Weber [2006], Johnson, Wilke and Weber [2004], and Weber, Blais and Betz [2002]). Slovic, a psychologist who has had a strong influence on behavioral economics, writes:

> *"Although knowledge of the dynamics of risk taking is still limited, there is one important aspect that has been fairly well researched—that dealing with the stability of a person's characteristic risk-taking preferences as he moves from situation to situation. Typically,*

[40] Revelle, Wilt and Condon [2010, p. 9].
[41] Pinker [2002].
[42] Thaler [2008].

> *a subject is tested in a variety of risk-taking tasks involving problem solving, athletic, social, vocational, and pure gambling situations. The results of close to a dozen such studies indicate little correlation, from one setting to another, in a person's preferred level of risk taking."* (Slovic [1972, p. 795])

Prospect theory proposed by Kahneman and Tversky [1979] is one attempt to reconcile results on cross-situational inconsistencies in risk behavior. The study of context effects such as framing, time inconsistency, and heuristics seemed to suggest that there are no stable underlying preferences present across all situations (Schoemaker [1982], Hershey, Kunreuther and Schoemaker [1982]).[43]

Since one strand of behavioral economics stresses cross-situational inconsistencies, it is, as is economics in general, largely silent about personality traits. It was natural to extrapolate and connect the finding of inconsistency with standard economic theory to the predominant view in psychology at the time behavioral economics was being created—the inconsistency in behavior across situations that was perceived to have been established in social psychology. A specimen of this mindset is found in Thaler and Ziemba [1988] who write:

> *"The more basic question is whether individuals display a consistent 'trait' that can be captured in an index of risk aversion or risk seeking. Psychologists have found that most such traits are highly context specific, and risk taking is no exception."* (Thaler and Ziemba [1988, p. 170, Footnote 12]).

While the approach used by behavioral economists may be valid given the focus on inconsistencies within a person, the strong focus on situational differences may have falsely led

[43] However, prospect theory does not explain why individuals in the same situation, framed the same way, differ in their reactions.

some to conclude that interpersonal differences are irrelevant, and that personality traits have very low predictability.

Revival

Recent interest in personality traits by psychologists has been partly motivated by advances in behavioral genetics and neuroscience. Behavioral genetic studies documenting substantial heritability of personality traits (often measured by the global trait measures such as the Big Five) provide evidence that enduring individual differences in personality exist.[44] Most twin studies report that genetic factors explain roughly one half of the variation in each of the Big Five domains across people; shared environment, including parenting, explains a trivial amount; and unshared environment explains half (Bouchard and Loehlin [2001]). Heritability estimates are higher among low-SES individuals and lower among high-SES individuals (Turkheimer, Haley, Waldron et al. [2003]).

Behavioral genetics studies using alternative paradigms reach slightly different conclusions. Studies comparing the personality of adopted children to the personality of biological children have found that heritability plays a smaller role than studies that use cohabitating twins (Krueger and Johnson [2008]). One study that used personality of twins as visually assessed by professionals found a much bigger role for shared environment: across the Big Five, the median explained variation due to shared environment was 26% (Borkenau, Riemann, Angleitner et al. [2001]).[45]

---

[44] Separating environmental and genetic influences on personality poses a fundamental identification problem. Behavioral genetics studies address the problem by comparing the traits of monozygotic (i.e., identical) twins who share the same genes to dizygotic (i.e., fraternal) twins who share only half of their genes. See Goldberger [1979].

[45] Genome-wide association studies (GWAS) have moved beyond using genetic variation within families to using identifiable differences in specific genes. Typically, these studies examine how small differences in DNA sequences

Heritability studies might be confounded by gene-environment interactions. Some analysts have argued that choice of environment and perception of environment are shaped by genetics (Rowe [1981; [1983]). Others have found that the environment affects genetic expression (epigenetic effects). For example, a particular gene is associated with antisocial behavior, but only in combination with parental maltreatment (Caspi, McClay, Moffitt et al. [2002]). In sum, the recent evidence on heritability suggests that people do have genetic predispositions towards certain personality traits, but that environment also plays an important role in shaping measured personality, both directly and indirectly through interactions with genes.[46]

Neuroscience has also renewed interest in personality by demonstrating which parts of the brain are associated with personality traits. The evidence comes in two forms. First, brain lesions and other region-specific brain damage allow researchers to identify the function of parts of the brain. Perhaps the most famous example is Phineas Gage, a construction foreman who survived an accident in which a railroad spike passed completely through his prefrontal cortex. While he maintained his cognitive functioning, his personality took a turn for the worse – he transformed from being polite and dependable to profane and irresponsible. Re-analysis of his skull suggests that the spike pierced regions associated with social functioning (Damasio [1994]). Phelps [2006] reviews evidence from patients with brain damage that suggests emotions

– known as single-nucleotide polymorphisms (SNPs) – are associated with outcomes. For example, one study finds that variations of a gene related to serotonin are associated with self-reported Neuroticism (Lesch, Bengel, Heils et al. [1996]). However, individual SNPs explain only a small amount of the variation in personality – one study found that no differences in any single gene or DNA sequence explained more than 1% of the variance in the heritability of Neuroticism (Shifman, Johannesson, Bronstein et al. [2008]). Therefore, it is likely that the heritability of these personality traits reflects many genetic differences, hence broad outcomes, such as Neuroticism, might be too coarse to have a close association with any single SNP.

[46] The recent research on gene-environment interactions also calls into question the evidence on heritability produced by traditional methods. See, e.g., Turkheimer, Haley, Waldron et al. [2003].

associated with personality traits are involved in learning, attention, and other aspects of cognition and that these relationships have a biological basis.

Second, the development of functional Magnetic Resonance Imaging (fMRI), a method that monitors blood flow in the brain, provides another way to link brain structure and function to personality. For example, Canli [2004] studies which regions of the brain are activated when people view pictures associated with negative and positive emotions. People who respond to positive images tend to be more extraverted, whereas those who responded more to negative images tended to be more neurotic. Similarly, DeYoung, Hirsh, Shane et al. [2010] find that the volumes of different brain regions systematically covary with all of the Big Five traits except Openness to Experience. For example, Conscientiousness was associated with a region related to planning and control over behavior. However, other parts of the brain are correlated with self-reported personality, suggesting that broad measures of personality might not be specific to a single portion of the brain.[47] See Canli [2008] for a review of how molecular biology and neuroscience have advanced understanding of personality.

Recent studies have connected biology to economic decision-making. Kosfeld, Heinrichs, Zak et al. [2005] find that people who are given nasal sprays of oxytocin exhibit more trust in a game-theoretic trust game. Reuter, Montag, Altmann et al. [2009] find that people with a particular gene variant respond to oxytocin, suggesting that trust has a genetic basis. Similarly, Figner, Knoch, Johnson et al. [2010] find that magnetic disruption of the left lateral prefrontal cortex increases discount rates, providing evidence that individual differences in brain structure and function can cause differences in economic decision-making across people.

---

[47] Many theories of personality are cognitively-oriented. For example, Mischel and Shoda [1995], Bandura [1999] and Revelle, Wilt and Condon [2010] suggest that behavior is driven by cognitive operations, beliefs, and representations of reality (how people process information, what they believe to be true, and how they interpret their perceptions).

Evolutionary science provides another impetus for the revival of personality theory. In general, natural selection promotes characteristics that better propagate genes. For example, jealousy may promote mate retention by arising when a mate is not faithful (Buss [2000]). More mysterious is why evolution leads to differences in people. Evolution can lead to heterogeneity in traits in at least three ways. First, some mutations might create traits that do not strongly affect genetic propagation, one way or the other. If heritable, these traits can persist in the population. This is like "white noise" (Buss [2000]). Similarly, single-period mutations could contribute to cross-sectional heterogeneity in traits.

Second, heterogeneity can arise from "balancing selection," evolutionary forces that actively promote heterogeneity. Different traits might be more useful under different environmental conditions. For example, some psychologists argue that people with higher levels of Extraversion could have thrived during relatively safe times, but could have suffered during hazardous ones (Nettle [2006]). Given that there is some variability in environmental conditions, differences in traits could have persisted over time. Third, different evolutionary strategies might be successful in the same environment.[48] Evolutionary theory helps guide thinking about why genetic mechanisms give rise to traits, but provides little hard evidence about the extent to which traits are heritable.

In summary, behavioral genetics, evolutionary theory, and neuroscience support the existence of personality traits. Personality psychologists have made parallel progress in demonstrating the predictive validity of personality for consequential, objectively measured life outcomes. In section 7, we summarize the evidence on the power of personality to predict academic, social, health, and economic outcomes.

[48] See Mealey [1995].

3. An Economic Framework for Conceptualizing Personality and Personality Traits

This section uses a series of progressively more comprehensive models to integrate concepts from personality psychology into economics.[49] Roberts' framework (Figure 2) captures the essence of a variety of models of personality psychology. It is useful to relate it, when possible, to standard models in economics. Psychology adds new and often more nuanced descriptions of human behavior.

In the nineteenth century, economics and psychology were closely aligned. Economists often spoke of the "hedonic calculus" used by people weighing choices.[50] One of the advances made in neoclassical economics in the first half of the twentieth century was to focus on choices and the objective (easily measured) factors (like prices and incomes) that determine choices. Revealed preference became a central tool of economics. It was implemented using the marginal rate of substitution between choices—a key parameter that emerged from the neoclassical revolution.[51] This parameter did not require measurable utility or knowledge of the mental states of the agent making choices. Mental states and measurability of utility, once the province of economists, were eliminated by Occam's Razor.

Measurable utility is used in utilitarian economics but fell out of favor (see Samuelson [1956], Foster and Sen [1997]). Preferences that fulfilled criteria for rationality were consistent with utility functions that were determined up to monotonic transformations. Measurable utility returned in a specific fashion with analyses of decision-making under uncertainty (see Savage

[49] We present additional frameworks for incorporating personality into economics in Section A3 of the Web Appendix.
[50] See, e.g., Schumpeter [1954].
[51] See Hicks [1946].

[1954]). Mental states are viewed as unnecessary baggage by most economists except insofar as they affect choices. Thus, the traits, abilities, and narratives discussed by Roberts are studied only if they affect choices by serving as constraints and by affecting information processing. Motives and values are captured in part by the economic preference parameters discussed in Section 6. Until recently, "happiness," and "aggregate utility," as well as other subjective mental states that do not affect behavior (choices) were off limits to most economists.[52]

The most direct way to introduce psychological variables into economic models is through constraints. Psychologists do not explicitly model constraints as they affect choice.[53] Thus IQ, achievement tests and personality variables affect earnings because they are productive traits (see, e.g., Bowles, Gintis and Osborne [2001a]), and, up to a point, more of a trait can generate more resources. A second way to introduce such variables is through preferences. A third way is through expectations. Many empirical studies in economics that are surveyed in Sections 6 and 7 relate outcomes to personality measures and appeal to preferences, expectations, or constraints. However, this approach lacks specificity. In this section, we consider specific classes of models which articulate more precise roles for the psychological variables.

A more specific approach based on the Roy model [1951] of comparative advantage is widely used in the empirical literature. In the Roy model, people purse their comparative advantage, Heckman, Stixrud and Urzua [2006] use the Roy model to introduce psychological variables into the study of social and economic outcomes.[54] Personality traits are treated as

[52] See, however, the revival of utility measurement in the happiness literature (see Layard [2005]). Perceptions on which one does not act, included in the domain of psychology, have recently entered economic studies through the happiness literature.
[53] This is done in Borghans, Duckworth, Heckman et al. [2008].
[54] See Roy [1951], Heckman and Sedlacek [1985], and Heckman and Honoré [1990].

endowments, choices are determined by personality traits, and other factors as they affect productivity in skills. Agents can perform one of J tasks with productivity P_j, $j \in \{1,\ldots,J\}$. The productivity in task *j* depends on the traits of agents θ, and the "effort" they expend on the task, e_j:

(1) $$P_j = \phi_j(\theta, e_j), \quad j \in \{1,\ldots,J\}.$$

The traits are the endowments of agents that govern behavior. Examples of traits include height, personality traits, problem solving IQ, and strength. θ is a public good as it is available in the same amount for all tasks. Productivity also depends on effort e_j. Effort is assumed to be divisible and fixed in supply. Effort and traits are often assumed to be measure so that over the relevant range

$$\frac{\partial P_j}{\partial e_j} \geq 0 \quad \text{and} \quad \frac{\partial P_j}{\partial \theta} \geq 0,$$

but neither condition is strictly required. Excess effort (over-exertion; too much attention to detail) may be counterproductive so that function ϕ_j may not be monotonic in e_j, contrary to what is assumed here. Indeed, as discussed in Section 5, certain psychopathologies are associated with extreme levels of traits that are quite productive at normal levels.[55] Different traits may have different productivities in different tasks, leading to comparative advantage in different tasks for people with different endowments.

Effort may complement capability $\left(\frac{\partial^2 P_j}{\partial e_j \partial \theta} > 0\right)$ or may substitute for it $\left(\frac{\partial^2 P_j}{\partial e_j \partial \theta} < 0\right)$.

Some people may solve complex math problems with no effort, while others may have to

[55] Cattan [2010] shows that sociability has negative returns in some sectors but positive returns in other sectors.

allocate considerable time and effort to achieve the same result. Effort can be a vector (time, mental energy, attention), and it is assumed to be a divisible private good with the feature that the more that is applied to task *j*, the less is available for all other tasks at any point in time. $\sum_{j=1}^{J} e_j = \bar{e}$, where $\bar{e}$ is the endowment of total effort. Baumeister, Bratslavsky, Muraven et al. [1998] interpret self-control as a component of e that is fixed over given time periods. A person who exerts more self-control in one task may be less self-controlled in another task.

Let R_j be the reward per unit productivity in task j. In the first case we analyze (case I), it is possible to productively engage in only one of the J tasks at any time. This can be interpreted as a case where effort can only be applied to one task. The agent faces the problem of picking $\hat{j}$ where

(2) $$\hat{j} = \arg\max_{j \in \{1,\ldots,J\}} \left\{ R_j P_j \left(\theta, \bar{e}\right) \right\}.$$

In this case, θ and $\bar{e}$ play the same role. People with different effort and capability endowments will generally choose different tasks.[56] See Heckman, Stixrud and Urzua [2006], Cattan [2010], and the large literature on comparative advantage in economic and social life that those papers cite that show the sorting of persons in tasks by virtue of their traits. Heckman, Stixrud and Urzua [2006] show how persons with different endowments of personality and intelligence sort into different occupations and levels of schooling. People low in certain traits may have better endowments of effort and may compensate by exerting effort. For certain tasks (e.g., creating

[56] A straightforward extension works with utilities and not rewards so $U\left(P_1,\ldots,P_J\right)$ and the agents picks the j that maximizes utility, with the other arguments zeroed out. Formally, define $d_{P_j} = 1$ if a person chooses to perform task j. Array the d_{P_j} into a vector d_P. Array the P_j into a vector P. Realized utility is thus $U\left(d_P \odot P\right)$ where $\odot$ is a Hadamard (component-wise) product, i.e. a product of two vectors of the same length where the operation is such that the result is the product of the first element of one vector with the first element of the second vector and so forth for each component.

new branches of mathematics), there may be threshold levels of θ such that for $\theta < \overline{\theta}_j$, $P_j\left(\theta, \overline{e}\right) = 0$ no matter what the level of $\overline{e}$. (The person needs a given level of capability no matter how hard they try.) The higher R_j, the more likely will the person select to perform task j. The particular choice of which j to perform depends on the productivity of traits in different tasks.

More generally, people perform multiple tasks at any point in time.[57] A less discrete version, which is our case II, builds on the same foundations, and allows people to perform multiple tasks at any time and postulates that $\phi_j\left(\theta, e_j\right)$ is concave and increasing in e_j.[58] The agent maximizes

$$\sum_{j=1}^{J} R_j P_j\left(\theta, e_j\right) \tag{3}$$

subject to $\sum_{j=1}^{J} e_j = \overline{e}$. The first order conditions for this problem are standard: $R_j \frac{\partial P_j}{\partial e_j} \geq \lambda$, $j = 1, \ldots, J$, where λ is the vector of multipliers associated with the effort constraint. Some people may make no effort in tasks. P_j may be zero if $e_j = 0$, but this is not strictly required.[59]

As rewards change favoring activity $j\left(R_j \text{ increases}\right)$, effort devoted to j will increase.[60,61] This model is consistent with compensation for shortfalls in endowments as well as

[57] This, of course, depends on the time unit. Agents may be able to do only one task at one time if the time unit is defined finely enough.
[58] Failure of concavity can take us back to case I.
[59] Again, it is straightforward to generalize this reward function to a general utility function $U\left(P_1, \ldots, P_J\right)$.
[60] $\frac{\partial^2 P_j}{\partial \theta \partial e_j} > 0$ is a force toward devoting more effort to task j. If effort is complementary with traits in all tasks, as traits expand, more effort will be expended in those tasks that are relatively more complementary in effort.

reinforcement, depending on the pattern of complementarity and substitutability. Different situations may be associated with different rewards for the same task. Such variation can produce the situational specificity of measured personality that was featured in the person-situation debate discussed in Section 2. One needs to standardize for the incentives to exert effort across tasks and for the endowment of effort in order to use measurements of personality to identify traits.

3.A. Identifying Personality Traits

The psychological traits in θ have not yet been specified. At the current level of generality, all traits can affect productivity in all tasks. However, some tasks may require only a single trait or a subset of all of the traits. Divide θ into "mental" (μ) and "personality" (π) traits: θ_μ and θ_π, each of which may be a vector.[62] This corresponds to the two types of traits in Roberts' model, presented in Figure 2.[63]

Psychological measurement systems sometimes use the productivity in different tasks to measure 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

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

---

[61] In case I, agents will pick j.

[62] Effort endowment might also be divided in the same fashion $\left(\overline{e}_\mu, \overline{e}_\pi\right)$, but we do not explicitly develop this possibility.

[63] Any such division may be quite arbitrary, but the division has an intuitive basis.

Even if we can measure productivity in j, P_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 $\theta_{1,j}$. 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.[65]

The activity of picking a task (or a collection of tasks) that measure a particular trait ($\theta_{i,j}$ in our example) is called *operationalization* in psychology. Demonstrating that a measure successfully operationalizes a trait is called ***construct validity***. If effort is involved in the performance of a task 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 measure 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. We present examples of such contamination by incentives in Section 5.

Operationalization and construct validation, while widely invoked in creating psychological measures, requires 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 j. Without further information one cannot infer which of the two traits produces the productivity in j. But in

[65] A weaker notion is to achieve relative ranks of a trait. One can define the rank of a trait holding fixed the ranks of all other influences.

general, even having two (or more) measures of productivity that depend on $\left(\theta_{1,\mu},\theta_{1,\pi}\right)$ is not enough to identify the separate components. Consider the following case of two productivity measures for the two tasks j and j':

$$P_j = \phi_j\left(\theta_{1,\mu},\theta_{1,\pi},e_j\right)$$
$$P_{j'} = \phi_{j'}\left(\theta_{1,\mu},\theta_{1,\pi},e_{j'}\right), \qquad j \neq j'.$$

Standardize measurements at a common level of effort $e_j = e_{j'} = e^*$.[66] If the system of equations satisfies a local rank condition, then one can solve for the pair $\left(\theta_{1,\mu},\theta_{1,\pi}\right)$ at e^*.[67] Note, however, that only the pair is identified. One cannot (without further information) determine which component of the pair is $\theta_{1,\mu}$ or $\theta_{1,\pi}$. In Section 5, we present an example where scores on achievement tests depend on both IQ and personality traits. In the absence of dedicated constructs (constructs that are generated by only one component of θ), there is an intrinsic identification problem that arises in using measures of productivity in tasks to infer traits.[68]

3.B. Extensions

The simple gross income maximizing framework of the Roy model is widely used in empirical work by economists studying the effects of personality on outcomes. It is amended in

---

[66] Note that if the support of e_j and $e_{j'}$ is disjoint, no e^* exists.

[67] Formally, it is required that the Jacobian of the system

$$\left|\begin{bmatrix} \dfrac{\partial^2 \phi_{j'}}{\partial \theta^2} & \dfrac{\partial^2 \phi_{j''}}{\partial \theta^2} \end{bmatrix}'\right|$$

be non-vanishing in open neighborhoods (see, e.g., Buck [2003]).

[68] There are various ways around this identification problem. For example, one might be able to choose configurations of data with low (or zero) values of one component. At high levels of effort, induced by a change in the reward, the effect of one component on productivity might vanish, etc.

many papers by attaching a cost $C_j\left(\theta, e_j\right)$ to obtaining the reward so that instead of criterion (2), the agent picks $\hat{j}$ that maximizes the net reward

$$\hat{j} = \arg\max_{j \in \{1,\ldots,J\}} \left\{ R_j P_j\left(\theta, \bar{e}\right) - C_j\left(\theta, \bar{e}\right) \right\},$$

and instead of (3), for case II the agent maximizes

$$\sum_{j=1}^{J} R_j P_j\left(\theta, e_j\right) - C_j\left(\theta, e_j\right).$$

This extension creates a further identification problem—whether the trait identified arises from its role in costs, productivity, or both. The identification problem deepens when we allow the costs to be psychic costs as in Heckman and Sedlacek [1985], Cunha, Heckman and Navarro [2005], or Heckman, Stixrud and Urzua [2006], and attempt to separate out productivity traits from preference traits.[69]

This framework is widely used in recent analyses of the role of personality and cognition. See, e.g., Heckman, Stixrud and Urzua [2006], Heckman, Humphries, Urzua et al. [2010], Báron and Cobb-Clark [2010], and Cattan [2010]. It has precedents in the work of Mandelbrot [1962], Heckman and Sedlacek [1985], and Heckman and Honoré [1990]. In most applications the $P_j\left(\theta, e_j\right)$ and $C_j\left(\theta, e_j\right)$ (or their logarithms) are assumed to be linear or log linear in θ and e_j:

$$\begin{aligned} P_j &= \alpha_\theta' \theta + \alpha_e' e_j \\ C_j &= \beta_\theta' \theta + \beta_e' e_j. \end{aligned}$$

The analyst models both the choice of the task and the output from the task chosen. A third (mixed) case (III) can arise in which some clusters of tasks are mutually exclusive so the agent

[69] Heckman and Navarro [2007] and Abbring and Heckman [2007] present conditions that allow identification of productivity and costs when there are direct measures of gross productivity, at least when there are measurements on P_j for individuals who select j.

can perform only one task within each cluster of tasks, but the agent can simultaneously engage in multiple tasks across clusters.

Preferences and goals (see Figure 2) may also shape effort.[70] This takes us to a fourth and more general case. There may be direct utility benefits or costs associated with exerting effort in each task. Array the effort across tasks in vector $e = (e_1, \ldots, e_J)$.

One might also attach direct value to the productivity in tasks arrayed in vector $P = (P_1, \ldots, P_J)$ with reward R_j. Output can produce income $\sum_{j=1}^{J} R_j P_j$ which can be spent on goods X with associated prices W. A utility function can be specified over P, e, and X with preference parameter vector η.[71] Thus, we may write

(4) $$U(X, P, e \mid \eta),$$

where the agent maximizes (4) subject to the constraints

(5) $$Y + R' P = W' X,$$

where Y is a flow of unearned income available to the agent in addition to his earnings from his productive activities, and

(6) $$\sum_{j=1}^{J} e_j = \bar{e}.$$

Preference specification (4) captures the notions that (a) agents have preferences over goods, (b) agents may value the output of tasks in their own right, and (c) agents may value the effort devoted to tasks.

[70] In some versions of the preceding models with costs, preferences can be embodied in psychic costs.
[71] Robson [1996; [2001] and Robson and Samuelson [2007; [2009] discuss the evolutionary origin of preference parameters.

The parameter η describes the parameters determining the tradeoffs in preferences among X, P, and e. In one interpretation, subjective measures of well-being (Layard [2005]) attempt to directly measure (4).[72] Parameters that affect subjective well-being but not choices can be identified from the measures of well-being, but not from choices.

The model can readily be extended to cover more general cases. There is no need to impose the linear reward structure $\left(R'P\right)$. The resources raised from productive tasks can be a nonlinear in P. Another simple extension of the model is the case where there is no financial gain from engaging in tasks, but the agent receives a direct utility benefit from doing so. In this case, constraint (5) is redefined as $Y = W'X$, but P remains as an argument of the utility function. One might also introduce goods as inputs into the ϕ_j functions.

One can extend all of the preceding models to account for learning and uncertainty. Let $\mathcal{I}$ be the information possessed by the agent and "E" denote the expectation operator. The agent can be interpreted as making decisions based on

$$E\left[U\left(X,P,e,\eta\right)|\mathcal{I}\right].^{73}$$

In a general specification, agents can be uncertain about their preferences (η), their traits $\left(\theta\right)$, the prices they face $\left(P\right)$, the outcomes of purchase decisions $\left(X\right)$, and their endowments of effort $\left(\bar{e}\right)$. Through choices to exert effort in different tasks, preferences (η) shape effort.

The use of the expectations operator begs the question of how subjective expectations are formed. There may be psychological traits θ that affect information perception and processing.

---

[72] However, the happiness literature is not strictly wedded to the notion that happiness is the same as U.
[73] Obviously, we can add many arguments to the utility function such as θ, i.e. a person may take direct utility from their traits.

For example, overconfidence may be a trait that causes persons to inflate their perceived productivity (see, e.g., Niederle, Möbius, Niehaus et al. [2010], Akerlof and Dickens [1982], Caplin and Leahy [2001], and Köszegi [2006]). A production function for information may depend on components of θ $(\theta_{\mathcal{I}})$ and the effort devoted to acquire information $e_{\mathcal{I}}$. Intelligent people may acquire information more readily than dull people. People more open to experience likely acquire more knowledge. Aggressive people may reduce their social interactions and impair their ability to learn from others.

3.C. An Economic Definition of Personality

Personality *traits* are components of e, θ and η that affect behavior. As previously discussed, identifying these traits may be a challenging problem. We only observe measured personality—behaviors generated by incentives, goals, and traits.

One might define measured personality as the performance $(\text{the } P_j)$ and effort $(\text{the } e_j)$ that arise from solutions to the optimization problems previously discussed. Thus, the derived productivity and effort functions would constitute measured personality as a response to constraints, information, and preferences, i.e. as functions that solve out for the P_j and e_j that agents choose. This approach would not capture the full range of behaviors considered by personality psychologists as constituting aspects of personality except as a reduced form expression. Actions include dispositions in performing tasks. The actions considered by psychologists include a variety of activities that economists normally do not study, e.g., cajoling, beguiling, bewitching, charming, etc. Thus, in selling a house, various actions might be taken, e.g., smiling, persuading people by reason, threatening, scowling, showing affection, etc.

Colloquially, "there are many ways to skin a cat," and the choice of which way to do so in any task defines the action taken.

To capture these more general notions, we introduce a set of "actions" broader than what is captured by e. Actions are *styles* of behavior that affect *how* tasks are accomplished. They include aspects of behavior that go beyond effort as we have defined it.

Any task can be accomplished by taking various actions. We denote the ith possible action to perform task j by $a_{i,j}$, $i \in \{1,...K_j\}$. Array the actions in a vector $a_j = \left(a_{1j},\ldots,a_{K_j j}\right)$. The actions may be the same or different across the tasks. Thus one can smile in executing all tasks or one may smile in only some. The productivity of the agent in task j depends on the actions taken in that task:

(7) $$P_j = f_j\left(a_{1j}, a_{2j},...a_{K_j j}\right).$$

The actions themselves depend on traits θ and "effort" $e_{i,j}$:

(8) $$a_{i,j} = \nu_{i,j}\left(\theta, e_{i,j}\right),$$

where

$$\sum_{i=1}^{I_j} e_{i,j} = e_j \text{ and } \sum_{j=1}^{J} e_j = \bar{e}.$$

Less effort may be required to perform a given action if a person has endowment θ that favors performance of the action. For example, a person may be "naturally sunny" and as such may find it easy to engage in social interactions. Stated this way, actions generalize the notion of effort to a broader class of behavior. Analytically, they play the same role as effort and some actions may be components of effort. There may be utility costs or benefits of effort exerted. A special case is when there are increasing returns to effort in each action. In that case, the agent

will simply apply all of his effort e_j in task j to the action which gives him the highest productivity, and the other possible actions are not taken.

Agents may have utility over actions beyond the utility they get from tasks. An agent may prefer accomplishing a task by working hard rather than by cheating. We can define the utility over actions. Let a denote the choice of actions applied to all tasks $\left(a=\left(a_1,\ldots,a_J\right)\right)$. The agent solves

$$\max U\left(a, X, P, \eta \mid \mathcal{I}\right)$$

with respect to X, e given the previously stated constraints. Actions may also directly affect $\mathcal{I}$, so the production of information can depend on θ, e and a. The choice of which actions to take depends on goals and values (captured by η) and on the available information.

One can extend the framework to introduce the effects of the situation in the person-situation debate, by considering specific situations indexed by $h \in \mathcal{H}$. These situations are assumed to affect productivity by affecting the set of possible actions and hence the action taken. Thus for a person with traits θ and effort vector e_j with action $a_{i,j}$, using the specification (8), the action function can be expanded to be dependent on situation h:

(9) $$a_{i,j,h} = \nu_{i,j,h}(\theta, e_{i,j}),$$

and productivity on a task can be specified solely as a functions of the action taken to perform the task

(10) $$P_{j,h} = f_j(a_{1,j,h},...,a_{K_j,j,h})$$

or in a more general specification where situation h, along with traits, has a direct effect on productivity in addition to their effects on actions taken.

(11) $$P_{j,h} = f_{j,h}(\theta, a_{1,j,h}, ..., a_{K_j,j,h}).$$ [74]

Situations could include physical aspects of the environment in which the agent is located or the network (and other social situations) in which the agent is embodied. The situation can include social factors such as peer effects.[75] Persons taking an achievement test sometimes do much worse if the significance of their scores for social perceptions of their group is stressed, as in the stereotype threat literature (Steele and Aronson [1998], Sackett, Hardison and Cullen [2004]).

The situation represents a key notion in the "person-situation" debate discussed in Section 2. Equations (9)-(11) capture the "if-then" notion of Mischel and Shoda [1995]. Under specification (11), agents with the same actions, the same efforts, and the same traits may have different productivities. Failure to control for situation h, just like failure to control for effort, will contaminate identification of traits using measures of actions or productivities. Situations may be forced on the agents or may be chosen. The situations are assumed to be mutually exclusive.[76]

Let T be the vector of traits $(\theta, \eta, \bar{e})$. They are endowments at any point in time. In the general case, the solution to the constrained maximization problem is to pick goods (X), situation (h), action $(a_{i,j})$, and effort (e_j), $j \in \{1, \ldots, J\}$ subject to the constraints. h is fixed if the situation is forced on the agent. For simplicity, we analyze this case. Relaxing it is straightforward but notationally cumbersome.

---

[74] A variety of other cases can be considered where instead of subscripting functions by "h," we explicitly characterize the arguments by measured characteristics.
[75] Included in situation h might be the act of being observed by third parties and other possible sources of social interactions.
[76] At the cost of further notation, we could make the set of possible situations task-specific.

For the case of fixed h, the solution to the maximization problem produces a set of response functions. Preference parameters (η) determine the choices and actions taken through their influence on the tradeoffs and goals that characterize consumer preferences:

(12) $$X = X(T, h, W, Y, R, \mathcal{I})$$

(13) $$e = e(T, h, W, Y, R, \mathcal{I})$$

(14) $$a = a(T, h, W, Y, R, \mathcal{I}).$$

Productivity P across tasks is derived from the actions, efforts, and traits of the agents.[77]

The behaviors that constitute personality are defined as a *pattern of actions* in response to the constraints, endowments, and incentives facing agents given their goals and preferences. This interpretation incorporates the notion that personality is a strategy response function (14). People may have different personalities depending on their trait endowments, constraints, and situations. The actions—not the traits—constitute the data used to identify the traits.

Introducing actions widens the set of data from which one can infer the components of *T*. Personality psychologists often use actions (e.g., "dispositions") to infer traits. The same identification issues previously discussed continue to arise but now apply to a broader set of measurements. The arguments made in the simple case carry through to the more general settings.

Using the framework of this section, we show how psychological notions such as disposition and situation can be represented within standard economic choice models. In Section 5, we use this framework to discuss a fundamental identification problem of using observed behaviors to measure the traits that generate the behavior.

---

[77] For the case of h chosen, we get a system of derived demands for $X, h, a_{i,j}, e_j$.

3.D. Life Cycle Dynamics

The analysis in the preceding subsection was for a particular point in time (e.g., a period). Traits are not set in stone. In a dynamic setting, one can think of traits traits (T), information $(\mathcal{I})$, situations $(\mathcal{H})$, and actions $(\mathcal{A})$ as state variables which evolve through aging, experience, and investment. As a result of experience (including social interactions), situations, biology (ontogeny), and investment, traits may change over the life cycle. We now briefly discuss the dynamics of trait and state formation.

To capture the evidence from a large and growing literature, we consider the dynamic evolution of traits.[78] Let T^{τ} be traits at age τ, $\tau \in \{1,\ldots,\mathcal{T}\}$. They may change through family and self investment (Cunha and Heckman [2007; [2009]), through schooling, through biology or through experience. Information $\mathcal{I}^{\tau}$ may be updated through various channels of learning. All task outputs, actions and goods inputs may be time dated.

Investment in period τ is an action or set of actions that an individual (or a person or group acting for the individual) may take in period τ. Investments have dynamic effects. The technology of skill formation (Cunha and Heckman [2007; [2009]) captures the notion that traits may evolve in response to the inputs of a vector of investments (IN^{τ}), and through aspects of the situation in which the agent is found, h^{τ}, where $S_{h^{\tau}}$ is the vector of attributes of the situation:

(15) $$T^{\tau+1} = f^{\tau}(\underbrace{T^{\tau}}_{\text{self-productivity}}, \underbrace{IN^{\tau}}_{\text{investment}}, S_{h^{\tau}}), \tau = 0,\ldots,\mathcal{T}-1$$

[78] We survey the evidence on the life cycle dynamics of traits in Section 8, focusing primarily on the traits θ that affect measured productivity.

where the first set of arguments arises from self and cross productivity (so skill begets skill; traits beget other traits and traits cross-foster each other; see Cunha and Heckman [2007; [2009]). The second set of arguments arises from investment. Investment is a broad concept and includes parental nurturance, schooling, learning by doing, and learning by imitation, etc. The third set of arguments arises from the situation in which the person is placed.[79]

Notice that if elements of T^{τ} are augmented over the life cycle through investment and practice, the actions and efforts required to achieve a given task can change. Thus, if θ^{τ} is enhanced over time, the amount of effort required to perform a task may be reduced. In this way, we can model habit formation and capture the notion of *arete*, effortless performance of actions, discussed in Aristotle [1956].[80]

As emphasized by Mischel and Shoda [1995], situations may change over time as a function of past actions, past situations, investment, information, and the like.

(16) $$h^{\tau+1} = \psi^{\tau}\left(h^{\tau}, IN^{\tau}, a^{\tau}\right).$$

Past actions may serve to determine the set of present situations. Those situations in turn may influence my current actions.

Information $\mathcal{I}^{\tau}$ may change over the life cycle through experimentation as well as through exogenous learning:

(17) $$\mathcal{I}^{\tau+1} = \phi^{\tau}\left(\mathcal{I}^{\tau}, a^{\tau}, T^{\tau}, IN^{\tau}, h^{\tau}\right).$$

This learning mechanism incorporates the beliefs of agents about the available data. Thus people may learn about the environments and themselves in part as a consequence of their own actions

[79] The actions taken by agents might also enter as arguments to this technology.
[80] See Lear [2004]. A habit can be defined as effortless performance of a task, i.e., an action that requires no effort. It is possible to build a stock of traits to sufficient level that one achieves actions effortlessly.

and in part as a consequence of the exogenous arrival of information. Equations of motion (15)-(17) are very general. We specialize them in Section 8 when we review the work on the life cycle dynamics of traits.

There is a rich and evolving literature on dynamic preferences when agents do not possess full knowledge of their future environments (see, for example, Hansen [2005], Hansen and Sargent [2008], Rust [2008], Epstein and Zin [1989], Epstein and Schneider [2003], and Skiadas [1998]) that is too large to summarize here. Preferences need not be separable over time, there may be time inconsistency of choices associated with hyperbolic discounting.[81] We discuss commonly used dynamic preference preference specifications in Section 5.F.

3.E. Relationship of the Model in This Section to the Existing Models in Personality Psychology

Personality psychologists are generally not as formal as economists in characterizing their models. Our formalization is, to our knowledge, the first mathematically precise definition of personality traits and measured personality. The models we have sketched in this section capture central features of the major models in personality psychology.

By its authors' own admission, the Costa-McCrae [2008] Five Factor Theory is not a fully articulated model. It emphasizes the role of traits (T) and, in particular, the Big Five factors in producing outcomes and agent actions and is sketchy about other details. Agents are assured to learn about their own traits, but precise learning mechanisms are not specified. Expression of traits is affected by the external environment and through social interactions in a not fully specified fashion. The concept of an evolving information set $(\mathcal{I}^{\tau})$ plays a central

[81] See Kirby and Herrnstein [1995] and Gul and Pesendorfer [2004].

role in Five Factor Theory. People learn about their traits through actions and experience, but how this occurs is not given. These notions are captured by equation (17). Situations may also evolve as a function of actions and experience, but no role is assigned to investment.

Thus, a restricted version of (16) formalizes aspects of the Five Factor Theory. The theory features "characteristic adaptations," which correspond to the actions and efforts of our model that also affect the productivity in tasks. The role of preferences is left unspecified. However, Costa and McCrae explicitly feature *rationality* (McCrae and Costa [2008, p. 161]) and reject the characterization of flawed human decision making that dominates social psychology and the field of behavioral economics that was spawned from social psychology. They explicitly reject a purely situationist explanation of the origin of actions, but they allow for situations to affect actions. Traits evolve through biological processes (ontogeny), but no role is assigned to investment or experience as a determinant of traits. Thus, the arguments of (15) are shut down, but traits may still exogenously evolve as a function of age and the biology of the individual. While traits evolve through the processes of maturation and biology, persons may learn about themselves (their traits) by taking actions and being acted on by the external environment.

A rival to trait theories based on the Big Five are "social cognitive" theories.[82] Central figures in this literature are Albert Bandura, Daniel Cervone, and Walter Mischel. Roberts' diagram (Figure 2) captures key aspects of this theory. This line of thinking stresses the role of cognition in shaping personality and the role of social context in shaping actions and self-knowledge. Authors writing in this school of thought explicitly reject the "cognitive-noncognitive" distinction that is often used in economics. A major role is assigned to agency—

[82] See Cervone and Pervin [2009].

individual goals and motives that produce actions. Goals and motives are captured by η . Although the personality psychology literature contrasts these two lines of thought, to us the lines are not distinct. In one extreme version of the social-cognitive theory, traits are entirely absent. Agent behavior is entirely shaped by situations. Mischel and Shoda [2008] focus on the role of situation in shaping actions, efforts, and productivities, but, as previously discussed, also allow for traits to influence actions.

Thus, both schools of thought accept specification (8) or its extension (9), and both would be comfortable with response systems (12)-(14). The relative importance of the factors emphasized by the two schools of thought can only be settled by empirical research. The social-cognitive theorists tolerate deviations from rationality in their theories, while trait theorists do not.

Both schools of thought entertain the possibility of learning about oneself. A major difference between the two groups comes in the role of investment. The social-cognitive theorists feature investment and social interactions as direct determinants of traits that are assumed to evolve as a function of the experiences of agents. The trait theorists do consider this possibility. Instead they emphasize self-learning about traits that may evolve by fixed biological principles unrelated to the experiences of individuals.

4. Measuring Personality

Unlike other personal traits, like height or weight, personality traits cannot be directly measured. The observed productivities, efforts, and actions are used to infer traits. We now discuss notions of validity of measurements used in psychology.

4.A. Linear Factor Models

Linear factor models are widely used in personality psychology and in psychometric models for mental testing. We review the use of these models in psychology. They are coming into widespread use in economics. To capture essential points, we consider measurements arising from productivity in tasks and thus can focus solely on outputs of tasks, abstracting from actions, efforts, and situations. With suitable notation we can encompass these ingredients in what follows. We assume additive separability of the arguments of equation (1). The stripped down model writes task performance of person n on task j in the following manner:

(18) $$P_{n,j} = \mu_j + \lambda_j' T_n + \Delta_{n,j},\ n = 1,\ldots,N,\ j = 1,\ldots,J,$$

where μ_j is the mean of the jth task, λ_j is a vector of factor loadings and $\Delta_{n,j}$ is other determinants of measured performance, including measurement errors. The number of components in T_n, L, has to be small relative to J ($L<<J$) if the factor model is to have any explanatory power. Otherwise for each task one can create a unique factor and the model becomes tautological. A purely cognitive task would be associated with zero values of the components of vector λ_j on elements of T_n that are associated with personality traits. Factor model (18) captures the notions that: (a) latent traits T_n generate a variety of outcomes, (b) task

outputs are imperfect measures of the traits $\left(T_n\right)$ because $\Delta_{n,j}$ also determines task output, and (c) tasks other than tests or observer reports may also proxy the underlying traits, i.e., latent traits generate both test scores and behaviors. Notice that tasks may depend on vector T_n and outcomes across tasks may be correlated even if the components of T_n are not. Thus a correlation of outcomes across tasks can arise because tasks depend on the same vector of traits.[83]

4.B. Convergent and Discriminant Validity

In this notation, personality psychologists largely focus on observer- and self-reports as measures of $P_{n,j}$. The measurements are designed to capture a particular trait. As discussed in Section 3, the choice of which collection of tasks is used to measure a capability ("operationalization and construct validity") is an inherently subjective activity. Many psychologists take a pragmatic, empirical point of view. Traits are what the measurements used capture. The danger with this empiricist definition is that it offers no guide to the choice of measurements, which are usually settled by conventions or intuitions.

The concept of "discriminant validity" of a collection of tasks (e.g., a set of test scores or a set of observer reports) is commonly used to test for construct validity. It exploits the notion that a particular battery of measurements captures a component of T_n, for example, $T_{n,l}$, and not other components. Many measurements may be taken on $T_{n,l}$ and having multiple measurements helps to control for measurement error.

[83] The strength of the correlation depends on the magnitudes of λ_j, $\lambda_{j'}$ across the two tasks, j and j'.

All measurements are really just outcomes on a type of task (although the effort applied may vary greatly across tasks). The literature in psychology assigns a special status to tests, self-reports, and observer-reports of latent traits, but we do not. Behaviors, tests, observer reports, and self reports all proxy the underlying traits.

Let $P_{n,l}^{q}$ be the qth measurement (by test or observer-report on task performance) on trait l for person n. Using a linear factor representation, the qth measurement of factor l for person K is represented as

(19)
$$P_{n,l}^{q} = \mu_l^q + \lambda_l^q T_{n,l} + \epsilon_{n,l}^{q},$$
$$q = 1,...,Q_l, \quad n = 1,...,N, \quad l = 1,...,L.$$

The factor $T_{n,l}$ is assumed to be statistically independent of the "measurement errors," $\epsilon_{n,l}^{q}$, $q = 1,...,Q_l$. Different factors are assumed to be independent ($T_{n,l}$ independent of $T_{n,l'}$ for $l \neq l'$). The measurement errors (or "uniquenesses") are assumed to be mutually independent within and across constructs.[84]

In fact, measurement $P_{n,l}^{q}$ may depend on other components of T_n, so that the measurement captures a composite of latent traits. A more general case is

(20)
$$P_{n,l}^{q} = \mu_l^q + \left(\lambda^q\right)' T_n + \epsilon_{n,l}^{q}, \quad q = 1,...,Q_l,$$

where λ^q is a vector with possibly as many as L nonzero components. The $\epsilon_{n,l}^{q}$ are assumed to be independent of T_n and mutually independent within and across constructs (l and l' are two

[84] Cunha and Heckman [2008] and Cunha, Heckman and Schennach [2010] present conditions under which this assumption is substantially weakened and identification of factors is still possible.

constructs). The task has *discriminant validity* for trait l if λ_l^q is the only nonzero component of T_n. The μ_l^q and λ_l^q can depend on measured characteristics of the agent, Q_n.[85]

A standard approach to defining constructs in personality psychology is based on factor analysis. It takes a set of measurements (including observer- and self-reports) that are designed to capture a construct arrived at through intuitive considerations and conventions, and measures within-cluster and across-cluster correlations of the measurements to isolate latent factors $t_{n,l}, l = 1,...,L$ or their distributions. The measurements and clusters of tests are selected on intuitive grounds or a priori grounds, and not on the basis of any predictive validity in terms of real-world outcomes (for example, success in college, performance on the job, earnings). This process gave rise to the taxonomy of traits that became the Big Five. Because of the somewhat arbitrary basis of these taxonomies, there is some controversy in psychology about competing construct systems as discussed in Section 5. In practice, as we document below, the requirement of independence of the latent factors across constructs (lack of correlation of tests across clusters) is not easily satisfied.[86] This fuels controversy among psychologists advocating competing taxonomies.

Conventional psychometric validity of a collection of items or test scores for different constructs thus has three aspects. (a) A factor T_l is assumed to account for the intercorrelations among the items or tests within a construct *l*. (b) Item-specific and random error variance are

[85] Hansen, Heckman and Mullen [2004] show how to allow Q_n to depend on T_l and still identify the model

[86] Indeed, as documented in Section 7, the factors associated with personality are also correlated with the cognitive factors.

low (intercorrelations among items are high within a cluster).[87] (c) Factor T_l for construct l is independent of factor $T_{l'}$ for construct $l' \neq l$. Criteria (a) and (b) define "convergent validity." Criterion (c) is "discriminant validity."

Nothing in these testing procedures guarantees that the measurements that satisfy convergent and discriminant validity identify a single trait. Multiple traits operating in the same fashion across many outcomes can produce outcomes that satisfy the criteria.

4.C. Predictive Validity

An alternative criterion for validating measurement systems is the predictive power of the tests for real world outcomes, that is, on behaviors measured outside of the exam room or observer system. The Hogan Personality Inventory,[88] the California Personality Inventory, and the Minnesota Multiphasic Personality Inventory were all developed with the specific purpose of predicting real-world outcomes. Decisions to retain or drop items during the development of these inventories were based, at least in part, upon the ability of items to predict such outcomes. This approach has an appealing concreteness about it. Instead of relying on abstract a priori notions about domains of personality and subjectively defined latent factors generated from test scores and self and observer personality assessments, it anchors measurements in tangible, real-world outcomes and constructs explicit tests with predictive power. Yet this approach has its own problems.

First, all measurements of factor $T_{n,l}$ can claim incremental predictive validity as long as each measurement is subject to error $\left(\epsilon^q_{n,l} \neq 0\right)$. Proxies for $T_{n,l}$ can appear to be separate

[87] Cronbach's alpha is a widely used measure of intercorrelation among test scores, that is, a measure of importance of the variance of the $\epsilon^q_{n,l}$ uniquenesses relative to the variance of the factors. See Hogan, Hogan and Roberts [1996] for a precise definition. Sijtsma [2009] discusses the severe limitations of Cronbach's alpha.
[88] See http://www.hoganassessments.com/products_services/hpi.aspx and also Hogan and Roberts [2001].

determinants (or "causes") instead of surrogates for an underlying one-dimensional construct or factor. Thus suppose that measurement system (19) is the correct specification and that a set of measurements display both convergent and discriminant validity. As long as there are measurement errors for construct *l*, there is no limit to the number of proxies for $T_{n,l}$ that will show up as statistically significant predictors of an outcome. This is a standard result in the econometrics of measurement error. (See, e.g., Aigner, Hsiao, Kapteyn et al. [1984]). For this reason, it is necessary to correct for measurement error in using predictive validity to define and measure traits.

A second problem is reverse causality. This is especially problematic when interpreting correlations between personality measurements and outcomes. Outcomes may influence the personality measures as well as the other way around. For example, self-esteem might increase income, and income might increase self-esteem. Measuring personality traits prior to measuring predicted outcomes does not necessarily obviate this problem. For example, the anticipation of a future pay raise may increase present self-esteem. Heckman, Stixrud and Urzua [2006] and Urzua [2008] demonstrate the importance of correcting for reverse causality arising from schooling affecting traits and traits affecting schooling in interpreting the effects of personality tests on a variety of socioeconomic outcomes. Application of econometric techniques for determining the causal effects of factors on outcomes makes a distinctive contribution to psychology.

Psychologists sometimes attempt to circumvent these problems by using early measures of traits determined long before the outcomes are measured. This is problematic if the traits evolve over time and the contemporary traits drive behavior. This practice trades a reverse

causality problem for an errors in variables problem. In our review of the literature in Section 7, we distinguish studies that attempt to control for reverse causality and those that do not.

Many psychologists focus on prediction, not causality. Establishing predictive validity will often be enough to achieve the goal of making personnel assignment and student placement decisions.[89] However, for policy analysis, including analyses of new programs designed to augment the skills of the disadvantaged, causal models are required.[90]

The papers of Heckman, Stixrud and Urzua [2006], Urzua [2008], and Cunha and Heckman [2008], are frameworks for circumventing the problems that arise in using predictive validity alone to define and measure personality constructs.[91] These frameworks recognize the problem of measurement error in the proxies for constructs. Constructs are created on the basis of how well latent factors predict outcomes. They develop a framework for testing *discriminant validity* because they allow the factors across different clusters of constructs to be correlated, and can test for correlations across the factors.

They use an extension of factor analysis to represent proxies of low-dimensional factors. They test for the number of latent factors required to fit the data and rationalize the proxies.[92] Generalizing the analysis of Hansen, Heckman and Mullen [2004], they allow for lifetime experiences and investments to determine, in part, the coefficients of the factor model and to affect the factor itself. They allow for the latent factor to determine investment and experience. They correct estimates of latent factors on outcomes for the effects of spurious feedback, and

---

[89] See, e.g., Hogan and Roberts [2001].
[90] See, for example, Hogan and Hogan [2007].
[91] Cudeck and MacCullum [2007] use a version of this procedure. In this regard, they appear to be an exception among personality psychologists. However, in psychometrics, there is a long tradition of doing predictive analysis based on factor analysis (see, for example, the essays in Hansen, Heckman and Mullen [2004]), but there is no treatment of the problem of reverse causality. Conti, Heckman, Lopes et al. [2010] discuss alternative approaches to selecting the number of latent factors. See also Cragg and Donald [1997].
[92] For example, Ones and Viswesvaran [1998] present classical statistical methods for determining the number of factors. In addition to their techniques, there are methods based on Bayesian posterior odds ratios.

separate proxies from factors. The factors are estimated to change over the life cycle as a consequence of experience and investment. We review these studies in Sections 7 and 8 of this chapter.

Another issue is that measurements of latent factors may be corrupted by "faking." There are at least two types of false responses: those arising from impression management and those arising from self-deception (Paulhus [1984]). For example, individuals who know that their responses on a personality questionnaire will be used to make hiring decisions may deliberately exaggerate their strengths and downplay their weaknesses.[93] Subconscious motives to see themselves as virtuous may produce the same faking behavior, even when responses are anonymous. Of course, it is possible to fake Conscientiousness on a self-report questionnaire whereas it is impossible to fake superior reasoning ability on an IQ test. To a lesser degree, a similar bias may also operate in cognitive tests. Persons who know that their test scores will affect personnel or admissions decisions may try harder. The effects of faking on predictive validity have been well-studied by psychologists, who conclude that distortions have surprisingly minimal effects on prediction of job performance (Hough, Eaton, Dunnette et al. [1990]; Hough and Ones [2002]; Ones and Viswesvaran [1998]). Correcting for faking using scales designed to measure deliberate lying does not seem to improve predictive validity (Morgeson, Campion, Dipboye et al. [2007]). Nevertheless, when measuring cognitive and personality traits, as noted in Section 3, one should standardize for incentives and environment.

The linear factor model does not capture a variety of interesting interactions are not captured. Cunha, Heckman and Schennach [2010] and the papers they cite develop a nonlinear

[93] See Viswesvaran and Ones [1999]; and Sternberg [2000; [2001].

non-normal factor analysis that allows for measurement errors to be correlated across measures and over time.

5. Implementing the Measurement Systems

This section discusses the major measurement systems for cognition and personality. For cognition, there is a fairly well-established terminology. Aptitude tests are designed to measure differences in the rate at which individuals learn (i.e., fluid intelligence). Achievement tests are designed to measure acquired knowledge (i.e., crystallized intelligence). For personality, a variety of synonymous terms are used, and this proliferation of terms can be confusing. We attempt to equate these alternative descriptions and, further, link them to measures of childhood temperament and psychopathology.

5.A. Cognition

Intelligence (also called cognitive ability and general mental ability) includes the "ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought" (Neisser, Boodoo, Bouchard et al. [1996, p. 77]). The term "IQ" is often used synonymously with intelligence but in fact refers specifically to scores on intelligence tests.[94] Notwithstanding a century of active study and general agreement about the sorts of tasks on which more intelligent individuals perform better, the construct of intelligence "resists a consensual definition."[95]

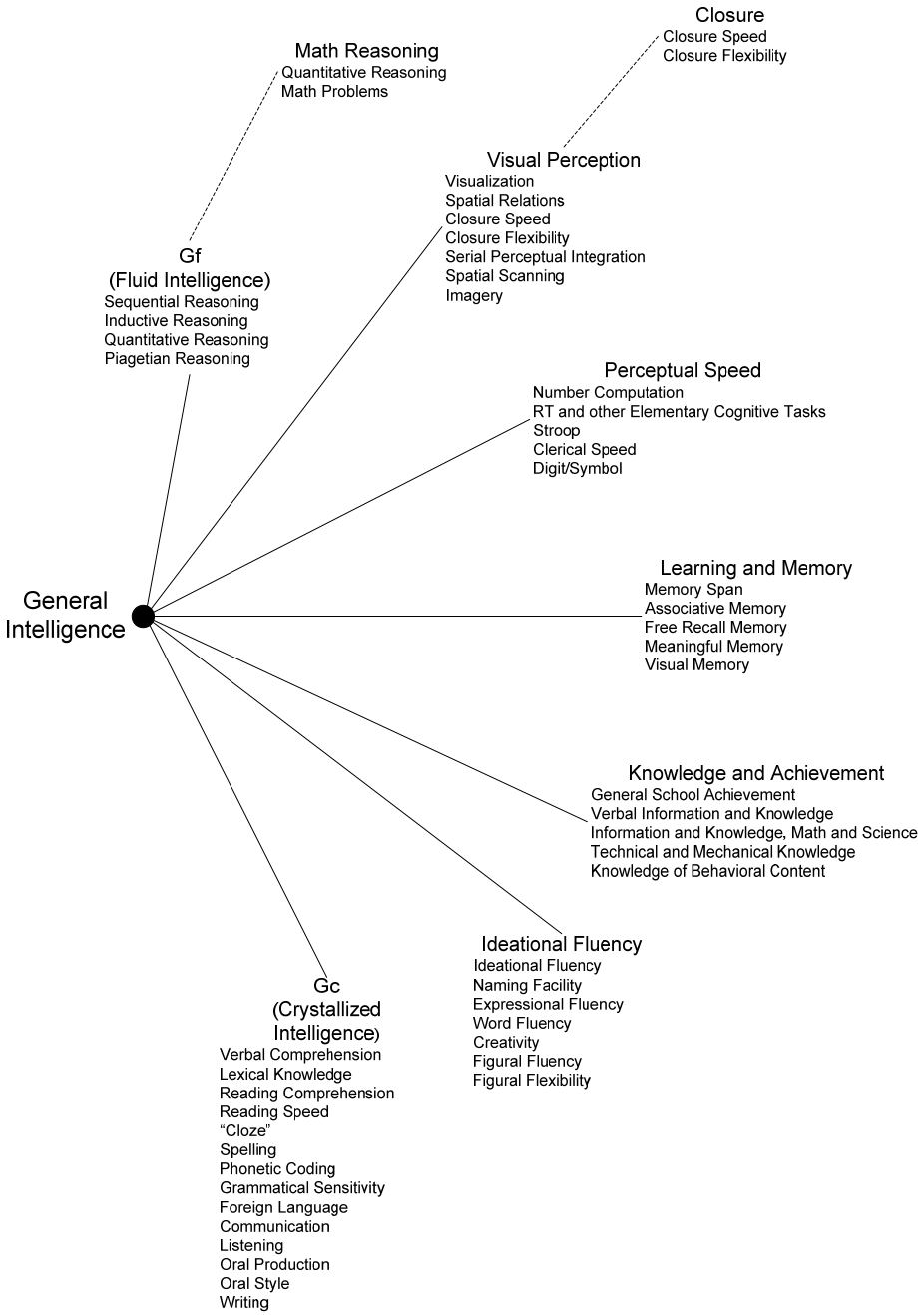
Scores on different tests of cognitive ability tend to be highly correlated, with half or more of the variance of diverse tests accounted for by a single general factor labeled "*g*" and

[94] Several psychologists have attempted to broaden the term intelligence to include other capacities. Most notably, Gardner [2004] suggests that the notion of intelligence should also include creativity and the ability to solve practical, real-world problems. Carroll [1993] includes in his theory of multiple intelligences, musical intelligence, kinaesthetic intelligence, and interpersonal and intrapersonal intelligence, among others.
[95] Wilhelm and Engle [2005].

more specific mental abilities loading on other factors.[96] *g* is widely interpreted as general mental ability.[97] An extreme version of *g*-theory that is no longer widely accepted is that *g* accounts for all of the correlations among different tests.[98] In this context, the order of a factor indicates its generality in explaining a variety of tests of cognitive ability chosen by an a priori notion of "construct validity," with different emphases (for example, verbal ability, numeracy, coding speed, and other tasks). A first-order factor is predictive in all cognitive tasks, $j=1,\ldots, J$ in equation (18) —*"g."* A lower-order factor is predictive in only some tasks . Lower-order factors can be correlated with the higher-order factors and may be correlated with each other. They have independent predictive power from the higher-order factors. Figure 3 reports one possible partition of general intelligence due to Ackerman and Heggestad [1997], who summarize the work of Carroll on the multiple facets of general intelligence.[99]

[96] Johnson, Bouchard, Krueger et al. [2004]; Jensen [1998]; Lubinski [2004]; Spearman [1904; [1927].
[97] Gottfredson [2002]
[98] See, e.g., Carroll [1993].
[99] Carroll's own organization of his evidence is somewhat different. He would organize the branches shown in Figure 3 as "lower order" manifestations of an hierarchically higher generalized and fluid intelligence.

Figure 3. An Hierarchical Scheme of General Intelligence and Its Components



Source: Recreated from Ackerman and Heggestad [1997].

Fluid vs. Crystallized Intelligence

There is less agreement about the number and identity of lower-order factors.[100] Carroll [1993] proposed a general intelligence factor, *g*, and several more specific second-order factors, including what Cattell [1971/1987] dubbed crystallized and fluid intelligence. Crystallized intelligence, Cattell proposed, comprises acquired skills and knowledge and thus is partly dependent upon educational opportunity and motivation. Fluid intelligence, by contrast, is a general "relation-perceiving ability" (p. 138). Cattell's student Rindermann [2007] elaborates:

> "*Fluid intelligence is the ability to "perceive complex relations, educe complex correlates, form concepts, develop aids, reason, abstract, and maintain span of immediate apprehension in solving novel problems in which advanced elements of the collective intelligence of the culture were not required for solution*" (p. 462).

In contrast, crystallized intelligence is the same class of skills, "but in materials in which past appropriation of the collective intelligence of the culture would give one a distinct advantage in solving the problems involved" (p. 462).

Carroll [1993] and Horn and McArdle [2007] summarize the large body of evidence against the claim that a single factor "*g*" is sufficient to explain the correlation structure of achievement and intelligence tests.[101] Two pieces of evidence are worth highlighting. First, crystallized intelligence tends to increase monotonically for most of the life cycle, whereas fluid

---

[100] Cattell [1971] analyzed 477 data sets and estimated a structure with *g* as the highest-order factor, eight second-order ability clusters, and over 70 more narrowly defined third-order abilities on a variety of different tests. Alternative hierarchical models, also with *g* as the highest-order factor, have been proposed (for example, Lubinski [2004]; Horn [1970]).

[101] Recent research by Ardila, Pineda and Rosselli [2000] shows that more than one factor is required to summarize the predictive power of cognitive tests in economic data. This could be due to the existence of multiple intellective factors or because personality factors affect the measurement of cognitive factors as we discuss later on in this section.

intelligence tends to peak in very early adulthood then to decline.[102] Second, the well-known Flynn effect, which documents the population-wide increase in performance on intelligence tests over the past half-century, is particularly dramatic for measures of fluid intelligence but much smaller for measures of crystallized intelligence.[103] SAT scores have declined rather than increased over the same period, requiring a renorming in the 1990s.

The relative weighting of fluid versus crystallized intelligence varies among tests according to the degree to which prior experience is crucial to performance. These factors operate as manifestations of the first-order factor, *g*, but contribute additional explanatory power to predicting some clusters of test score outcomes. Achievement tests, like the Armed Forces Qualifying Test used by economists and psychologists alike, are heavily weighted towards crystallized intelligence,[104] whereas tests like the Raven Progressive Matrices [1962] are heavily weighted towards fluid intelligence.[105] Several studies have shown that fluid intelligence is much more strongly related to *g* than are measures of crystallized intelligence.[106] Moreover, lay intuitions of intelligence (i.e., what most people mean by "being smart") correspond more closely with the ability to learn than with already acquired knowledge.[107] Thus, it seems to us useful to reserve the term "intelligence tests" for tests that primarily measure fluid intelligence, and the term "achievement tests" for tests that primarily measure crystallized intelligence.

---

[102] McArdle, Hamagami, Meredith et al. [2000].
[103] Dickens and Flynn [2001].
[104] Roberts, Goff, Anjoul et al. [2000].
[105] Raven, Raven and Court [1988]. Conti and Pudney [2007] uses data on intelligence and achievement tests across nations to show that a single factor accounts for 94-95 percent of the variance across both kinds of tests. The high correlation between intelligence and achievement tests is in part due to the fact that both require cognitive ability and knowledge, even if to different degrees, that common developmental factors may affect both of these traits, and that fluid intelligence promotes the acquisition of crystallized intelligence.
[106] Cattell [1971/1987]; Gustafsson [1988]; Kvist and Gustafsson [2007].
[107] Gottfredson [1998].

Predictive Validity of Tests of Cognition

How well do IQ and achievement tests predict success in life? This is a hard question to answer. Many different skills are required to achieve success in any task.[108] Different tasks in life require different skills in different degrees.[109] Table 3 shows the domains of validation and the estimated validities of a number of widely used tests of cognition. Notice that the domains of validation differ greatly. For IQ tests, the validities are usually established by comparing test scores with other test scores or with grades in school and not success in life. Nevertheless, it is well-established that standardized tests of ability and achievement predict objectively measured academic, occupational, and life outcomes.[110]

The SAT college entrance exam is moderately successful in predicting grades in college—which was what the SAT was designed to do.[111] However, high school grades are better predictors of college performance.[112] Geiser and Santelices [2007] The rival American College Test (ACT) is validated in a similar fashion but uses somewhat broader measures of college performance, such as grades in higher years of college rather than just freshman year grades.[113] The Graduate Record Exam is validated by performance in graduate school.[114] The Armed Forces Qualifying Test (AFQT) is validated by performance in the military. Performance is measured by success in military training schools and performance standardized tasks like fixing a rifle or repairing a radio.[115] One can interpret *The Bell Curve* by Herrnstein and Murray as conducting a validity study of that test using real world outcomes of the sort shown in Table 1.

---

[108] Mandelbrot [1963].
[109] See, e.g., Roy [1951], Mandelbrot [1961], Willis and Rosen [1979], Heckman and Sedlacek [1985], and Heckman, Stixrud and Urzua [2006].
[110] Kuncel, Ones and Sackett [2010].
[111] See Young and Kobrin [2001].
[112] Bowen, Chingos and McPherson [2009a], Geiser and Santelices [2007].
[113] ACT [2007].
[114] Kuncel and Hezlett [2007].
[115] See McHenry, Hough, Toquam et al. [1990].

The correlation of AFQT with wages is a moderate $r = 0.3$. The General Aptitude Test Battery (GATB) predicts success at work as measured by supervisor ratings in over 12,000 occupations and participation in training programs.[116]

[116] Schmidt and Hunter [1983; [1998], Hartigan and Wigdor [1989] and McHenry, Hough, Toquam et al. [1990].

Table 3. Predictive Validities of Various Tests of Fluid and Crystallized Intelligence

Cognitive Achievement and IQ Tests				
Test	**Domain over which it is validated**	**Estimated Validities**	**Source**	**Notes**
SAT	First year college GPA	0.35 to 0.53	Validity of the SAT for Predicting First-Year College Grade Point Average	
ACT	Grades in early years of college	0.42	ACT Technical Manual	
Stanford-Binet	Correlations with other intelligence tests	0.77 to 0.87 with WISC-R	Rothlisburg (1987); Greene, Sapp, Chissom (1990)	
WISC (Wechsler Intelligence Scale for Children)	Correlations with academic achievement	WISC: 0.443 to 0.751 with WRAT tests, 0.482 to 0.788 with 1st grade grades, 0.462 to 0.794 with 2nd grade grades; WISC-R: 0.346 to 0.760 with WRAT tests, 0.358 to 0.537 with 1st grade grades, 0.420 to 0.721 with 2nd grade grades	Hartlage and Steele (1977)	WRAT = Wide Range Achievement Test; Ranges are given because correlations vary by academic subject
WAIS (Wechsler Adult Intelligence Scale)	Correlations with other intelligence tests, achievement tests, and outcomes	0.67 (median) with verbal tests, 0.61 (median) with nonverbal tests, 0.69 with education attained, 0.38 to 0.43 with college grades, 0.62 with high school grades, 0.14 with nursing grades	Feingold (1982)	
Raven's Standard Progressive Matrices	Correlations with other intelligence tests	0.74 to 0.84 with WAIS-R	O'Leary, Rusch, Guastello (1991)	
GATB (General Aptitude Test Battery)	Supervisor rating performance in training programs and in job performance	0.23 to 0.65	Hunter (1986)	Large range due to variety of jobs
ASVAB (Armed Services Vocational Aptitude Battery)	Performance in military training programs and military attrition rates	0.37 to 0.78 for training (mean=0.56); 0.15 for attrition	Schmidt (1988) for performance in training programs; Sticht et al (1982) for attrition rates	Large range in training correlations due to a variety of jobs
GED (General Educational Development)	Test difficulty is normed against graduating HS seniors. Test scores of high school seniors and grades of high school seniors	0.33 to 0.49 for HS Senior GPA	Technical Manual: 2002 Series GED Tests	
DAT (Differential Aptitude Tests)	Correlations with academic achievement	0.13 to 0.62 for college GPA	Omizo (1980)	Large range is due to varying validity of eight subtests of DAT
WIAT (Wechsler Individual Achievement Test)	Correlation with other achievement tests; teacher ratings of student achievement	0.80 with grade 4 CAT/2, 0.69 with grade 5 CAT/2, 0.83 with grade 6 CAT/2; 0.67 with teacher ratings	Michalko and Saklofske (1999)	CAT=California Achievement Test

Notes: Feingold [1982], Greene, Sapp and Chissom [1990], Hartlage and Steele [1977], Hunter [1986], Kobrin, Patterson, Shaw et al. [2008], Michalko and Saklofske [1996], O'Leary, Rusch and Guastello [2006], Omizo [1980], Rothlisburg [1987], Schmidt, Hunter and Larson [1988], Sticht [1982], American Council On Education [2007], American Council on Education [2009].

5.B. Personality Traits

As noted in Section 2, it is misleading to sharply contrast aspects of cognition with personality traits. Consider, for example, so-called "quasi-cognitive" traits (Kyllonen, Walters and Kaufman [2005]). These include creativity (Csikszentmihalyi [1996]), emotional intelligence (Mayer and Salovey [1997]), cognitive style (Stanovich [1999]; Perkins and Tishman [2001]), typical intellectual engagement (Ackerman and Heggestad [1997]), and practical intelligence (Sternberg [2000]). Furthermore, the Big Five factor of Openness to Experience overlaps, conceptually and empirically, with the construct of intelligence (McCrae and Costa [1997a]; Noftle and Robins [2007]). We note in Section 5.C that personality can affect performance on tests of fluid intelligence. Personality traits also affect acquired skills and knowledge (i.e., crystallized intelligence).[117] It is therefore not surprising that Conti, Heckman and Urzua [2010a] document a correlation of $r = 0.54$ between their cognitive and non-cognitive latent factors in the British Longitudinal Cohort (Conti, Heckman and Urzua [2010a; b]). This correlation is higher than what is reported in the psychology literature. For many personality traits and for measures of cognition that are based on fluid intelligence, the correlations are much closer to zero, as we note below.

Finally, consider the construct of *executive function*. "Cognitive control" and "executive function" are terms used interchangeably, primarily in the neuroscience literature. Both have been defined as the voluntary, effortful blocking of a habitual behavior in order to execute a less familiar behavior (Matsumoto and Tanaka [2004]). Some authors (e.g., Gray [2004]) also use the terms "cognitive control" and "self-control" interchangeably, though self-control is traditionally

[117] See Chamorro-Premuzic and Furnham [2005] for an extended discussion of this topic.

considered a personality trait rather than an aspect of cognition. While tasks requiring executive function are related to questionnaire measures of self-control, the size of these associations is only about $r = 0.11$ to 0.14.(Duckworth and Schulze [2009]).

Areas of the dorsolateral prefrontal cortex (PFC), in conjunction with the nearby anterior cingulate cortex (ACC), are now understood as responsible for "executive control" over lower order processes.[118] That is, executive control entails top-down, intentional control of behavior and is not necessary for the performance of simple, automatic tasks (Miller and Cohen [2001]). The PFC achieves structural and functional maturity later than other (e.g., sensorimotor) brain regions (Casey, Tottenham, Liston et al. [2005]). Specific executive functions attributed to the PFC include abstract reasoning, planning, decision making, working memory, attention, conflict monitoring, task switching, and inhibition of prepotent (i.e., dominant, habitual) impulses. While many functions have been attributed to the PFC, Miller [2000] has observed that "there is little agreement on the cardinal prefrontal functions" (p. 449). Nevertheless, there is some consensus that one can distinguish between working memory on the one hand, and response inhibition and task switching on the other (Garon, Bryson and Smith [2008], Miyake, Friedman, Emerson et al. [2000]). This distinction is important because working memory is highly related to fluid intelligence.[119] Thus, working memory is a common component of the constructs of both executive function and general intelligence.[120]

While the construct of executive function demonstrates the inadequacy of terms such as "cognitive" and "non-cognitive", many personality traits nevertheless are conceptually and empirically easily distinguished from general cognitive ability. Most personality traits are in fact

[118] Notably, the volume of dorsolateral prefrontal cortex (PFC) is correlated with Big Five Conscientiousness (DeYoung, Hirsh, Shane et al. [2010]).
[119] Carpenter, Just and Shell [1990]; Heitz, Unsworth and Engle [2005].
[120] Friedman, Miyake, Corley et al. [2006].

very weakly correlated with IQ (Webb [1915]; McCrae and Costa [1994]; Stankov [2005]; Ackerman and Heggestad [1997]). Thus, regardless of the terms used to describe individual differences that determine life outcomes, one thing is clear: human capital entails more than intelligence. Personality traits, however defined do matter and they can be distinguished from intelligence.

5.C. Operationalizing the Concepts

Intelligence tests are routinely used in a variety of settings including business, education, civil service, and the military.[121] Psychometricians attempt to use test scores to measure a factor (a component of T_ℓ). The working hypothesis in the intelligence testing business is that specific tests measure only a single component of T_ℓ, and that tests with different "content domains" measure different components. We first discuss the origins of the measurement systems for intelligence and we then discuss their validity.[122]

IQ Tests

Modern intelligence tests have been used for just over a century, beginning with the decision of a French minister of public instruction to identify retarded pupils in need of specialized education programs. In response, Alfred Binet created the first IQ test.[123] Other pioneers in intelligence testing include Cattell [1890] and Galton [1883], both of whom developed tests of basic cognitive functions (for example, discriminating between objects of different weights). These early tests were eventually rejected in favor of tests that attempt to tap higher mental processes. Terman [1916] adapted Binet's IQ test for use with American populations. Known as the

[121] Siegler [1992] provide a detailed overview of the different types of applications of psychological testing.
[122] See Roberts, Markham, Matthews et al. [2005] for a more complete history of intelligence testing.
[123] In 1904, *La Société Libre pour l'Etude Psychologique de l'Enfant* appointed a commission to create a mechanism for identifying these pupils in need of alternative education led by Binet. See Herrnstein and Murray [1994] for an overview of Binet's life and work.

Stanford-Binet IQ test, Terman's adaptation was, like the original French test, used primarily to predict academic performance. Stanford-Binet test scores were presented as ratios of mental age to chronological age multiplied by 100. IQ scores centered at 100 as the average are now conventional for most intelligence tests.

Wechsler [1939] noted two major limitations of the Stanford-Binet test. First, it was overly reliant on verbal skills and, therefore, dependent upon formal education. Second, the ratio of mental to chronological age was an inappropriate metric for adults (Boake [2002]). Wechsler created a new intelligence test battery divided into verbal subtests (e.g., similarities) and performance subtests (e.g., block design, matrix reasoning). He also replaced the ratio IQ score with deviation scores that have the same normal distribution at each age. This test, the Wechsler Adult Intelligence Scale (WAIS) – and, later, the Wechsler Intelligence Scale for Children (WISC) – produces two different IQ subscores, verbal IQ and performance IQ, which sum to a full-scale IQ score. The WAIS and the WISC have for the past several decades been by far the most commonly used IQ tests.

Similar to Wechsler's Matrix Reasoning subtest, the Raven Progressive Matrices test is a so-called "culture-free" IQ test because it does not depend heavily on verbal skills or other knowledge explicitly taught during formal education. Each matrix test item presents a pattern of abstract figures.[124] The test taker must choose the missing part.[125] If subjects have not had exposure to such visual puzzles, the Raven test is an almost pure measure of fluid intelligence. However, the assumption that subjects are unfamiliar with such puzzles is not typically tested. It is likely that children from more educated families or from more developed countries have more

[124] See John and Srivastava [1999a] for a discussion of the Raven test.
[125] See Figure A1 in Section A5 of the web appendix.

exposure to such abstract puzzles (Blair [2006]). Our view is that to varying degrees, IQ and achievement tests reflect fluid intelligence, crystallized intelligence, and motivation.

5.D. Personality Constructs

Dominant theories of personality assume a hierarchical structure analogous to that found for intelligence. However, despite early efforts to identify a *g* for personality (for example, Webb [1915]), even the most parsimonious personality models incorporate more than one factor. The most widely accepted taxonomy of personality traits is the Five-Factor model.[126] The so-called Big Five factors are obtained from conventional factor analysis using a version of (18) where the "tests" are measures of different domains of personality based on observer reports or self reports.

The five-factor model has its origins in Allport and Odbert's [1936] lexical hypothesis, which posits that the most important individual differences are encoded in language. Allport and Odbert combed English dictionaries and found 17,953 personality-describing words, which were later reduced to 4,504 personality-describing adjectives. Subsequently, several different psychologists working independently and on different samples concluded that personality traits can be organized into five superordinate factors.

Table 4 reviews the Big Five factors introduced in Section 2. It summarizes the 30 lower-level facets (six facets for each of five factors) identified in the Revised NEO Personality Inventory (NEO-PI-R, Costa and McCrae [1992b]), shorthand for *N*euroticism, *E*xtroversion, *O*penness to Experience—*P*ersonality *I*nventory—*R*evised. Of course, these lower-level facets (e.g., "impulsive") can be further subdivided into even more narrow traits ("impulsive about junk food," "impulsive about smoking"). The more narrowly defined a trait, the more specific the contexts in which the trait operates. In parentheses in the third column of this table, we have

[126] See Goldsmith, Buss, Plomin et al. [1987] for an historical overview of the development of the Big Five.

included a strongly related trait adjective from an adjectival personality questionnaire. In the fourth column of Table 4, we present other traits in each family. In the fifth column, we propose relations of the Big Five to children's temperament traits studied by developmental psychologists.

Table 4. The Big Five domains and their facets

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

Notes: Facets specified by the NEO-PI-R personality inventory (Costa and McCrae [1992b]). Trait adjectives in parentheses from the Adjective Check List (Gough and Heilbrun Jr. [1983]). *These temperament traits may be related to two Big Five factors.
Source: Table adapted from John and Srivastava [1999b].

Temperament is the term used by developmental psychologists to describe the behavioral tendencies of infants and children.[127] Because individual differences in temperament emerge so early in life, these traits have traditionally been assumed to be biological (as opposed to environmental) in origin.[128] However, findings in behavioral genetics suggest that, like adult personality, temperament is only partly heritable, and as discussed in Section 8, both adult and child measured traits are affected by the environment.

Temperament is studied primarily by child and developmental psychologists, while personality is studied by adult personality psychologists. The past decade has seen some convergence of these two research traditions, however, and there is evidence that temperamental differences observed during the preschool years anticipate adult personality and interpersonal functioning decades later (for example, Caspi [2000]; Newman, Caspi, Moffitt et al. [1997]; Shiner and Caspi [2003]). Column 5 of Table 4 displays temperament traits that have been associated both theoretically and empirically with adult personality traits.

Historically, many temperament researchers examined specific lower-order traits rather than broader, higher-level factors that characterize studies of adult intelligence and personality.[129] Shiner [1998] suggests that "there is therefore a great need to bring order to this vast array of studies of single lower-level traits" (p. 320). Recently, taxonomies of temperament

---

[127] See Caspi and Shiner [2006] and Zentner and Bates [2008] for a discussion of varying perspectives on temperament, including a summary of points where major theorists converge.

[128] Indeed, some psychologists use the term "temperament" to indicate all aspects of personality that are biological in origin. They study temperament in both children and adults.

[129] Measuring temperament presents unique methodological challenges. Self-report measures, by far the most widely used measure for adult personality, are not appropriate for young children for obvious reasons. One strategy is to ask parents and teachers to rate the child's overt behavior (for example, California Child Q-sort), but informants can only guess what a child might be thinking and feeling. Infants present a special challenge because their behavioral repertoire is so limited. One strategy is to place infants in a standard situation and code reactions under a standardized scenario (for example, the Strange Situation, which is used to distinguish infants who are securely attached to their caregiver versus insecurely attached). Young children can be interviewed using puppets or stories. For obvious reasons, all measures of temperament are more difficult and more expensive to collect than adult self-report measures. This may explain their absence in large-sample studies.

have been proposed that group lower-order traits into higher-order dimensions; several of these taxonomies resemble the Big Five (for example, John, Caspi, Robins et al. [1994]; Putnam, Ellis and Rothbart [2001]; Rothbart, Ahadi and Evans [2000]; Shiner and Caspi [2003]). However, compared to adults, there seem to be fewer ways that young children can differ from one another. Child psychologists often refer to the "elaboration" or "differentiation" of childhood temperament into the full flower of complex, adult personality. The lack of direct correspondence between measures of temperament and measures of adult personality presents a challenge to researchers interested in documenting changes in personality over the full life cycle. Developing the required measures is an active area of research.

Alternatives to the Big Five

The Five-Factor model is not without its critics. Alternative systems have been proposed. For example, Eysenck [1991] offers a model with just three factors (i.e., Neuroticism, Extraversion, and Psychoticism). Cloninger [1987] and Tellegen [1985] offer different three-factor models. Figure 4 shows the commonalities across some competing taxonomies and also areas of divergence. Solutions with more factors can increase the prediction of outcomes including job performance, income, and change in psychiatric status (Mershon and Gorsuch [1988]). On the other hand, more parsimonious models in which the five factors are reduced to two "metatraits" have also been suggested (Digman [1997]). Further, the facet-level organization of any given Big Five factor is subject to debate and controversy.

Figure 4. Competing taxonomies of personality

Evsenck Big Three	Costa & McCrae NEO-PRF Big Five	Tellegen MPQ	Zuckerman	Cloninger	Big Nine
Neuroticism	**Neuroticism**	**Negative Emotionality**	**Neuroticism-Anxiety**	**Harm Avoidance**	**Adjustment**
Anxious	Anxiety	Stress reaction			
	Vulnerability				
Depression	Depression				
Guilt-feeling					
Low self-esteem					
Tense					
Irrational					
Shy	Self-consciousness	Alienation			
Moody					
Emotional	*Impulsiveness*				
Psychotism					
Aggressive	Hostility	Agression	**Aggression-Hostility**	**Cooperativeness**	**Agreeableness**
Cold					
Egocentric					**Rugged Individualism**
Impersonal	**Agreeableness**				
Anti-social	Altruism				
Unempathic	Compliance				
Tough-minded	Tendermindedness				
	Straightforwardness				
	Trust				
	Modesty				
	Conscientiousness	**Constraint**		**Self-Directedness**	**Dependability**
	Deliberation				
	Dutifulness				
Impulsive	Self-discipline	Control			**Locus of Control**
	Order				
	Competence			**Persistence**	**Achievement**
	Achievement striving				
		Traditionalism			
Extraversion	**Extraversion**				
Sensation-seeking	Excitement seeking	*Harm Avoidance*	**Impulsive Sensation Seeking**	**Novelty Seeking**	
Venturesome					
Active	Activity		**Activity**		
Surgent					
Carefree					
		Positive emotionality			
		Achievement		**Reward Dependence**	
		Social Closeness			
Sociable	Gregariousness		**Sociability**		**Affiliation**
Lively					
Assertive	Assertiveness				
Dominant		Social Potency			**Potency**
	Positive emotions	Well-being			
	Warmth				
					Intelligence
	Openness			**Self-Transcendence**	
	Fantasy				
	Aesthetics	Absorption			
	Feelings				
	Actions				
	Ideas				
	Values				

Source: Figure reproduced from Bouchard and Loehlin [2001], with kind permission from Springer Science and Business Media.

Recent research suggests that the rush to accept the Big Five may be premature.[130] The first studies of the Big Five were based primarily on English speaking samples. And, although the Big Five structure appears to replicate across many different cultures (McCrae and Costa [1997b]), examination of the structure of natural language lexicons that derive from many different cultures show that taxonomies known as the Big Six (Ashton, Lee, Perugini et al. [2004]) or the Multi-Language Seven (ML7; Saucier [2003]), may better represent the personality domain. While they add one or two dimensions to the Big Five and shift the meaning of the Big Five slightly, they are not very different from the Big Five.[131] However, the Big Six and ML-7 appear to replicate more readily across different cultures than the Big Five.

One of the most stinging criticisms of the Five-Factor model is that it is atheoretical (Block [1995]). While research is under way on determining the neural substrates of the Big Five (see Canli [2006] and DeYoung, Hirsh, Shane et al. [2010]), the finding that descriptions of behavior as measured by tests, self-reports, and reports of observers cluster reliably into five groups has not so far been satisfactorily explained by a basic theory. McAdams [1992; [2006] suggests that while the Big Five usefully organizes broad dispositional traits by which individuals can be compared to one another, it does not provide a theoretical framework for understanding all dimensions of personality.

A more pragmatic concern is that restricting the conceptualization and measurement of personality to the Big Five factors can obscure relations between specific facets and outcomes Hough [1992]; Hough and Oswald [2000]; Paunonen and Ashton [2001]. Given that each Big Five factor is heterogeneous, the predictive validities for a given outcome of particularly consequential facets are diluted when analyses consider only factor-level scores. For instance,

[130] This discussion draws heavily on Roberts [2006].
[131] See Roberts [2006] for a description of these shifts.

Paunonen and Ashton [2001] compared Big Five Conscientiousness and Openness to Experience with two related facets, need for achievement and need for understanding. In each comparison, the lower-level facet predicted course grades among undergraduates better than its higher-level factor measure.

The Five-Factor model is largely silent on an important class of individual differences that do not receive much attention in the recent psychology literature: motivation, i.e., the η in utility function (4). The omission of motivation (that is, what people value or desire) from measures of Big Five traits is not complete, however. The NEO-PI-R, for example, includes as a facet "achievement striving". Individual differences in motivation are more prominent in older (now rarely used) measures of personality. The starting point for Jackson's Personality Research Form (PRF; Jackson [1974]), for example, was Murray's [1938] theory of basic human drives. Included in the PRF are scales for (need for) play, order, autonomy, achievement, affiliation, social recognition, and safety. The Schwartz Values Survey (Schwartz [1992]) is another self-report measure of motivation which yields scores on ten different motivations including power, achievement, benevolence, and conformity. Some motivation theorists believe that one's deepest desires are unconscious and, therefore, may dispute the practice of measuring motivation using self-report questionnaires (see McClelland, Koestner and Weinberger [1989]). For a brief review of this debate and an overview of how motivation and personality trait measures differ, see Roberts, Harms, Smith et al. [2006].

A practical problem facing the analyst who wishes to measure personality is the multiplicity of personality questionnaires. The proliferation of personality measures reflects, in part, the more heterogeneous nature of personality in comparison to cognitive ability, although,

as we have seen, various types of cognitive ability have been established in the literature.[132] The panoply of measures and constructs also points to the relatively recent and incomplete convergence of personality psychologists on the Big Five model, as well as the lack of consensus among researchers about identifying and organizing lower-order facets of the Big Five factors (see DeYoung [2007] and Hofstee, de Raad and Goldberg [1992]). For example, some theorists argue that impulsivity is a facet of Neuroticism (Costa and McCrae [1992b]), others claim that it is a facet of Conscientiousness (Roberts, Chernyshenko, Stark et al. [2005]), and still others suggest that it is a blend of Conscientiousness, Extraversion, and perhaps Neuroticism (Revelle [1997]). Figure 4 shows in italics facets whose classification is in debate. Another reason for the proliferation of measures is the variety of methodologies for verifying tests discussed in Section 4.

5.D.1. Self-Esteem and Locus of Control are related to Big Five Emotional Stability

The traits of self-esteem and locus of control deserve special attention given their inclusion in many large-sample longitudinal studies.[133] Self-esteem refers to an individual's subjective estimation of his or her own worth. An example item from the widely-used Rosenberg Self-Esteem Scale (Rosenberg [1989]) asks respondents to indicate their agreement with the statement, "I feel that I am a person of worth, at least on an equal plane with others." Locus of control refers to one's belief about whether the determinants of one's life events are largely internal or external. Those with an internal (as opposed to external) locus of control believe that life events are typically caused by their own actions. An example item from the widely-used Rotter Locus of Control Scale (Rotter [1966]) requires respondents to choose between "Many of

[132] See, for example, Carroll [1993].
[133] See, e.g., NLSY79 based studies Heckman, Stixrud and Urzua [2006] and Heckman, Humphries, Urzua et al. [2010].

the unhappy things in people's lives are partly due to bad luck" and "People's misfortunes result from the mistakes they make."

For the most part, researchers who study self-esteem and locus of control have carried out their work in isolation of each other and without reference to the Big Five taxonomy. However, Judge and colleagues (Judge, Bono, Erez et al. [2005]; Judge, Erez, Bono et al. [2002]; Judge and Hurst [2007b]) have proposed that locus of control, self-esteem, and Big Five Emotional Stability are indicators of a common construct, termed *core self-evaluations*. They point out that measures of these three traits, as well as generalized self-efficacy (the belief that one can act effectively to bring about desired results) demonstrate high convergent validity, poor discriminant validity, and poor incremental predictive validity. Positive core self-evaluations indicate a generally positive and proactive view of oneself and one's relationship to the world. Accordingly, we have, in Table 4, associated aspects of core self-evaluations with the Big Five factors of Neuroticism and Emotional Stability.

5.D.2. Relating the Big Five to Measures of Psychopathology

Psychopathology is defined by the APA dictionary as "patterns of behavior or thought processes that are abnormal or maladaptive." Very recently, several attempts have been made to integrate taxonomies of psychopathology and normal personality into a single framework. In some cases, we may think of psychopathologies as extreme expressions of personality traits. Used interchangeably with the terms mental illness and mental disorder, psychopathology is studied by psychiatrists and clinical psychologists. Historically, the study of psychopathology was carried out in near complete isolation from the study of "normal" variation in personality.

The Diagnostic and Statistical Manual (DSM) of the *American Psychiatric Association* distinguishes between Axis I disorders, which are acute disorders requiring clinical attention

(e.g., depression, schizophrenia) and Axis II disorders, personality disorders that are more chronic and, generally, less impairing of overall functioning. Research has documented that Big Five Neuroticism is a non-specific correlate of various Axis I disorders, and that various other reliable associations can be documented (e.g., the positive emotionality facet of Extraversion is associated with bipolar disorder); however, the direction of causality is difficult to ascertain in what have typically been cross-sectional studies (Bagby, Bindseil, Schuller et al. [1997]; Cloninger, Svrakic, Bayon et al. [1999]; Gunderson, Triebwasser, Phillips et al. [1999]).

More research has examined relations between Axis II disorders and normal personality variation. For example, Markon, Krueger and Watson [2005] proposed a Big Four taxonomy (the Big Five minus Openness to Experience). Watson, Clark and Chmielewski [2008] proposed that a fifth factor called Oddity was needed to model traits related to eccentricity. Others have argued that the Big Five structure itself, without modification, can account for Axis II personality disorders (Widiger and Costa Jr. [2002]; Widiger, Trull, Clarkin et al. [2002]). For instance, Widiger, Trull, Clarkin et al. [2002] suggest that all Axis II personality disorders can be "translated as maladaptively extreme variants of the 30 facets" of Big Five personality factors. Table 5 summarizes their discussion relating the measures of the Big Five to measures of psychopathology. See also Costa, McCrae and Siegler [1999] for an empirical study associating facets of the Big Five to Axis II personality disorders. Most personality disorders are described by multiple personality factors.[134]

[134] Schizotypal and avoidant personality disorders combine Neuroticism with low Extraversion. Dependent personality disorders combine Neuroticism with Agreeableness.

Table 5. *DSM-IV* Personality Disorders and the Five-Factor model

	PRN	SZD	SZT	ATS	BDL	HST	NCS	AVD	DPD	OCP
Neuroticism (vs. **emotional stability**)										
Anxiousness (vs. unconcerned)			H	L	H			H	H	H
Angry hostility (vs. dispassionate)	H			H	H		H			
Depressiveness (vs. optimistic)			H		H					
Self-consciousness (vs. shameless)				L	H	L	L	H	H	
Impulsivity (vs. restrained)				H	H	H				L
Vulnerability (vs. fearless)				L	H			H	H	
Extraversion (vs. **introversion**)										
Warmth (vs. coldness)	L	L	L				L		H	
Gregariousness (vs. withdrawal)	L	L	L	H		H		L		
Assertiveness (vs. submissiveness)				H			H	L	L	
Activity (vs. passivity)		L		H		H				
Excitement seeking (vs. dullness)		L		H		H	H	L		L
Positive emotionality (vs. anhedonia)		L	L			H				
Openness (vs. **closedness**)										
Fantasy (vs. concrete)						H				
Aesthetics (vs. disinterest)										
Feelings (vs. alexithymia)		L			H	H	L			L
Actions (vs. routine)	L	L		H	H	H	H	L		L
Ideas (vs. closed-minded)			H							L
Values (vs. dogmatic)	L									L
Agreeableness (vs. **antagonism**)										
Trust (vs. mistrust)	L		L	L	L	H	L		H	
Straightforwardness (vs. deception)	L			L			L			
Altruism (vs. exploitation)				L			L			
Compliance (vs. opposition, aggression)	L			L	L		L		H	
Modesty (vs. arrogance)				L			L	H	H	
Tender-mindedness (vs. tough-minded)	L			L			L			
Conscientiousness (vs. **disinhibition**)										
Competence (vs. ineptitude)									L	H
Order (vs. disordered)			L							H
Dutifulness (vs. irresponsibility)				L						H
Achievement striving (vs. lackadaisical)										H
Self-discipline (vs. negligence)				L		L				H
Deliberation (vs. rashness)				L	L	L				H

Notes: PRN, paranoid; SZD, schizoid; SZT, schizotypal; ATS, antisocial; BDL, borderline; HST, histrionic; NCS, narcissistic; AVD, avoidant; DPD, dependent; OCP, obsessive-compulsive. H, high; L, low.
Source: Reproduced from Widiger and Mullins-Sweatt [2009].

5.E. IQ and Achievement Test Scores Reflect Incentives and Capture Both Cognitive and Personality Traits

We now elaborate on the discussion of Section 3 on the difficulty of isolating a pure measure of intelligence. Performance on intelligence and achievement tests depends in part on certain personality traits of the test taker, as well as their motivation to perform.[135] A smart child unable

[135] It is likely that performance on personality tests can also depend on cognitive ability, but that is less well documented. For example, it is likely that more intelligent people can ascertain the rewards to performance on a personality inventory test. Motivation is sometimes, but not usually, counted as a personality trait.

to sit still during an exam or uninterested in exerting much effort can produce spuriously low scores on an IQ test.

Almost 40 years ago, several studies called into question the assumption that IQ tests measure *maximal* performance (that is, performance reflecting maximal effort).[136] These studies show that among individuals with low IQ scores, performance on IQ tests could be increased up to a full standard deviation by offering incentives such as money or candy, particularly on group-administered tests and particularly with individuals at the low-end of the IQ spectrum.[137] Engaging in complex thinking is effortful, not automatic (Schmeichel, Vohs and Baumeister [2003]), and therefore motivation to exert effort affects performance. Zigler and Butterfield [1968] found that early intervention (nursery school, for example) for low-SES kids may have a beneficial effect on motivation, not on cognitive ability per se. In their study, the benefits of intervention (in comparison to a no-treatment control group) on IQ were not apparent under testing conditions where motivation to perform well were maximized. Raver and Zigler [1997] present further evidence on this point.

Table 6 summarizes evidence that extrinsic incentives can substantially improve performance on tests of cognitive ability, especially among low-IQ individuals.[138]

[136] Carroll [1993] does not accept this definition of IQ.
[137] The incentives for invoking effort vary across studies.
[138] The studies do not include direct measures of personality traits.

Table 6. 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 (avg 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 (avg IQ = 92.8)	Tokens given in experimental condition for right answers exchangeable for prizes	$t = 5.9$	
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 cdtn; 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 0.33 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 standardconditions 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 non-nursery groups was attributable solely to motivational factors…" (p. 10)
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)

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. $t(60) = 10.67$, $p < 0.001$	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

Segal [2008] shows that introducing performance-based cash incentives in a low-stakes administration of the coding speed test of the Armed Services Vocational Battery (ASVAB) increases performance substantially among roughly one-third of participants. Less conscientious men are particularly affected by incentives. Segal's work and a large body of related work emphasize heterogeneity in the motivations that affect human performance. Borghans, Meijers and ter Weel [2008] show that adults spend substantially more time answering IQ questions when rewards are higher, but subjects high in Emotional Stability and Conscientiousness are less affected by these incentives. Similarly, Pailing and Segalowitz [2004] find that an event-related potential (ERP) indexing the emotional response to making an error increases in amplitude when

incentives are offered for superior test performance.[139] This effect is smaller for individuals high in Conscientiousness and Emotional Stability. Thus, IQ scores do not accurately reflect maximal intellectual performance for individuals who are low in Conscientiousness and Emotional Stability. Performance on IQ tests encodes, in part, how effective persons may be in application of their intelligence, that is, how people are likely to perform in a real-world setting. However, it is far from obvious that motivation on an exam and motivation in a real-world situation are the same.

Like low motivation, test anxiety can significantly impair performance (Hembree [1988]). That is, subjects do worse when they worry excessively about how they are performing and when their autonomic nervous system over-reacts by increasing perspiration, heart rate, and so on. Because individuals who are higher in Big Five Neuroticism are more likely to experience test anxiety, there is another reason, beyond incentives, why Emotional Stability can impact IQ scores (Moutafi, Furnham and Tsaousis [2006]).

Many IQ tests require factual knowledge acquired through schooling and life experience, which are, in part, determined by the motivation, curiosity, and persistence of the test taker. Thus, personality traits can also affect IQ scores indirectly through the knowledge acquired by individuals who are higher in Big Five Openness to Experience and Big Five Conscientiousness. Cunha and Heckman [2008] show a correlation between cognitive and personality factors of the order of $r = 0.3$. Hansen, Heckman and Mullen [2004], and Heckman, Stixrud and Urzua [2006] show how schooling and other acquired traits substantially *causally* affect measured cognitive and personality test scores. We discuss this research in Section 8. Cattell's 1971] investment theory [1971] anticipates recent findings that knowledge and specific complex skills

[139] An ERP is an electrophysiological response of characteristic form and timing to a particular category of stimuli.

depend not only on fluid intelligence but also on the cumulative investment of effort and exposure to learning opportunities.

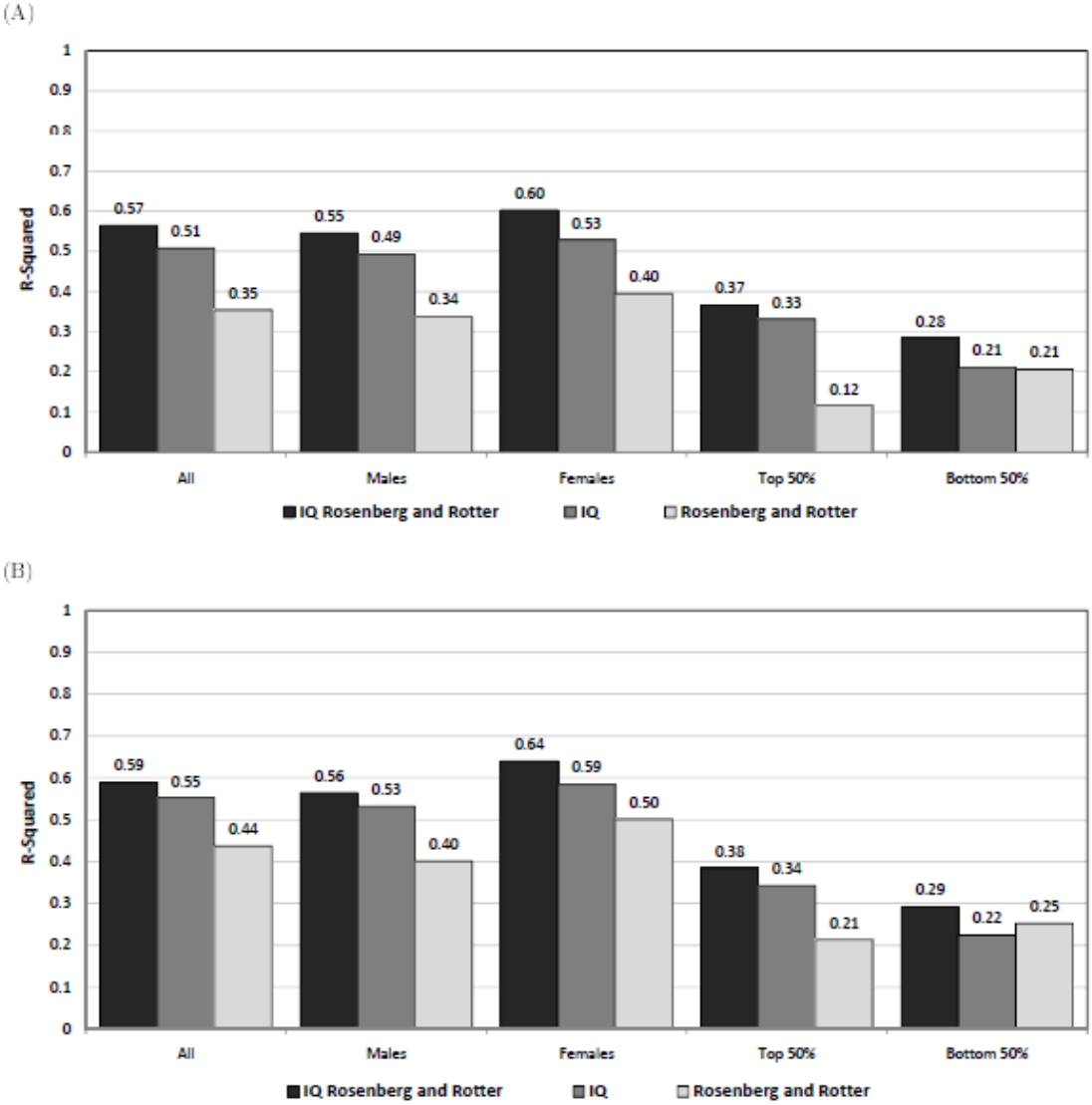
How, then, should one interpret a low IQ score? Collectively, the evidence surveyed here suggests that IQ test performance reflects not only pure intelligence, but also personality traits (including anxiety), intrinsic motivation, and reactions to extrinsic incentives to perform well, as indicated in our discussion of Section 3. It also reflects the knowledge acquired up to the date of the test, which reflects personality and motivational traits. The relative impurity of IQ tests likely varies from test to test and individual to individual. Little effort to date has been made to standardize the context and incentives of tests. To capture pure intelligence, it is necessary to adjust for incentives, motivations, and context in which the measurements are taken, using the framework discussed in Section 3.

Just as personality traits and incentives can affect IQ scores, they can also affect standardized achievement tests that are commonly used as proxies for pure intelligence. Figure 5 and Figure 6, below, show how scores on two achievement tests, the Armed Forces Qualifying Test (AFQT) and the Differential Aptitudes Test (DAT), are decomposed into IQ and personality (personality is measured by locus of control and self-esteem).[140] A substantial portion of the variance in both the AFQT score and DAT scores is independently explained by these two facets of personality. The variance explained is less than the variance independently explained by IQ scores, but it is still substantial. Furthermore, the facets are incrementally valid in that they explain the variance above and beyond the variance that IQ explains when all three are included in a regression. These findings caution the interpretation that these commonly used tests proxy mental ability; they likely proxy aspects of personality as well. Ironically, the measure of

[140] AFQT and DAT scores are highly correlated ($r = 0.76$). See Borghans, Golsteyn, Heckman et al. [2010], Kilburn, Hanser and Klerman [1998], Sticht [1995], and Wang [1993].

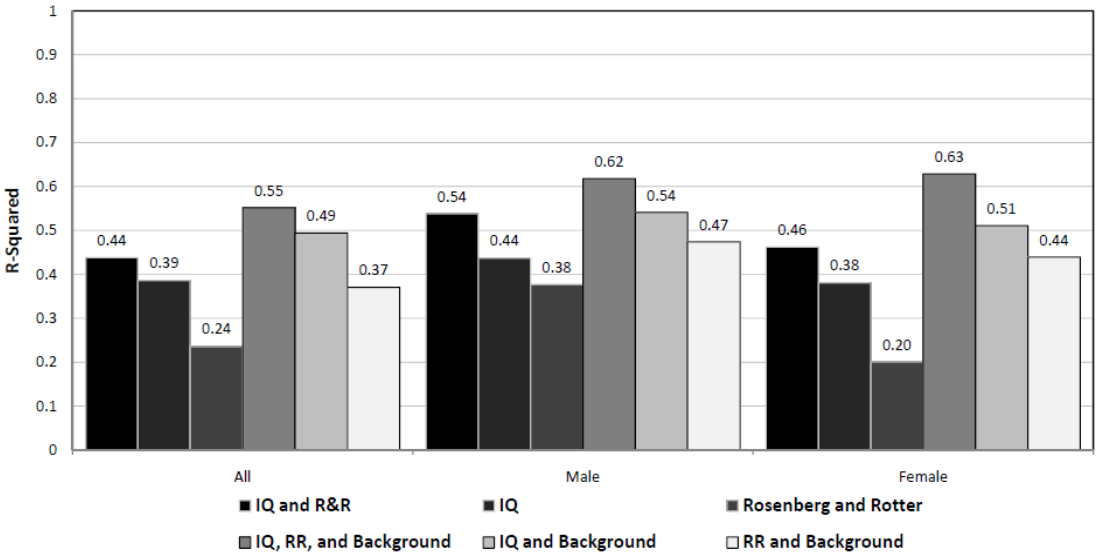
intelligence used by Herrnstein and Murray in *The Bell Curve* to predict a variety of social and economic outcomes is substantially affected by personality measures. We further discuss evidence about personality and standardized achievement tests in Section 7.

Figure 5. AFQT Score Decomposed by IQ, Self-Esteem and Locus of Control



Notes: The data come from the NLSY. Rosenberg, and Rotter were administered in 1979. The ASVAB was administered in 1980.To account for varying levels of schooling at the time of the test, scores have been adjusted for schooling at the time of the test conditional on final schooling using the method developed in Hansen, Heckman and Mullen [2004]. AFQT is constructed from the Arithmetic Reasoning, Word Knowledge, Numeric Operations, and Paragraph Comprehension ASVAB subtests. DAT and DAT percentile, IQ, and GPA are from high school transcript data. IQ is pooled across several IQ tests using IQ percentiles. GPA is the individual's core-subject GPA from each year of school. Sample excludes the military over-sample. Background variables include mother's highest grade completed, father's highest grade completed, southern residence at age 14, urban residence at age 14, living in a broken home at age 14, receiving newspapers in the household at age 14, receiving magazines in the household at age 14, and the household having a library card at age 14. Top 50% and Bottom 50% are based on AFQT scores from the population representative sample of NLSY79. Sample restricted to the non-military sub-sample.
Source: Borghans, Golsteyn, Heckman et al. [2010].

Figure 6. DAT Score Decomposed by IQ, Self-Esteem and Locus of Control



Notes: The data come from the NLSY79. Rosenberg and Rotter were administered in 1979. The ASVAB was administered in 1980. To account for varying levels of schooling at the time of the test, scores have been adjusted for schooling at the time of the test conditional on final schooling using the method developed in Hansen, Heckman and Mullen [2004]. AFQT is constructed from the Arithmetic Reasoning, Word Knowledge, Numeric Operations, and Paragraph Comprehension ASVAB subtests. DAT and DAT percentile, IQ, and GPA are from high school transcript data. IQ is pooled across several IQ tests using IQ percentiles. GPA is the individual's core-subject GPA from each year of school. Sample excludes the military over-sample. Background variables include mother's highest grade completed, father's highest grade completed, southern residence at age 14, urban residence at age 14, living in a broken home at age 14, receiving newspapers in the household at age 14, receiving magazines in the household at age 14, and the household having a library card at age 14. Sample restricted to the non-military sub-sample.
Source: Borghans, Golsteyn, Heckman et al. [2010].

5.F. The Evidence on the Situational Specificity Hypothesis

Since the publication of Mischel's [1968] book , psychologists have addressed the situational specificity hypothesis, i.e. that situations help explain the variations across people in actions, effort and behavior. Boiled down to its essence, this hypothesis says little more than that situations affect actions and efforts in a nonlinear fashion, i.e., that in equations (13) and (14), situational variables enter in a nonlinear fashion. This interaction effect is the Mischel-Shoda [1995] "if-then" relationship. To our knowledge, there are no studies available that parse the contributions of situations and traits to observed efforts, actions, and productivities.

Suppose that we observe the set of actions taken in performance of task j in situation h, $a_{i,j,h}$ which depend on θ and $e_{i,j}$:

$$a_{i,j,h} = \nu_{i,j,h}(\theta, e_{i,j}), \quad i \in \mathcal{A}, \quad h \in \mathcal{H}, \quad j \in \{1, \ldots, J\}.$$

The "average action" (i.e., the "personality") for the situation h in task j averages over, "integrates out" (or sums over) the θ and $e_{i,j}$:

(21) $$\bar{a}_{h,j} = \int_{\mathcal{S}_{j,h}(\theta, e_{i,j})} \nu_{i,j,h}(\theta, e_{i,j})\, g_{\theta,e}(\theta, e_{i,j} \mid h)\, d\theta\, de_{i,j}$$

where $\mathcal{S}_{j,h}(\theta, e_{i,j})$ is the support of $\theta, e_{i,j}$ for a given h, i.e., the domain of definition of $\nu_{i,j,h}$ function and $g_{\theta,e}(\theta, e_{i,j} \mid h)$, the density of $\theta, e_{i,j}$ given h and j. This is what psychologists mean by actions in a "typical situation" in task j, i.e., one that averages across θ and $e_{i,j}$. By the mean value theorem for integrals (Buck [2003]), $\bar{a}_{h,j}$ is the value of $a_{i,j,h}$ at a particular point of evaluation of θ and $e_{i,j}$.[141] One could use $\bar{a}_{h,j}$ as a definition of the situation h-typical

[141] The point of evaluation is implicitly determined by the integral.

action. Notice that if $v_{i,j,h}$ is separable in h,[142] there would be no person-situation interaction. Averaging over tasks $(j = 1,\ldots,J)$ in an analogous fashion produces the average action produced by a situation, $\bar{a}_h$.

By parallel reasoning, the average action for trait vector θ in task j can be defined as

(22) $$\bar{a}_{\theta,j} = \int_{\mathcal{S}_\theta(h,e_{i,j})} v_{i,j,h}(\theta, e_{i,j})\, g_{h,e_{i,j}}(h, e_{i,j} \mid \theta)\, dh\, de_{i,j}$$

where $\mathcal{S}_\theta(h, e_{i,j})$ is support of $v_{i,j,h}$ and $g_{h,e_{i,j}}(h, e_{i,j} \mid \theta)$, where $g_{h,e_{i,j}}$ is the density of $h, e_{i,j}$ given θ. One can think of $\bar{a}_{\theta,j}$ as one definition of the "enduring actions" of agents across situations in task j, i.e., the average personality for trait θ. One can average over tasks to produce an average action for trait vector θ. Again, if $v_{i,j,h}$ is separable in θ, the marginal effect of θ on actions is the same in all situations.

In light of the importance of existence of such interactions for personality psychology, it is surprising that there are so few quantitative estimates of their importance. A recent summary of the person-situation debate is provided in a series of papers in the *Journal of Research in Personality* [January, 2009, Vol. 43] that offer a retrospective on the controversy. Virtually all papers in that special issue acknowledge the existence of stable personality traits whose manifestations are tempered by situations and incentives. The editors of this special summarize a main message of the special issue with the following words:

> "*All personality psychologists should be unified when it comes to asserting that personality differences are worthy of scientific study, that individual differences*

[142] $v_{i,j,h}$ has the same marginal effect of h on actions regardless of the values of θ and $e_{i,j,h}$.

> *are more than just error variance and that not all behavior is simply a function of the situation*" – Lucas and Donnellan [2009, p. 147]

They go on to note that part of the intensity of disagreement on the issue arises from political disagreements. Situationists tend to be more liberal and interventionist and are more likely to assert that situations are major determinants of outcomes. Advocates of trait theory tend to be more conservative and more likely to assign responsibility of outcomes to individuals.

An important paper by Epstein [1979] defines stability of personality generated by traits across situations using measurements that average across tasks. He notes that in the presence of nonlinearities, agents with the same traits will take different actions in different situations. In four different studies, he presents compelling empirical evidence that, averaging over tasks and situations at a point in time, persons act in a predictable fashion with a high level of reliability (R^2 of 0.6-0.8) of average behavior ("measured personality") across situations. He uses a variety of measures based on objective behavior, self ratings, and ratings by others. He also establishes consistency (high levels of correlation) across the different types of measures. In any given situation, personality may not play a particularly powerful role but averaging over many situations, stable patterns emerge. Fleeson [2001] and Moskowitz [1982] present additional evidence on this question. Fleeson and Noftle [2008] summarize a substantial body of evidence on the stability of behaviors across tasks and situations, and the evidence of consistency of different measurements of personality (e.g., self reports, observer reports).

In one of the most ambitious recent studies of this question, Borkenau, Mauer, Riemann et al. [2004] establish a correlation of 0.43 of personality traits measured by the Big Five (self rated and observer rated) across 15 very different tasks. The range of correlations is from 0.51 to

0.29.[143] Wood and Roberts [2006] present further evidence on the persistence of traits across a variety of situations. Roberts [2009] provides a valuable overview of the latest research. Funder [2008] provides another useful overview of the debate and the evidence on the existence of a stable personality trait that at a point in time predicts behavior in a variety of different situations. Mischel's [1968] claim that there is no stable personality trait across situations does not hold up against a large array of data.

[143] Achenbach, McConaughy and Howell [1987] summarize correlations between children's problem behavior ratings by parents and teachers. Their meta-analysis produces an estimate of $r = 0.28$ and suggests consistency and variation in behavior and assessment across home and school situations. Whether this arises from parental bias or from situational specificity is not clear.

6. Personality and Preference Parameters

In Section 3, we argued that psychological traits are only identified from personality measures under strong assumptions. This section makes a similar point about preferences. In order to identify preferences in economics, strong assumptions are required about what agents care about, and the functional forms that characterize their preferences. This section demonstrates that the empirical evidence on the relationship between preference parameters and personality traits suggests that economists need to take into account how traits constrain agent choices, and the importance of accounting for nonseparabilities in preferences. If we do not take these features into account when measuring preference parameters, analysts may obtain misleading estimates of preference parameters. The literature on the links between preferences and psychological traits suggests some directions for future research directed toward understanding underlying preferences. It highlights the necessity of a coherent framework such as the model described in section 3.

6.A. Leisure, time, risk, and social preferences

In order to discuss the literature on the links between preferences and traits, we briefly define the aspects of preferences typically measured by economists. We demonstrate that there is nothing fundamental about these parameters, but rather that they represent convenient modeling choices. The behavioral economics literature and the literature on personality traits notes that the conventional parameters may not describe reality.

Measuring preferences boils down to measuring tradeoffs. Consider preferences over K goods for agent n

(23) $$U_n\left(X_1, X_2, ..., X_K\right)$$

Marginal rates of substitution are given by

$$MRS_{m,l}\left(X_1,\ldots,X_K\right)=\frac{\partial U_n\left(X_1,X_2,...,X_K\right)/\partial X_m}{\partial U_n\left(X_1,X_2,...,X_K\right)/\partial X_l}$$

The marginal rate of substitution is a fundamental concept in economic analysis. If preferences exist and are twice continuously differentiable, we can always define it for any set of choices with no additional assumptions. An important question is which basic choices preferences are defined over, and how stable they are, or rather, which range of choices will allow us to define marginal rates of substitutions as functions of stable parameters. Describing preferences amounts to defining the most fundamental life tradeoffs.

If we are to judge by the amount of interest in the empirical literature, the most important tradeoffs are the tradeoffs between leisure and consumption, between consuming now or later, and between less uncertainty and higher expected returns. In addition, social preferences and altruism have been studied, in particular the tradeoffs between equity, equality, efficiency, and own consumption, in the multiple ways that these may conflict. However, we stress that these aspects represent a choice by the scholars studying the subject. In a sense, this is not much different from the concept of operationalization and construct validity in the psychology literature. The relevant behaviors are by definition those which we choose to study. We look for our keys under our self-defined street lamps.

Preferences (23) can be general. We can define the preferences typically measured by economists by considering goods $X_1,...,X_K$ to be consumption and leisure, consumption at K different points in time, consumption in K states, or the consumption of K persons. For example, if X_1 is leisure and X_2 is consumption, we could measure leisure preference by $MRS_{1,2}$. Likewise, if we let the subscripts denote time, and n the preferences of the agent at time n, time discounting is described by $MRS_{m,l}^{n}$. In our model of personality, subscripts may

represent the performance on tasks and the exertion of effort. $MRS_{m,l}^{n}$ would then represent the tradeoff between better performance on a task and less disutility from effort. Note that the n superscript allows us to think of preferences as dependent on the time period, the current state, or the consumption of person n himself.

We are typically not only interested in how standard bundles of goods are traded off against each other, but in how lower variability in consumption is traded off against higher consumption. Risk preference, preferences for income distribution, and willingness to substitute over time reflect tradeoffs of this type. These preferences all depend on the curvature of the utility function; hence economists use measures such as absolute and relative risk aversion, and certainty equivalence which define curvature. These features are captured by the sensitivity of the MRS to changes in relative consumption:

$$\frac{\partial MRS_{m,l}}{\partial X_m / X_l} .$$

A more direct measure of risk aversion is the certainty equivalent of a lottery. We could define social preferences in a similar way; letting the subscripts on goods denote how much an agent values the consumption of other agents, we could define an "inequality equivalent" as a measure of how much an agent is willing to pay to avoid inequality.

The marginal rate of substitution as well as its elasticity might depend on any number of things, and it may not be possible to parameterize it with simple models. A reasonable aim for measuring preferences might be to find relevant tradeoffs that allow analysts to represent these measures in terms of stable parameters. The evidence on personality presented in this section suggests that standard parameters such as the time discount factor in exponential discounting and risk tolerance do not satisfy this requirement.

6.B. Leisure, time, risk, and social preferences

The typical features of preferences in the economics literature that receive the most attention—time discounting, risk aversion, leisure preference, and social preferences—have analogues in the psychology literature. See Table 7 for an overview. Time preferences relate closely to psychological factors associated with impatience and ability to plan for the future. Using data from an experiment involving college students, Daly, Delaney and Harmon [2009] find that a factor that loads heavily on psychometric measures such as self-control, consideration of future consequences, elaboration of consequences, affective mindfulness, and Conscientiousness is negatively associated with the discount rate. Figure 7 reports correlations between experimental measures of time preference, Big Five factors, and measures of cognition. Interestingly only cognitive measures are correlated with time preference. Risk aversion is related to the personality traits of impulsivity and sensation seeking, a trait proposed by Zuckerman [1994] and defined as "the tendency to seek novel, varied, complex, and intense sensations and experiences and the willingness to take risks for the sake of such experience."[144] Sensation seeking predicts risky driving, substance use and abuse, smoking, drinking, unprotected sex, juvenile delinquency, and adult criminal behavior (see Zuckerman [2007] for a review). Further, the Balloon Analogue Risk Task (BART) (Lejuez, Read, Kahler et al. [2002], Lejuez, Aklin, Zvolensky et al. [2003]), a computer game in which participants make repeated choices between keeping a certain smaller monetary reward and taking a chance on an incrementally larger reward, correlate with real-world risk behaviors such as smoking, stealing, and not wearing a seatbelt. BART scores also correlate with sensation seeking. Preferences for leisure are closely related to several personality measures. The Big Five includes an

[144] See Zuckerman [1994].

Achievement Striving subscale of Conscientiousness, which describes ambition, the capacity for hard work, and an inclination toward purposeful behavior. Jackson's Personality Research Form [1974] includes an achievement scale measuring the aspiration to accomplish difficult tasks and to put forth effort to attain excellence, as well as an endurance scale, measuring willingness to work long hours and perseverance in the face of difficulty, and a play scale, measuring the inclination to participate in games, sports, and social activities "just for fun." A recent factor analysis of seven major personality inventories identified industriousness as one of six facets of Conscientiousness (Roberts, Chernyshenko, Stark et al. [2005]). Social preferences also have analogues in the personality literature. Warmth and gregariousness are facets of Extraversion, trust, altruism, and tender-mindedness are facets of Agreeableness, and hostility is a facet of Neuroticism. However, these facets have not yet been empirically linked to preference parameters. However, the broader Big Five traits have been found to be related to several measures of social preferences. Dohmen, Falk, Huffman et al. [2008] use an experimental measure of trust and find that people who are more conscientious or neurotic trust less, whereas people who are more agreeable or open to experiences trust more. Agreeableness and Conscientiousness are associated with more positive reciprocity and less negative reciprocity, whereas Neuroticism is associated with more negative reciprocity.

The empirical association between preferences and personality is not uniform across studies. Figure 7 and Figure 8 show correlations between experimental measures of time and risk preference, Big Five factors, and measures of cognition.[145] Cognitive measures are correlated with both risk and time preference. Risk preference is related to Openness to Experience. Other Big Five factors do not seem to be related to preference parameters. The

[145] Figures A 2 and A 3 in Section A6 of the Web Appendix display correlations among the survey measures in the GSOEP data used to generate Figure 7 and Figure 8.

empirical links between preference parameters and personality traits seem to depend heavily on the data used. In this section, we discuss how preferences are measures and how this may influence results.

Table 7. Standard preference parameters and conceptually similar measures in the psychology literature.

Preference parameter	Personality measures
Time preference	Conscientiousness
	self-control
	affective mindfulness
	consideration of future consequences
	elaboration of consequences
	time preference
Risk aversion	Impulsivity
	Sensation seeking
	Balloon Analogue Risk Task
Leisure Preference	Achievement Striving
	Endurance
	Industriousness
Social preference	Warmth
	Gregariousness
	Trust
	Altruism
	Tender-mindedness
	Hostility

Figure 7. Pairwise Correlations between Time Preference (Impatience), Personality, and Cognitive Ability for Males and Females

Time Preference							
0.008 0.86	O						
0.068 0.13	0.109 0.00	C					
0.066 0.14	0.416 0.00	0.105 0.00	E				
0.083 0.06	0.083 0.01	0.290 0.00	0.049 0.12	A			
0.060 0.18	0.193 0.00	0.064 0.04	−0.013 0.69	−0.012 0.71	N		
−0.119 0.01	0.162 0.00	−0.152 0.00	0.146 0.00	−0.137 0.00	−0.064 0.05	Symbol Test	
−0.101 0.03	0.154 0.00	−0.130 0.00	0.117 0.00	−0.183 0.00	−0.082 0.02	0.404 0.00	Word Test

Notes: O=Openness to Experience, C=Conscientiousness, E=Extraversion, A=Agreeableness, N=Neuroticism. The top value in each box is the pairwise correlation and the bottom value is p-value. Darker shaded boxes have lower *p*-values. The measures of the Big Five are based on 3 questions each. The measures of cognitive ability (symbol test and word test) are based on timed modules similar to the Wechsler Adult Intelligence Scale (WAIS). Time preference was elicited through a real-stakes experiment.
Source: Authors own calculations using experimental data similar to GSOEP.

Figure 8. Correlations between Risk Tolerance, Personality, and Cognitive Ability for Males and Females

Risk Tolerance							
0.136 0.00	O						
−0.072 0.13	0.109 0.00	C					
0.076 0.11	0.416 0.00	0.105 0.00	E				
−0.094 0.05	0.083 0.01	0.290 0.00	0.049 0.12	A			
−0.029 0.54	0.193 0.00	0.064 0.04	−0.013 0.69	−0.012 0.71	N		
0.177 0.00	0.162 0.00	−0.152 0.00	0.146 0.00	−0.137 0.00	−0.064 0.05	Symbol Test	
0.191 0.00	0.154 0.00	−0.130 0.00	0.117 0.00	−0.183 0.00	−0.082 0.02	0.404 0.00	Word Test

Notes: O=Openness to Experience, C=Conscientiousness, E=Extraversion, A=Agreeableness, N=Neuroticism. The top value in each box is the pairwise correlation and the bottom value is p-value. Darker shaded boxes have lower *p*-values. The measures of the Big Five are based on 3 questions each. The measures of cognitive ability (symbol test and word test) are based on timed modules similar to the Wechsler Adult Intelligence Scale (WAIS). Risk tolerance was elicited through a real-stakes experiment.
Source: Authors own calculations using experimental data similar to GSOEP.

6.C. Mapping preferences into personality

Despite the empirical and intuitive conceptual links between preferences and traits, a precise mapping between the measures is not available. It is often unclear whether two variables, such as a measure of time preference and a measure of Conscientiousness, measure the same thing, whether one generates the other, or whether both are generated from underlying, more complex preferences. In section 3, we argued that measured personality is generated by underlying

endowments and preference parameters. However, the preferences measured by economists are often chosen such as to ensure identification on particular types of data on choices and may be misspecified. Further, studies documenting relationships between preferences and traits typically only study correlations without being motivated by an underlying model. Hence any talk about causality based on these studies is, at this stage, pure speculation. There are two main reasons for the disconnect between measures in economics and measures in psychology.

First, economists are interested in marginal rates of substitution, and they measure them over relevant ranges via choices. Personality psychologists are generally not interested in these tradeoffs and often do not measure choices. Most approaches to measuring preferences in economics, whether observational or experimental, use some variation of revealed preference given observed choices. In contrast, psychologists measuring traits ask questions about preferences, information or "typical" actions. Some measures involve questions about how subjects would feel about a given outcome, without giving an alternative. While such questions may elicit some (unspecified) feature of preferences when compared to other evaluations, it is far from clear exactly what is being measured. This difference in focus makes it harder to compare preference parameters and personality measures.

Second, traditional preference parameters in economics may not adequately describe underlying preferences and may not span the entire space of human decisions measured by psychologists. Time, risk, social, and leisure preferences may not capture the only tradeoffs in life. While time preference, risk aversion, leisure preference, and social preference have analogues in psychology, many personality psychologists do not perceive self-control and delay of gratification, risk-taking behavior and sensation-seeking, and motivation and ambition as the most important aspects of human decision making.

Economists make strong simplifying assumptions to make models tractable and to ensure identification. The parameters so generated are used to build models, evaluate policy, and create counterfactual worlds. The most widely used ways of specifying tradeoffs are through parameterizations assuming separability, and assuming that the marginal rates of substitution are each represented by a *single* parameter. Personality psychologists do not have the same incentives as economists to describe behavior by simple specifications as they are often content to stop with rich descriptions and do not use their estimated relationships in subsequent policy analysis. Thus, they allow for a more complex range of behaviors. The choice of measured traits is large and often defies a simple, tractable, explanation. As discussed in the previous sections, to economists these often appear to be arbitrary.

Hence, while economics may be able to contribute to the literature on personality psychology with more coherent methods of measurement, we focus in this chapter on how personality psychology may contribute to economics with richer specifications of preferences and capability constraints.

6.D. Measuring preferences

When measuring preferences, economists are not interested in levels, but in characterizing the tradeoffs between goods over a relevant range, i.e., in measuring $MRS_{m,l}$. Most approaches to measuring preferences, whether observational or experimental, apply some variation of revealed preference theory to observed choices, prices, and incomes. Ideally, we would like to know the marginal rates of substitution between all goods over the entire choice set. Given a price ratio between two goods, we know the marginal rate of substitution at one point. Using observational data, economists typically assume some distribution of agents with identical preferences, facing different prices, and experimental economists ask a series of

questions asking subjects to choose between several options, varying the price ratio between questions. Hence if we assume nothing about preferences, we would have to measure choice at all combinations of incomes and price ratios between all goods. However, if we assume that preferences are such that if the MRS is governed by only one parameter, in principle, without any measurement error, an observation at a single price ratio is enough to identify preferences. Without any restrictions on preferences, identifying parameters over the entire space would be a challenging task, so restrictions on preferences are typically chosen such that the main aspects of choice are captured while ensuring that parameters are identified.

Two types of restrictions are typically made on preference specifications. First, functional separability is often assumed, since this ensures that we can ignore all other goods when measuring marginal rates of substitution between any pairs of goods. Second, the functional form is chosen such that the marginal rate of substitution is described by as few parameters as possible while still capturing main psychological features of choice. Assuming separability is often very restrictive, as we will show in the evidence given below. People are affected by various incentives and influences when choosing between goods. Ignoring these features can lead to large variation in estimated parameters. Likewise, additive separability implies dependencies between aspects of preference such as time and risk, which may not be closely connected from a psychological standpoint (Gorman [1968]). Simple parameterizations hide the fact that some aspects of preferences are multi- rather than uni-dimensional. We will review this evidence below.

Standard preference specifications over time, risk, and income distributions often assume additive separability and simple parameterizations. The parameterization is typically chosen to

capture two features of preferences, the *relative importance* of goods and *variability* across goods. To see this more clearly, consider the preferences given by

$$U\left(X_1, X_2, \ldots X_K\right) = w_1 u_1\left(X_1\right) + w_2 u_2\left(X_2\right) + \ldots + w_l u_l\left(X_l\right) + \ldots + w_K u_K\left(X_K\right).$$

The marginal rate of substitution for this case is

$$MRS_{m,l} = \frac{w_m}{w_l}\frac{u_m'\left(X_m\right)}{u_l'\left(X_l\right)}.$$

First, note that the separability ensures that we can disregard the consumption of all other goods when measuring this marginal rate of substitution. Further, the relative importance of goods is captured by the relative weights $\frac{w_m}{w_l}$, and the preference for variation across goods is captured by the curvature in the subutility function (i.e. the u_l). The standard preference specifications for time, risk, and social preferences are special cases of this function. First, expected utility is given by this preference specification $w_l = p_l$ for all l, and $u_l = u_{l'}$ for all l, l'. The weights are the probabilities on each of the states of the world. The curvature represents risk aversion, disutility from variation across states.

Time preferences in the form of standard exponential discounting is given by $n = 1$ and $w_j = \beta^j$. $u_l = u_{l'}$ for all $l \neq l'$. Here, β is the discount factor, and the agent assigns zero weight to all periods in the past. The curvature is his intertemporal elasticity of substitution, how much variability he is willing to accept across time. With CES utility, risk preference and the intertemporal elasticity of substitution are the same parameter. Further, one can think of these preferences as representing social preferences. Take the case $u_l = u_{l'}$ for all l, l'. Each subutility function represents the agent's utility of another person's consumption. The w_l are the weights the agent assigns to his own consumption, as well as to the people around him. If the

curvature of u is very high, the agent will not allow for much inequality. Also note that the discount factor β is usually assumed to be subjective, while the weights in the expected utility are usually assumed to be objective probabilities, if these are available.

It is implicitly assumed that preference parameters are constant over any domain, any time period, any social tradeoff or tradeoff between paid work and consumption. Another way of stating this is that we are assuming separability between, e.g., risk tradeoffs and the type of lottery, or between two time periods and the period in which the tradeoff is evaluated. Functional forms are constructed in this way to capture, in the simplest way possible, some desirable features, rather than to give the most complete description of reality. The additively separable model, although convenient, is thus a very restrictive preference specification. We have no inherent reason to believe that preferences are separable, linear, or that the weights are given by probabilities of states, or constant discount factors.

6.E. Do measured parameters predict real world behavior?

One test of whether economists are measuring stable parameters is whether the conventional economic parameter predict behavior in other contexts. Dohmen, Falk, Huffman et al. [2009a] experimentally measure risk preference in a subsample of the German Socio-Economic Panel (GSOEP) and find that it predicts self-reported risky behaviors, such as holding stocks, being self-employed, participating in sports, and smoking, but it does not predict as well as a survey question about 'willingness to take risks in general'. However, this may simply be because both the self-reported behaviors and questions about willingness to take risk are noisy contemporaneous survey measures. Barsky, Juster, Kimball et al. [1997] measure risk tolerance, time preference, and the intertemporal elasticity of substitution and find that risk tolerance predicts smoking and drinking, holding insurance and stock, and decisions to immigrate and be

self-employed. However, a large part of the variation in risky behaviors is unexplained by measured risk tolerance.

Mischel's famous experiment from the 60's, measuring children's delay discounting and relating it to later life outcomes found that children with lower discount rates had better social, cognitive, and mental health outcomes in later life. (For an overview of these studies see Mischel, Ayduk, Berman et al. [2010].)

Benz and Meier [2008] compare measures of social preferences with charitable giving in a field experiment and find that experimental measures do not predict real life behavior well. Levitt and List [2007] and List [2009] discuss the more general discrepancy between results from the lab and the field and argue that this is not necessarily because people behave inconsistently, but because experimenters are not controlling for relevant aspects of the choice situation. This is just a version of the person-situation debate. Falk and Heckman [2009] present a different interpretation of the value of experiments. We discuss the evidence below.

6.F. Towards a better understanding of preferences

There is a large literature in economics on how standard models of preferences can be improved to better reflect reality, and the literature on personality psychology and preferences contributes to this line of research. In particular, the field of behavioral economics has highlighted many so-called anomalies, ways in which standard preferences do not give a very good description of human behavior. We can divide these attempts into two main approaches.

First, behavioral economists have tried to improve models of behavior by developing more flexible functional forms for preferences. Some of the now standard examples are given below, such as loss aversion, hyperbolic discounting, and reciprocity. These are not anomalies with respect to rationality, but in those schemes agents rank choices in ways that conflict with

standard models of preferences. For example, the time inconsistent actions induced by hyperbolic discounting defined below, are often described as "errors", but they are not. The agent is simply optimizing non-standard preferences.

Secondly, behavioral economists have introduced the concept of bounded rationality—what we refer to as cognitive constraints. These are behaviors for which there is no reasonable preference specification such that the behavior in question is rational. They are called anomalies or biases relative to conventional economic choice frameworks. Examples include failure to predict the winner's curse, mental accounting, framing effects, failure to apply Bayesian updating, and default effects. We will think of these as mental constraints, or traits, along the lines of the model in Section 3. They are consistent with evidence reviewed below on the interaction between cognitive ability and preference parameters.

Note that while some of the nonstandard features of preferences may seem compelling, the higher level of generality tend to make parameters unidentified on the data commonly used by economists. See the discussion in Hansen [2005]. In the following, we will review the evidence on misspecification and capability constraints in detail.

6.G. Capability constraints

If we attempt to measure preferences without considering all constraints agents face, we will generally get biased estimates. In the model of personality in section 3, we describe how agents not only react to their preference parameters, but also to their traits. The literature in behavioral economics and the literature on personality has focused, in particular, on bounded rationality, or what we might call cognitive constraints. The marginal rate of substitution is identified through price variation. However, suppose the true shadow price ratio depends not

only on observed prices, but also on unobserved traits T of the individual. Failure to control for T leads to bias.

The empirical literature has focused on the interaction between cognition and preference parameters. Virtually all methods of estimating time preference assume that respondents are equally numerate, but Peters, Vastfjall, Slovic et al. [2006] show that this assumption is often untrue. Furthermore, more numerate individuals are less susceptible to framing effects and draw stronger and more precise affective meaning from numbers and comparisons using numbers. The confound with numeracy may explain why more intelligent (or educated) individuals often display lower discount rates when decisions require complex calculations to compare subtly different delays or reward amounts (for example, de Wit, Flory, Acheson et al. [2007]; Dohmen, Falk, Huffman et al. [2007]) but does not shed light on why smarter individuals also have lower discount rates when choosing between relatively simple cash sums (Funder and Block [1989]) and between non-cash rewards (such as smaller vs. larger candy bars in Mischel and Metzner [1962]).[146] A meta-analysis by Shamosh and Gray [2007] of 24 studies in which both IQ and discount rates were measured shows the two traits are inversely related $(r = -0.23)$. The complexity entailed by comparing the present and future values of rewards suggests that the inverse relationship between discount rates and intelligence is not just an artifact of measurement. One explanation for this could be that cognitive ability is related to the ability to direct attention. Daly, Delaney and Harmon [2009] find that lower discount rates are associated with cognitive mindfulness, which includes the ability to control attention. Further, an individual with poor working memory and low intelligence may not be capable of accurately calculating or even perceiving the value of a deferred reward. At the least, making such calculations is more

[146] Heckman [1976] shows that more educated people have lower discount rates. More able people are more likely to attend more years of school.

effortful (that is, costly) for individuals of low cognitive ability. If the cost of making calculations exceeds the expected benefit of such deliberation, the individual may choose by default the immediate, certain reward. However, it is important to be aware of reverse causality, since more patient individuals may also invest more in cognitive ability.

Cognitive constraints also matter for risk preference. There appears to be an inverse relationship between cognitive ability and risk aversion, where higher-IQ people have higher risk tolerance (Benjamin, Brown and Shapiro [2006]; Dohmen, Falk, Huffman et al. [2007]).[147] Reference dependence can lead subjects to be susceptible to framing because they will perceive two identical bets differentially if one is framed as a loss and the other is framed as a gain. There is some evidence that individuals with higher cognitive ability and education are less risk averse. Burks, Carpenter, Goette et al. [2009] find that higher IQ individuals are more consistent in their choices between a lottery and fixed sums. They hypothesize that agents with higher cognitive ability are better at translating their preferences into choices between lotteries.

Since ambiguity aversion can lead to irrational choices, such as in Ellsberg's paradox, we may think of this as an interaction between risk and cognitive constraints, However, Borghans, Golsteyn, Heckman et al. [2010] find that while risk aversion is related to personality traits, ambiguity is not. In particular, IQ does not explain how subjects choose between a risky and an ambiguous urn.

Further, personality can serve as a constraint in the sense that lacking ability to experience empathy (e.g. psychopathy) will constrain the expression of social preferences. Whether this should be viewed as a constraint or as a part of preferences is not clear.

[147] The two cognitive ability tests used by Dohmen, Falk, Huffman et al. [2007] were a coding speed and vocabulary test.

6.H. Misspecification of functional forms

We may organize the various ways of modifying standard functional forms into two broad and sometimes overlapping categories. First, some of the domains which we traditionally perceive as fundamentally different, such as between risk and time preference, social and risk preference, and between leisure and time preference, may be closely related. Conversely, some of the parameters typically measured may be functions of multiple parameters. Second, there are nonseparabilities between the tradeoffs in question and other variables affecting utility, for example, age may affect how consumption is discounted over time.

6.H.1. Multidimensionality

Marginal rates of substitution are often assumed to be generated by only one or two parameters, for example the discount factor and the intertemporal elasticity of substitution. This of course ensures that parameters are identified given sparse data, and, if it is a sensible specification of preferences, it gives a very convenient description of behaviour. The problem is that one or two parameters may not describe behaviour well. Conversely, some of the concepts analyzed separately in the literature may be governed by the same parameters.

Frederick, Loewenstein and O'Donoghue [2002], in discussing the concept of time discounting, argue that "the cumulative evidence raises serious doubts about whether there is, in fact, such a construct—a stable factor that operates identically on, and applies equally to, all sources of utility" (p. 392). Instead, Frederick et al. suggest that time preference is tri-dimensional, comprising three separate underlying motives: *impulsivity,* the tendency to act spontaneously and without planning; *compulsivity,* the tendency to stick with plans; and *inhibition,* the ability to override automatic responses to urges or emotions. However, they do not provide a precise map between ρ and these personality traits, nor do they show how exactly

their three traits affect choice behavior. There may be multiple interpretations of this assertion.

First, it may be that the tradeoff between different time periods is simply described by several parameters. Second, it may be that impulsivity, inhibition, and compulsivity are traits which we should think of as constraints, i.e., as something which affects shadow prices of consumption in different time periods. Third, it may be that the relevant tradeoff is not between different time periods but, for example in the case of impulsivity, between various levels of sensation, a behavior which is also related to risk seeking.

Like time preference, risk preference may be multidimensional. As noted by Rabin [2000], the simple expected utility framework does not explain risk aversion over small stakes, since it would imply an implausibly high curvature of the utility of wealth function. See Starmer [2000] for a review of the literature on departures from expected utility. Weber [2001] shows that risk preference varies by domain, and a scale that assesses risk taking in five different domains shows low correlations across these domains (Weber, Blais and Betz [2002]). One can be quite risk-averse when it comes to financial decisions but risk-loving when it comes to health decisions (Hanoch, Johnson and Wilke [2006]). Weber's risk-return model of risk taking (Weber and Milliman [1997]; Weber and Hsee [1998]) finds that low correlations among risk-taking preference across domains can be explained by domain-specific perceptions of riskiness and return. Dohmen, Falk, Huffman et al. [2009a] find that a survey question on willingness to take risks within a domain predicts self-reported behaviors within each domain. One reason for this domain specificity may be that sensation-seeking, enjoyment of risk per se, is an important aspect of risk preferences.[148]

[148] Zuckerman [2007] suggests that sensation seeking is related more closely to Big Five Conscientiousness (inversely), but there is obvious conceptual overlap with excitement seeking, a facet of Big Five Extraversion on the NEO-PI-R questionnaire, as well as with Big Five Openness to Experience.

Another aspect of risk preferences which may help explain the seeming inconsistencies is ambiguity aversion, the disutility from model uncertainty. This is measured as the tradeoff between lower expected return and higher model uncertainty. Ambiguity aversion can lead to Ellsberg's paradox: people tend to prefer an urn with a .5 probability of winning to an urn with an unknown probability where they are allowed to choose which side to bet on. One version of these preferences is Gilboa and Schmeidler [1989] max-min preferences, where the agent maximizes an expected utility function which has been minimized with respect to the prior probabilities, i.e.

$$U\left(X_1, X_2, ..., X_K\right) = \min_{(p_1, p_2 ... p_K)} p_1 u\left(X_1\right) + p_2 u\left(X_2\right) + ... + p_K u\left(X_K\right).$$

Borghans, Golsteyn, Heckman et al. [2009] measure ambiguity aversion and risk aversion in a group of Dutch high school students and show that this aspect of choice is distinct from risk aversion.

As for preferences for leisure, agents may not simply distinguish between "leisure" and "work" but also between the different kinds of work. Thus, the concepts of intrinsic and extrinsic motivation have been studied by behavioral economists (Benabou and Tirole [2003]). Introducing extrinsic incentives may erode the intrinsic motivation for work.[149] These distinctions may be important since having strong preference for more interesting work will increase the incentive to invest in human capital. This implies that constraints on educational choice may lead to higher measures of preference for leisure, since an individual who is not able to reach his full potential in terms of human capital investment may be constrained to work

[149] Psychologists have long discussed the contrast between extrinsic and intrinsic rewards. See Deci and Ryan [1985]. See also Kohn [1999].

which he does not enjoy. This means that some people will appear to have a low preference for leisure, while it is only the case when the work they are doing is interesting.

One aspect of preferences where there is no consensus on the parameters governing choices is the notion of social preferences. Social preferences is a broad term used to refer to any explanation for non-selfish behavior, usually as measured in a dictator game where people have to divide a sum between themselves and another person. Typically, more than 60 percent of proposers give positive amounts, averaging 20 percent of the sum. A variation of this game is the classic ultimatum game where a giver has to divide a sum between himself and another subject, the receiver, and where the other subject can accept or decline the sum. If he declines, both will lose their money. Studies typically find that receivers decline if offered less than 20 percent. These results cannot be explained by pure selfishness. In the dictator game, the giver is willing to forgo his own consumption in order to increase another person's consumption, and in the ultimatum game, the receiver is willing to forgo his own consumption in order to decrease the giver's consumption if he pays him too little. There is a large literature on finding the deeper parameters that govern these behaviors. The literature on these preferences includes several tradeoffs, notably between the utility of oneself and that of others, efficiency, and fairness. The notion of fairness covers various concepts, including equality and rewards in proportion to talent, effort, kindness, or intentions. For reviews of this literature, see List [2009], and Camerer and Fehr [2004].

In the linear, separable model, where each good X_i is the consumption of person i, we can think of the weights as caring or altruism, the fact that people often care about other people's utility or consumption. See Meier [2007] for a review. Fehr and Schmidt [1999] analyze

inequality aversion in which people dislike inequality rather than valuing the consumption or utility of agents per se.

Caring or altruism has been shown to decrease with social distance. People typically care more about themselves than about others, and they are less altruistic the less well they know other people. The social preference of reciprocity has been studied. Fehr and Gächter [2000], and Falk and Fischbacher [2006] present evidence on reciprocity and conditional cooperation, in which agents act in a pro-social or antisocial manner depending on the behavior of others with whom they interact. People exert positive reciprocity if they tend to reward others for kindness but negative reciprocity if they tend to punish others for unkindness. More precisely, they are willing to incur a cost on themselves in order to reward or punish others. Falk and Fischbacher [2006] develop a theory of reciprocity where utility depends on the kindness of others, which is a function not only on the outcome from the another person's action, but also on the perceived intentions. Reciprocity then reflects how much value a person puts on rewarding kindness. We could reflect these features by letting the person-specific weights on the subutilities depend on social distance and past actions of others. Reciprocity is often measured using a gift-giving game where the proposer offers a wage to a responder, who then subsequently chooses a level of effort. However, List [2009] argues that the effect of fairness preferences may have been overstated in the literature, that many of the observed results are due to concerns over either reputation or scrutiny by experimenters. Several studies have shown that the reciprocity effect may die out when the study is conducted over a longer time frame rather than the short duration of lab experiments (Gneezy and List [2006], Hennig-Schmidt, Rockenbach and A [2010], Kube, Maréchal and Puppe [2006]). Andreoni's [1995] warm glow model of altruism suggests that people do not simply care about others, but about the act of giving.

Inequality aversion is distinct from caring in the sense that A's utility may be decreasing in B's consumption if it is higher than A's. (See Fehr and Schmidt [2006] for a review.) Fehr and Schmidt [1999] suggest the following asymmetric specification for the utility of agent η :

$$U_n(X) = X_n - \alpha_n \frac{1}{K-1}\sum_{j\neq n} \max|X_j - X_n, 0| - \beta_n \frac{1}{K-1}\sum_{j\neq n} \max|X_n - X_j, 0|$$

where the weights satisfy $\beta_n \leq \alpha_n$ and $0 \leq \beta_n \leq 1$. The additively separable model does not encompass this. In this case, MRS changes from 1 to -1 at $X_A = X_B$. Higher weight is given to own consumption than to that of others, but when asked about the distribution of rewards in games where subjects do not have any interest in the outcome, there can be considerable variation in preferences.

People seem to be more accepting of inequality if they believe that it represents a difference earned through effort and talent, but effort seems to be more important than talent (See Tausch, Potters and Riedl [2010] for a review). This may be related to the notion of reciprocity. The distinction may be whether the preference is for people who have earned their reward for doing something "for me" or something admirable in general.

A different aspect of social preferences include "trust". On can think of this as the belief about how own actions affect those of others; Dohmen, Falk, Huffman et al. [2008] use the German Socio-Economic Panel (GSOEP) and find that most people exert positive reciprocity; positive reciprocity and negative reciprocity are only weakly correlated; and people who are negatively reciprocal are less willing to trust others. Altmann, Dohmen and Wibral [2008] conduct an experiment that implements a trust game and find that trust and positive reciprocity are positively related.

In standard economic models, it is assumed that preferences are separable across domains, and that time, risk, leisure, and social preferences are fundamentally distinct concepts which do not interact with each other. However, this need not be true. Consider the example of additively separable preferences over risk, time, and distribution with a decreasing marginal utility of wealth function. Suppose in each time period, own income and income of another agent is drawn from a known independent and identical distribution for all agents. Preferences are

$$\sum_{t=0}^{\infty} \beta^t E\left[w_1 u\left(c_1\right) + w_2\left(c_2\right)\right],$$

where w_1, w_2 represent how much he cares about himself and the other agent, respectively. In this case, the intertemporal elasticity of substitution (eis) is the same as the agent risk preference.[150] However, note that we can use the same argument to show that it is also equal to the agent's inequality aversion as measured by his elasticity of substitution between himself and the other agent. (See Atkinson [1970]). This results relies on the additive separability of preferences, and it is possible to construct preferences where these are distinct concepts. However, Hansen [2005] shows that in many cases, even for more flexible preference specifications, time and risk preferences are observationally equivalent.

Despite the common discounting process demonstrated above, however, Barsky, Juster, Kimball et al. [1997] find no evidence that the intertemporal elasticity of substitution is correlated with risk tolerance. However, the sample on which they measure these parameters is small. Green and Myerson [2004] argue that risk and time belong to different underlying psychological processes. As evidence, they point out that the two constructs react differently to

[150] This result is due to Gorman [1968]. See Browning, Hansen and Heckman [1999] for a definition of eis and a survey of evidence (through 1998) on the magnitude of eis.

the same effect: for example, an increase in the size of reward generally decreases the discount on time but increases the discount rate when rewards are probabilistic.[151] This is evidence against the standard intertemporally separable model of risk aversion.

Risk aversion may affect the way multiple other goods are traded off. Altmann, Dohmen and Wibral [2008] find that people who are less risk averse are also more willing to give money to an investor in a trust game where he will only be repaid if the investor decides to return the favor. In this type of game, it seems natural that trust is closely related to risk and ambiguity aversion, that a person who is more prone to accept uncertainty is also more likely to trust others. Great care has to be taken in trying to separate trust from risk aversion given the observational equivalence results noted above. Kosfeld, Heinrichs, Zak et al. [2005] find that people who receive oxytocin exhibit more trusting behavior in a real-stakes trust game. Oxytocin, however, does not make subjects more generous, suggesting that trust is not simply altruism. Additionally, oxytocin does not affect people's decision over risky outcomes when playing against a computer rather than a human. Combined, these findings suggest that there is a unique characteristic that affects willingness to trust, distinct from altruism and risk aversion. Fehr [2009] posits that this missing element might be "betrayal aversion." Using survey data from Germany, Fehr [2009] finds that risk preferences, betrayal aversion, and altruism (as expressed through volunteering) predict people's self-reported willingness to trust others.

Likewise, Borghans, Golsteyn, Heckman et al. [2009] show that risk-aversion is positively associated with Neuroticism, which contains measures of fear and strong emotional responses to bad outcomes. They also find that risk aversion is negatively associated with

[151] Further support for this disassociation comes from a cross-cultural study by Du, Green and Myerson [2002], in which Chinese graduate students discounted delayed rewards much more steeply than Japanese students, but Japanese students discounted probabilistic rewards more steeply than did the Chinese. Barsky, Juster, Kimball et al. [1997] report that their estimates of time preference and risk tolerance are independent.

ambition, a trait which may involve investment in uncertain opportunities. Further, Agreeableness is positively associated with risk aversion.

6.H.2. Nonseparabilities

The most restrictive version of the additively separable model suggests that the marginal rate of substitution between two goods is unaffected by the consumption of other goods. Browning, Hansen and Heckman [1999] show that there is much evidence against this assumption. If nonseparability is not accounted for, this will induce apparent situational inconsistencies. Further, estimates will suffer from omitted variable bias.

One type of nonseparability is between goods and the state or time period in which the marginal rate of substitution is measured. The additively separable model does allow for this type of dependence, represented by the subscript i on the utility function. While exponential discounting is still the most common representation of time preferences, experiments show that people tend to put higher weight on the present than on future periods. This is the motivation for hyperbolic discounting. The most often used specification is (β,δ)-preferences where β is the usual discount factor while δ is an additional discounting of all future periods.

$$U_\tau\left(X_\tau, X_{\tau+1},...\right) = u\left(X_\tau\right) + \delta\beta u\left(X_{\tau+1}\right) + \delta\beta^2 u\left(X_{\tau+2}\right) + ...$$

the consequence of these preferences is that the tradeoff between period τ and period $\tau+1$ is not evaluated the same way from the perspective of period $\tau-1$ and period τ, leading to time inconsistency. Other possibilities are that the discount rates change with age. Hyperbolic and age-dependent discounting makes use of the subscript τ on the utility function. We may think of an agent in multiple periods as several agents who play a game with each other. The agent today has to take into account what future agents might do. Further, discount rates appear to vary inversely with the size of reward (Green, Fry and Myerson [1994b]; Kirby [1997]), and vary

with the type of reward offered (Chapman, Nelson and Hier [1999]; Chapman and Coups [1999]; Estle, Green, Myerson et al. [2007]; Bickel, Odum and Madden [1999]; Bonato and Boland [1983]).

As previously noted, the expected utility form for risk preferences does not explain risk preferences over small stakes, as argued by Rabin [2000]. If subutility functions represent utility of lifetime wealth in different states, people should be approximately risk neutral for small stakes. However, people often avoid more than fair small bets. If this is explained by expected utility, then the curvature of the utility of wealth function would have to be implausibly high. Kahneman and Tversky [1979] suggested that people are loss averse, i.e., that losses weigh higher than gains in the utility function. This would imply that people have state-dependent preferences, which can be expressed as

$$U_n\left(X_1, X_2, \ldots X_n, \ldots X_K\right) = p_1 u\left(X_1 - X_n\right) + p_2 u\left(X_2 - X_n\right) + \ldots p_K u\left(X_K - X_n\right)$$

where $u'(y)$ is higher for negative y than for positive y. Note that this specification is very similar to that of inequality aversion discussed above, and it is not captured by the additively separable model. Both models share the feature that people do not have stable preferences over levels, but over differences.

The concepts of loss aversion, reference point dependence, and endowment effects (Thaler [1980], Kahneman and Tversky [1979]) are variations on this theme. If an agent has had an object in his possession for even a short amount of time, it affects how he trades it off against other goods. List [2003] has shown that this effect disappears when agents have market experience. However, loss aversion will interact with many choices in life, for example, how agents evaluate lotteries. Another example of reference dependence has been demonstrated by List [2007] and Bardsley [2008]. They let agents play a dictator game with a $10 endowment,

but they now tell subjects that they are allowed to take up to $10 from the other player. This modification nearly eliminated giving.

Experimental measures of social preferences vary greatly across studies. Levitt and List [2007] and List [2009] make the well known point that the degree of scrutiny in the lab as opposed to in the real world may make subjects behave more pro-socially (Bandiera, Barankay and Rasul [2005], List [2006]) and argue against the "realism" of experimental data.[152] Further, several studies have found that people tend to be more selfish when the stakes of the game increase (Carpenter, Verhoogen and Burks [2005], Slonim and Roth [1998], Parco, Rapoport and Stein [2002]).

There is evidence of substantial heterogeneity in preferences both between socioeconomic groups and within groups. Marginal rates of substitution depend on other things people value such as education, youth, cultural values, etc. See the evidence in Browning, Hansen and Heckman [1999]. This evidence supports the claim that people are different at a basic level, since preferences govern the choices that shape life. However, preferences may be experience dependent. While most studies view life outcomes as the result of choices governed by exogenous preferences, and hence infer preferences from outcomes, it may be the case that initial conditions determine both preferences and constraints on the available choices.

The motivation for preference specifications in economics is typically introspection, axioms about rationality, and convenience, rather than empirical evidence. When measuring preferences, functional forms are chosen in an attempt to minimize approximation error subject to identification. However, economists also seem to be constrained regarding which fundamental attributes agents have preferences over. Time, risk, and social preferences may not

[152] See, however, Falk and Heckman [2009].

be the right dimensions of choice over which parameters are stable. Each of these domains seem to be guided by multiple parameters, and some of these parameters seem to matter for each of the domains. While the marginal rates of substitution economists measure are correct at observed prices, they may not be easily mapped into the conventional preference specifications. Personality psychology may help in guiding economists as where to look for more fundamental parameters. However, at the time of this writing, the potential is largely unexplained.

Table 8 summarizes the empirical evidence discussed in this section. It presents a summary of the main papers relating economic preference parameters to psychological measurements. We list the authors, preferences studied, the way causality is determined, and the main results.

Table 8: Link between Personality Traits and Preferences

Author(s)	**Main Variable(s)**	**Data and Methods**	**Causal Evidence**	**Main Result(s)**
Altmann, Dohmen and Wibral [2008]	Outcome(s): *trust* – amount the first-player sends in a real-stakes experimental trust game Explanatory Variable(s): *reciprocity* – amount returned by the second player in a real-stakes experimental trust game; *risk aversion* – certainty equivalent as measured by real-stakes choices over lotteries	Data: Collected by authors; 240 students from the University of Bonn Methods: OLS	Controls: gender Timing of Measurements: The measures are contemporaneous. Theory: People might generally value adhering to social norms associated with trust and reciprocity.	Reciprocity and trust are positively related ($p<0.01$). Risk aversion and trust are positively related ($p<0.05$).
Borghans, Golsteyn, Heckman, & Meijers [2009]	Outcome(s): *risk aversion* – choices over real-stakes lotteries; *ambiguity aversion* – comparison of the willingness to bet on lotteries when the probability distribution is unknown Explanatory Variable(s): *gender*; *personality* – self-reported measures of The Big Five, ambition, flexible thinking, and self-control	Data: Collected by authors; 347 students aged 15 to 16 from a Dutch high school Methods: OLS, F-test	Controls: n/a Timing of Measurements: The measures are contemporaneous. Theory: Risk aversion and ambiguity aversion represent different preferences and might reflect different personality traits.	Men are less risk averse than women ($p<0.001$) but more ambiguity averse ($p<0.05$). Risk-aversion is mediated by personality ($p<0.05$), while ambiguity aversion is not. Risk-aversion is positively associated with Agreeableness and Neuroticism and is negatively associated with ambition ($p<0.05$).

Borghans, Meijers and ter Weel [2008]	Outcome(s): *cognitive ability* – number of correct answers on an IQ test; *effort* – time spent on each question Explanatory Variable(s): *risk aversion* – survey response to lotteries; *time preference* – survey response to trade-offs across time; *leisure preference* – survey response; *experiment incentives* – payment for correct answers to the IQ test; *personality* – self-reported Big Five, performance motivation, positive and negative fear of failure, locus of control, social desirability, curiosity, resilience, enjoyment of success, attitude toward work	Data: Collected by authors; 128 university students from a Dutch University Methods: probit	Controls: type of cognitive test, the amount of incentive pay, and time constraints Timing of Measurements: They measured IQ both before and after providing incentives. Theory: People with different personalities and preferences might be willing to expend different amounts of mental effort during a test.	Performance motivation, fear of failure, internal locus of control, curiosity, low discount rates, and risk aversion are positively associated with more correct answers ($p<0.05$). Negative fear of failure, Extroversion, Openness to Experience, and Agreeableness are negatively associated with answering the question correctly ($p<0.05$). Incentives did not affect the number of questions answered correctly. Intrinsic motivation, curiosity, internal locus of control, Emotional Stability, Conscientiousness, and discount rates are negatively associated with responsiveness to incentives ($p<0.05$). Risk aversion is negatively associated with responsiveness to incentives ($p<0.10$). Leisure preference and Openness to Experience are positively associated with responsiveness ($p<0.05$).
Burks, Carpenter, Goette et al. [2009]	Outcome(s): *risk aversion* – choices over real-stakes lotteries; *time discounting* – choices over real-stakes payments at different times; *inconsistent risk and time preference* – making at least one inconsistent choice in the experiments eliciting preferences; *job performance* – whether a worker leaves before the end of the first year Explanatory Variable(s): *cognitive ability* – IQ as measured by an adaptation of Raven's Standard Progressive Matrices	Data: Collected by authors, administrative data; 892 trainee truckers from a U.S. trucking company (2005-2006) Methods: OLS, interval regressions, linear probability model, Cox proportional hazard	Controls: race, age, age squared, education, household income, absorption, achievement, aggression, alienation, control harm avoidance, social closeness, social potency, stress reaction, traditionalism, and well-being Timing of Measurements: The measures are contemporaneous, except for job-turnover which was evaluated after the experiment. Theory: People with higher IQ can better forecast the future.	An increase in IQ from the bottom quartile to the top quartile is associated with an increase in risk-taking consistency of 25 percentage points ($p<0.001$), an increase of intertemporal consistency of 15 percentage points ($p<0.001$), a decrease in discount rate ($p<0.001$), and a decrease in risk aversion ($p<0.001$). People in the lowest quartile of IQ are about twice as likely to leave the job within the first year ($p<0.001$).

Daly, Delaney and Harmon [2009]	Outcome(s): *time preference* – discount rate measured by a real-stakes choices over delayed payments Explanatory Variable(s): *health* – blood pressure, body fat, blood glucose, weight, height, heart rate; *personality* – questionnaire measures of The Big Five, self-control, consideration of future consequences, elaboration of potential outcomes, emotional regulation, cognitive and affective mindfulness, suppression of unwanted thoughts, experiential avoidance	Data: Collected by authors; 204 students from Trinity College Dublin Methods: factor analysis, OLS	Controls: age and sex Timing of Measurements: The measures are contemporaneous. Theory: Personality traits and health indicators might be associated with willingness to delay gratification.	Age and sex do not predict the estimated discount rate. A factor that loads heavily on self-control, consideration of future consequences, elaboration of consequences, affective mindfulness, and Conscientiousness is negatively associated with the discount rate ($p<0.01$). A factor that loads on blood pressure is positively associated with the discount rate ($p<0.10$).
Dohmen, Falk, Huffman et al. [2009a]	Outcome(s): *experimental risk measure* – measured by real-stakes choices over lotteries and cash payments Explanatory Variable(s): *survey risk measure* – survey responses on an 11-point scale, relating to general risk preference and risk preference relating to car driving, financial matters, leisure and sports, career and health	Data: Collected by the authors; 450 adults from Germany Methods: OLS	Controls: gender, age, height, and other personal characteristics Timing of Measurements: The measures are contemporaneous. Theory: Survey and experimentally-elicited risk measure the same concept	Survey measures of general risk attitude predict incentive compatible, experimentally elicited measures of risk attitude ($p<0.01$).
Ding, Hartog and Sun [2010]	Outcome(s): *experimental risk measure* – measured by real-stakes choices over lotteries and cash payments Explanatory Variable(s): *survey risk measure* – responses on an 11 point scale, relating to general risk preference and risk preference relating to car driving, financial matters, leisure and sports, career and health, survey responses to hypothetical lotteries	Data: Collected by the authors; 121 students of PKU in Beijing who participated in an experiment (2008) Methods: OLS, correlations	Controls: major, gender, family income, and class rank Timing of Measurements: The measures are contemporaneous. Theory: There could be an underlying risk parameter that applies in all situations.	The survey measures of risk explain at most 10 percent of the variance in the experimental measures of risk (general risk attitude and financial risk are the best). Self-assessed risk depends much on the domain or context; the highest correlation between context-based survey questions is $r=0.55$. Women are more risk averse than men; risk-aversion decreases with parental income; and risk attitudes depend on domain (context). People view winning and losing money differently.

7. The Predictive Power of Personality Traits

This section discusses the empirical evidence on the power of personality in predicting life outcomes. It is intuitive that personality is important, but as noted in Section 2, some psychologists and behavioral economists do not agree. Nevertheless, a growing body of evidence suggests that personality traits—especially those related to Conscientiousness, and, to a lesser extent, Neuroticism—predict a wide range of outcomes. The predictive power of any particular personality trait tends to be less than the predictive power of IQ, but in some cases rivals it. Table 9 shows that adolescent measures of personality (locus of control and self-esteem) are associated with a diverse set of meaningful outcomes.[153] The fourth column shows that personality has incremental validity beyond IQ for many outcomes. As is the case in the literature as a whole, the associations vary in strength across outcomes. In this section we further explore the evidence on the link between personality and meaningful later life outcomes, discussing separately education, labor market success, health, and crime. Tables A8 – A16 in the Web Appendix complement the discussion and provide more details about the studies.[154]

[153] Tables A3 and A4 in Section A7 of the Web Appendix present correlations of the outcomes with the Big Five.
[154] See http://jenni.uchicago.edu/personality_economics/.

Table 9. Correlations, Partial Correlations, and Explained Variance of IQ and Personality with Later-life Outcomes

	Correlations and Partial Correlations			Explained Variance (R^2)		
Outcome/Variables Included	**IQ**	**Locus of Control/ Self-Esteem**	**All**	**IQ**	**Locus of Control/ Self-Esteem**	**All**
High School Diploma	0.037***		0.017	0.001	0.005	0.005
		0.049***	0.045***			
		0.044***	0.039***			
Highest Grade Completed	0.390***		0.358***	0.151	0.042	0.165
		0.013	-0.049***			
		0.199***	0.118***			
12th Grade GPA	0.486***		0.464***	0.236	0.035	0.242
		0.018	-0.065***			
		0.180***	0.075***			
Hourly Wage at 35	0.198***		0.179***	0.039	0.001	0.041
		0.063***	0.033**			
		0.058***	0.013			
Weeks Unemployed at 35	-0.137***		-0.137***	0.019	0.002	0.021
		0.023*	0.046***			
		-0.043***	-0.009			
Any Welfare at 35	-0.235***		-0.211***	0.055	0.014	0.058
		-0.038***	-0.003			
		-0.103***	-0.051***			
Depression at 40	-0.097***		-0.070***	0.009	0.011	0.016
		-0.024*	-0.012			
		-0.098***	-0.079***			
Physical Health at 40	-0.040***	-0.143***	-0.019	0.002	0.020	0.021
		0.024*	-0.138***			
			0.027**			
Mental Health at 40	0.023*		0.007	0.001	0.004	0.004
		0.056***	0.054***			
		0.021	0.019			

Note: The first column shows the later life outcome. The second column shows the correlation of the outcome with IQ. The third column shows the partial correlations of the outcome with locus of control (Rotter) and self-esteem (Rosenberg) entered jointly. The fourth column shows the partial correlation with IQ, self-esteem, and locus of control entered jointly. The fifth, sixth, and seventh columns show R^2 for the models associated with the second, third, and fourth columns, respectively. IQ is a percentile score obtained by equating IQ across different IQ tests from NLSY79 transcript data following the procedure in Borghans, Golsteyn, Heckman et al. [2010]. The sample excludes military personnel.
*statistically significant at the10 percent level; **statistically significant at the 5 percent level; ***statistically significant at the 1 percent level
Source: Authors own calculations using the NLSY79.

For three main reasons, summarizing the large literature on the predictive power of personality on outcomes is a daunting task. First, the measurements of personality and cognition differ among studies. As noted in Section 5, not all psychologists use the Big Five. We attempt

to cast all measures into the Big Five categories. When this is not possible, we discuss the measures used and how they may differ from the Big Five measures.

Second, different studies use different measures of predictive power. Many studies report only simple correlations or simple standardized regression coefficients.[155] Such estimated relationships do not control for other factors that may influence outcomes. This is particularly problematic for estimated relationships between personality and other outcomes that do not control for cognition, situation, or the effect of other personality traits. Where possible, we report both simple and partial correlations.

We also consider a measure of predictive validity that extends traditional conceptions of variance explained. Recent work by economists relaxes the normality and linearity assumptions that underlie the use of simple partial correlations and standardized regression coefficients. It measures the predictive power of variables by the slopes of percentile changes on outcomes and not by variance explained. If outcomes are characterized by substantial measurement error, a low R^2 for a predictor may still be consistent with a substantial effect of the predictor on means and quantiles. For example, Heckman, Stixrud and Urzua [2006] report the effects of percentile changes in cognitive and personality traits on a variety of outcomes over the full range of estimated relationships, relaxing traditional normality or linearity assumptions and not relying directly on measures of variance explained. This practice is increasingly being applied by economists (see, e.g., Piatek and Pinger [2010]). We report both simple and partial correlations and we note when reported empirical relationships do not control for variables likely to be

[155] Standardized regression coefficients report regressions of outcomes divided by their standard deviations. This produces correlation coefficients in bivariate regressions and partial correlation coefficients in multivariate regressions. See, e.g., Goldberger [1968].

important (e.g., studies of the effects of personality that do not control for cognition). Where possible, we do not impose linearity and normality assumptions.

Third, many studies do not address the question of causality, i.e., does the measured trait cause (rather than just predict) the outcome? Empirical associations are not a reliable basis for policy analysis. Problems with reverse causality are rife in personality psychology. Contemporaneous measures of personality and outcomes are especially problematic. For example, does greater Neuroticism lower earnings, is it the other way around, or do they mutually influence each other?

Few economists or psychologists address the issue of causality, and when they do so, it is usually by employing early measures of cognition and personality to predict later outcomes. Using early measures of psychological traits to predict later outcomes raises problems of its own. Arguably, it is the contemporaneous values of traits that affect outcomes. In Section 8, we present evidence that traits can change over the life cycle due to external factors. Using lagged values may help solve the reverse causality problem, but at the cost of inducing an errors-in-the-variables problem.[156] We carefully delineate how each study addresses causality.

A main finding of our survey, consistent with the claims of the psychologists cited in Section 2, is that Conscientiousness predicts many outcomes. Other personality traits play a role for some outcomes. In academic outcomes, Conscientiousness is important for educational attainment and achievement. Openness to Experience is independently important for broad measures of educational attainment in some samples. Traits related to Neuroticism also affect educational attainment, but the relationship is not always monotonic. For many labor market outcomes, of all of the Big Five traits, Conscientiousness is the best predictor of overall job

[156] In addition, in some versions of the economic model of personality, agents anticipate future outcomes and these are partly embodied in the lagged psychological measures.

performance and wages but is less predictive than cognitive ability. Personality is, however, more ubiquitously important across occupational categories and education levels. Additionally, for a variety of labor market outcomes, traits related to Neuroticism (e.g. locus of control and self-esteem) are also powerful predictors. For health, Conscientiousness can better predict longevity than does intelligence or background. The little evidence on crime suggests that traits related to Conscientiousness and Agreeableness are important predictors of criminality.

7.A. Educational Attainment and Achievement

We now turn to evidence for the predictive power of personality traits for educational outcomes, separately considering educational attainment, grades, and tests scores.

Educational Attainment

Despite recent increases in college attendance, American high school dropout rates remain high. About one in four American students drops out of formal schooling before receiving a high school diploma, and in recent decades the dropout rate has increased slightly (Heckman and LaFontaine [2010]). A growing body of research finds that personality is associated with educational attainment, suggesting that further study of personality and its determinants might shed light on the recent stagnation in educational attainment. We begin by reviewing evidence about the relationship of psychological measures with years of schooling and then consider specific aspects of educational achievement.

Traits such as perseverance and preferences related to an interest in learning might lead people to attain more total years of schooling. Indeed some evidence suggests that this might be the case. Table 10 presents associations between years of schooling and the Big Five from three

nationally representative samples. The studies yield different results, possibly because they control for different covariates or because they come from different countries. The first study controls for age, sex, and gender and finds that of the Big Five, Openness to Experience and Conscientiousness are most related to years of schooling attained (Goldberg, Sweeney, Merenda et al. [1998]) . The second study – which also controls for parental education and father's occupational status – reports a strong relationship with Openness to Experience but a much weaker relationship with Conscientiousness than the first study, suggesting that parental background might mediate some of the effects of Conscientiousness (van Eijck and de Graaf [2004]).

The first two samples lack information on cognitive ability. Openness to Experience, however, is the only Big Five factor with moderate associations with general intelligence (r = 0.33 in a meta-analysis by Ackerman and Heggestad [1997]), and intelligence is associated with years of education (r = 0.55 in Neisser, Boodoo, Bouchard et al. [1996]). Thus Openness to Experience may proxy for intelligence. However, as Figure 9 illustrates, controlling for rudimentary measures of crystallized intelligence and fluid intelligence, does not affect the coefficients on the Big Five within the third sample.[157] This sample differs from the others, because Openness to Experience is not strongly associated with years of education unconditional on intelligence, possibly because it is based on a smaller inventory of questions. Conscientiousness, however, is associated with years of schooling to a similar degree as intelligence. In each study, schooling and personality are measured at the same point in time, so that for older individuals, personality is measured long after schooling has been completed, complicating the interpretation if schooling affects personality.

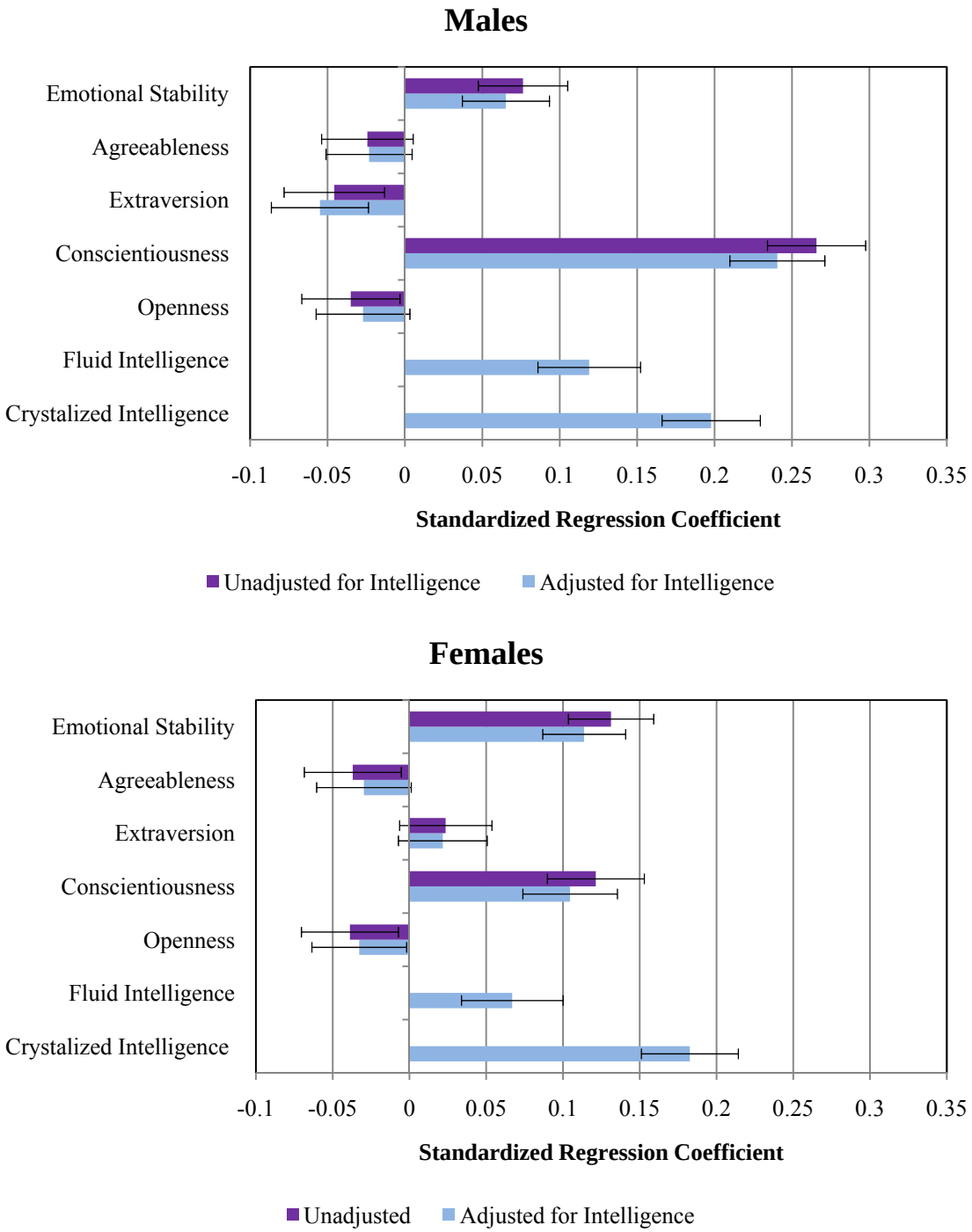
[157] Table A5 in Section A7 of the Web Appendix presents the full results from this regression. Table A6 in Section A7 of the Web Appendix presents analogous results for high school graduation.

Table 10. The Relationship between Years of Educational Attainment and Big Five Traits

Source	Sample	Timing of Measurement and Outcome	Controls	Metric	Results	
Goldberg, Sweeney, Merenda et al. [1998]	Representative sample of U.S. working adults aged 18-75 (N=3,629)	All the variables were measured in the same year, but years of schooling were cumulative.	age, gender, ethnicity	Partial Correlation with Years of Schooling (r)	Openness	0.31***
					Conscientiousness	0.12***
					Extraversion	-0.04**
					Agreeableness	-0.08***
					Neuroticism	-0.03
van Eijck and de Graaf [2004]	Representative sample of Dutch adults aged 25-70 (N=1,735)	All the variables were measured in the same year, but years of schooling were cumulative.	age, gender, father's education, mother's education, and father's occupational status	Standardized Regression Coefficient (β)	Openness	0.14***
					Conscientiousness	0.05***
					Extraversion	-0.07***
					Agreeableness	-0.07**
					Neuroticism	-0.09***
GSOEP (2004-2008), own calculations.	Representative sample of Germans aged 21-94 (N=2,381)	The Big Five were measured 3 years prior to the measurement of schooling, but years of schooling were cumulative.	age, age^2, gender, crystallized intelligence, fluid intelligence	Standardized Regression Coefficient (β)	Openness	-0.03
					Conscientiousness	0.18***
					Extraversion	-0.02
					Agreeableness	-0.03
					Neuroticism	-0.09***

**statistically significant at the 5 percent level; ***statistically significant at the 1 percent level

Figure 9. Association of the Big Five and intelligence with years of schooling in GSOEP



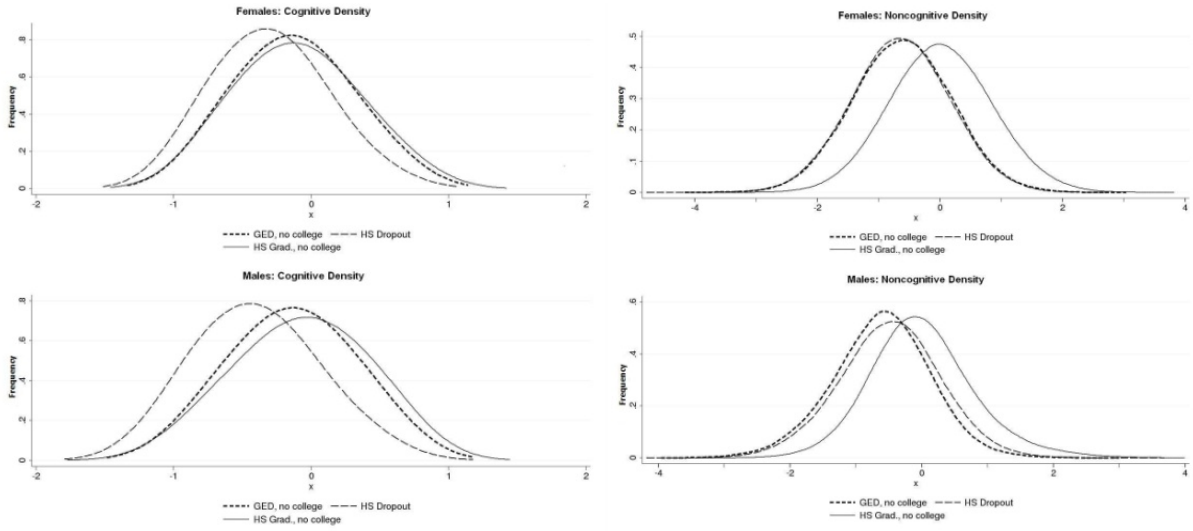
Note: The figure displays standardized regression coefficients from multivariate of years of school attended on the Big Five and intelligence, controlling for age and age-squared. The bars represent standard errors. The Big Five coefficients are corrected for attenuation bias. The Big Five were measured in 2005. Years of schooling were measured in 2008. Intelligence was measured in 2006. The measures of intelligence were based on components of the Wechsler Adult Intelligence Scale (WAIS). The data is a representative sample of German adults between the ages of 21 and 94.
Source: *GSOEP, waves 2004-2008*, own calculations.

Nevertheless, the components of Openness to Experience representing an intrinsic interest in ideas and learning may affect educational attainment not measured by total years of schooling such as the student's difficulty with classes and attendance. Consistent with this supposition, a longitudinal study of talented high school students showed that when controlling for PSAT score, students who expressed more intrinsic motivation in learning took more difficult math courses one year later $(\beta = 0.30, p < 0.05)$, two years later $(\beta = 0.31, p < 0.05)$, and three years later $(\beta = 0.26, p < 0.10)$ but did not have higher grades (Wong and Csikszentmihalyi [1991]). Likewise, of the Big Five, Openness to Experience is most consistently associated with fewer contemporaneously measured school absences in 7th grade $(r = -0.31, p < 0.01)$, 10th grade $(r = -0.19, p < 0.01)$, and 12th grade $(r = -0.27, p < 0.01)$ (Lounsbury, Steel, Loveland et al. [2004]). Still, interest in learning is not the whole story. Using prospective data, Lleras [2008] finds that controlling for cognitive ability, three Conscientious behaviors (completing homework, working hard, arriving promptly to class) in tenth grade predicted educational attainment ten years later, whereas relating well to others, a behavior related to Extraversion and Agreeableness, did not.

Examining discrete educational decisions, rather than total years of education, gives a more nuanced picture. The decision to obtain a GED is a particularly telling example. Many view GED certification as equivalent to earning a high school diploma. Indeed GED recipients have the same distribution of measured achievement test scores as high school graduates who do not attend college. However, controlling for cognitive ability, GED recipients have lower hourly wages and annual earnings and attain fewer years of education, suggesting they may "lack the abilities to think ahead, to persist in tasks, or to adapt to their environments" (Heckman and Rubinstein [2001, p. 146]). Figure 10, taken from Heckman, Humphries, Urzua et al. [2010],

shows that GED recipients have cognitive skills similar to students who obtain high school diplomas but do not attend college. However, GED recipients have noncognitive skills (personality traits) similar to high school dropouts.

Figure 10. Distribution of Cognitive and Non-Cognitive Skills by Education Group



Notes: The data come from the National Longitudinal Study of Youth 1979 (no college sample, all ethnic grous). The distributions above represent noncognitive ability factors estimated using measures of early violent crime, minor crime, marijuana use, regular smoking, drinking, early sexual intercourse, and educational attainment as laid out in Hansen, Heckman and Mullen [2004]. The sample is restricted to the cross-sectional subsample for both males and females. Distributions show only those with no post-secondary educational attainment. The noncognitive ability factors are separately normalized to be mean zero standard deviation one.
Source: Reproduced from Heckman, Humphries, Urzua et al. [2010].

Supporting the evidence from the GED program that personality plays an important role in educational attainment in adolescence, several prospective studies have shown that facets of Conscientiousness (e.g., self-control, distractibility) and facets of Neuroticism (e.g., internal locus of control) predict successful graduation from high school (Bowman and Matthews [1960]; Gough [1964]; Hathaway, Reynolds and Monachesi [1969]; Janosz, LeBlanc, Boulerice et al. [1997]; Kelly and Veldman [1964]; Whisenton and Lorre [1970]). Table 11 presents findings

from three more recent studies examining the relationship between locus of control, a trait related to Emotional Stability, and high school graduation. While the level of statistical significance varies across studies, the studies report remarkably similar estimates. When controlling for basic demographics a one standard deviation increase in locus of control is associated with effects of between 4.5 and 6.8 percentage point increase in graduating from high school. Two of the studies control for cognitive ability and find that the coefficient drops to between 1.4 and 1.5. However, the measures of cognitive ability (course grades and AFQT score) are partly determined by locus of control, as discussed later in this section.

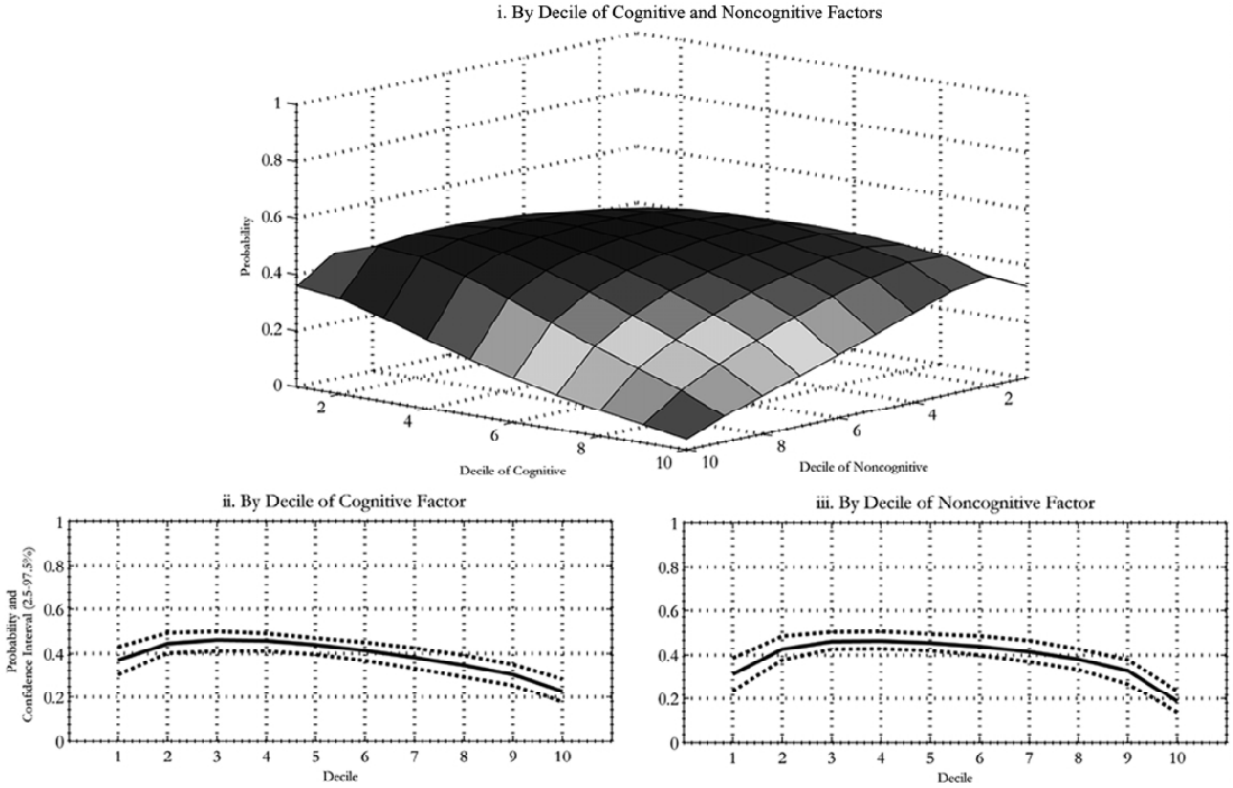
Table 11. The Relationship between Probability of High School Graduation and Locus of Control

Source	Sample	Timing of Measurement and Outcome	Controls	Metric	Results
Báron and Cobb-Clark [2010]	Australians born in 1987 or 1988 (N=2,065)	Contemporaneous	welfare receipts, family structure, sex, parental education, parental immigration status, parental involvement in education, indigenous background, and born early for their grade	The effect of a standard deviation increase in locus of control on the probability of high school graduation (b)	Locus of control 4.5*
Cebi [2007]	Nationally representative sample of students in the US (N=1,394)	Locus of control was measured in 10th or 11th grade	(1) race, gender, urban, parental education, family structure (2) race, gender, urban, parental education, family structure, home life, AFQT	The effect of a standard deviation increase in locus of control on the probability of high school graduation (b)	Locus of control (1) 4.6*** Locus of control (2) 1.5
Coleman and DeLeire [2003]	Nationally representative sample of students in the US (N= (1) 13,720 and (2) 12,896)	Locus of control was measured in 8th grade	(1) race, gender (2) race, gender, 8th grade Math Score, 8th grade reading score, 8th grade GPA, parent's education, parenting controls, family structure	The effect of a standard deviation increase in locus of control on the probability of high school graduation (b)	Locus of control (1) 6.8 Locus of control (2) 1.4**

Notes: The (#) in the control column indicates the controls used in different specifications. The (#) preceding the number of observations and estimate indicates the model used.
*statistically significant at 10 percent level; ** statistically significant at 5 percent level; *** statistically significant at 1 percent level

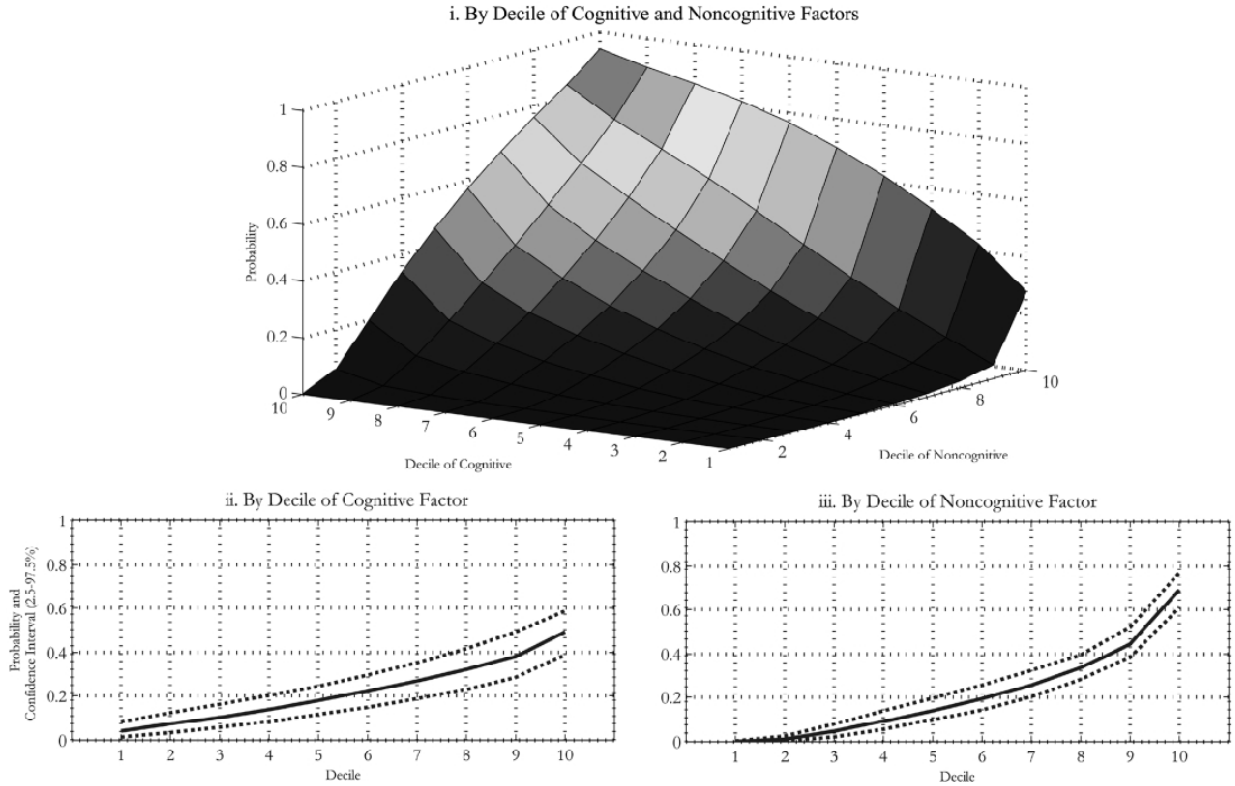
Several recent studies using methods that address measurement error and reverse causality corroborate the evidence that traits related to Neuroticism affect educational attainment. For example, Heckman, Stixrud and Urzua [2006] account for the effect of family background on test scores; correct for the influence of schooling on personality; and address measurement error in test scores. (Their estimates of the effect of schooling on these traits and on cognitive measures are discussed in Section 8.) Figure 11 shows that better adolescent personality traits – as measured by locus of control and self-esteem – increases the probability of graduating from high school (and stopping at high school) for males at the lowest quantiles of the personality distribution. However, at the higher quantiles, the probability of stopping education at high school graduation is decreasing in non-cognitive skills, because those students continue on to college. As discussed in Section 3, the effects of traits on outcomes need not be monotonic. As Figure 12 shows, both higher cognitive ability and better personality traits have strong effects on graduating from a 4-year college at all deciles. Moving from the lowest decile to the highest decile in the noncognitive skill distribution increases the probability of graduating from college more than a similar change in the cognitive skill distribution. These examples show why considering broad measures of education might obscure important relationships between skills and educational attainment and why assuming a linear – or even monotonic – relationship between skills and educational attainment might be incorrect. Applying similar methods that account for reverse causality, Piatek and Pinger [2010] find that moving from the first to the last decile of the locus of control distribution results in a 0.30 point increase in the probability of attaining higher education for males and a 0.23 increase for females.

Figure 11. Probability of Being a High School Graduate at Age 30, males



Notes: The data are simulated from the estimates of the model and the NLSY79 sample. Higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals. The upper curve is the joint density. The two marginal curves (ii) and (iii) are evaluated at the mean of the trait not being varied.
Source: Heckman, Stixrud and Urzua [2006, Figure 19].

Figure 12. Probability of being a 4-year-college graduate at age 30, males



Notes: The data are simulated from the estimates of the model and the NLSY79 sample. Higher deciles are associated with higher values of the variable. The confidence intervals are computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals. The upper curve is the joint density. The two marginal curves (ii) and (iii) are evaluated at the mean of the trait not being varied.
Source: Heckman, Stixrud and Urzua [2006, Figure 21].

Cunha, Heckman and Schennach [2010] use a dynamic factor model (described in Section 8) to investigate the development of both cognitive skills and personality traits during childhood, allowing for endogenous investment in skills and dynamic complementarities. They find that adolescent personality – as measured by a variety of behavior inventories – accounts for 12% of the variation in educational attainment, whereas adolescent cognitive ability accounts for 16% of the variability.

A separate, but related literature examines the importance of early attention (a trait related to Conscientiousness) and aggression (a trait related to low Agreeableness) in

determining graduation from high school. Some studies find that aggression is particularly important compared to attention. Duncan and Magnuson [2010] find that when controlling for intelligence and demographic variables, antisocial behavior, but not attention measured in childhood, predicts high school completion. Likewise, Fergusson and Horwood [1998] find that teacher and parent ratings of conduct problems at age 8 are negatively related to predicted high school completion at age 18. In contrast, Vitaro et al. [2005] examines individuals in a population-based sample of Quebec children and find that kindergarten teacher ratings of hyperactivity-inattention (inversely) predicted completion of high school better than did aggressiveness-opposition. Both attention and aggression likely play roles, but there is no consensus on their relative importance.

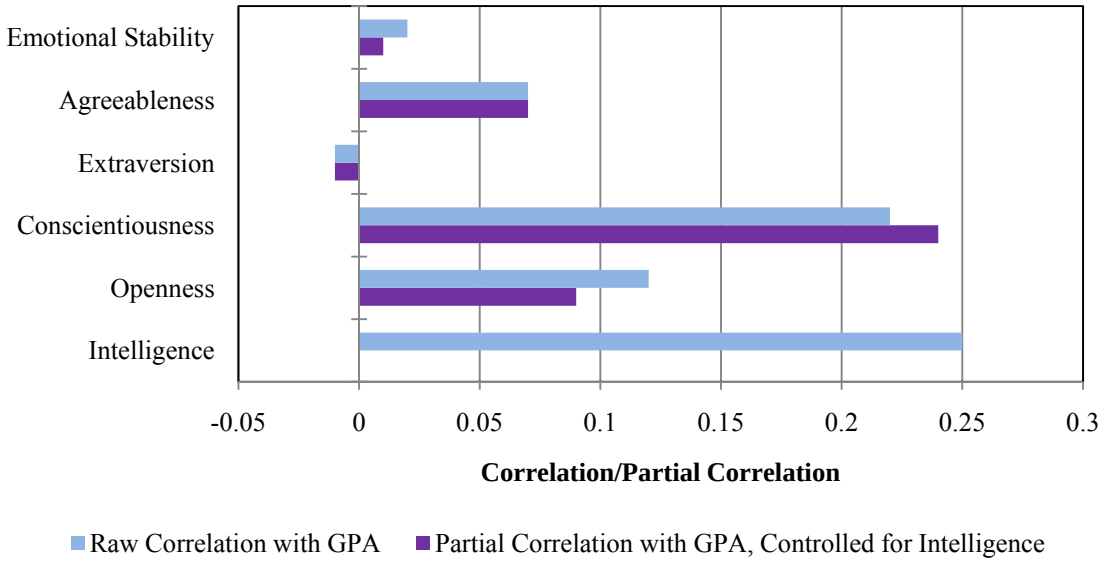
In sum, traits related to Big Five Openness to Experience and Conscientiousness are particularly important in determining how many total years of education individuals complete in their lifetimes. However, Openness to Experience might partially serve as a proxy for cognitive ability. Two traits related to Neuroticism, locus of control and self-esteem, play a particularly important role for adolescent schooling decisions. Their effects differ across schooling attainment levels, suggesting that analysts should be wary of using years of schooling attained as the outcome variable compared to using the probability of attainment at different grades. Attention and early aggression, traits related to Conscientiousness and Agreeableness, are also predictive.

Course grades

Conscientiousness is the most robust Big Five predictor of course grades. Poropat [2009] conducted a meta-analysis of Big Five personality traits and course grades in primary, secondary,

and post-secondary education, presented in Figure 13. Associations between grades and Conscientiousness are almost as large those between grades and cognitive ability. Associations with grades are substantially smaller for other Big Five factors, the largest of which is Openness to Experience.

Figure 13. Correlations of the Big Five and Intelligence with Course Grades



Notes: All correlations are significant at the 1% level. The correlations are corrected for scale reliability and come from a meta analysis representing a collection of studies representing samples of between N=31,955 to N=70,926, depending on the trait. The meta-analysis did not clearly specify when personality was measured relative to course grades.
Source: Poropat [2009].

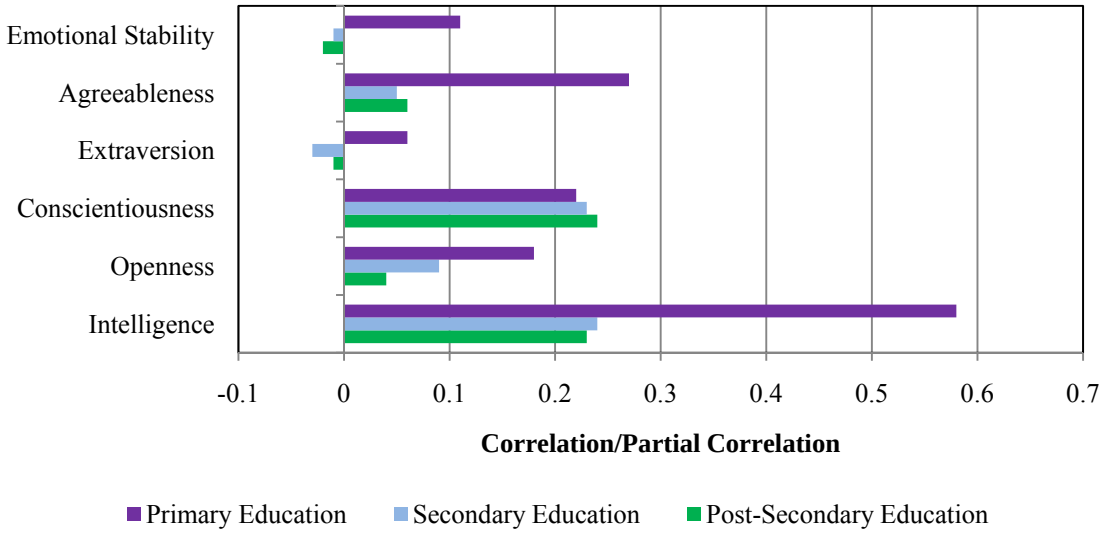
A few prospective, longitudinal studies, have estimated the effect of Conscientiousness on course grades when controlling for baseline levels of grades. These studies help isolate the effects of personality on grades by reducing the potential for omitted variable bias and misleading halo effects – the propensity for teachers to favor students based on traits unrelated to academic achievement. In general, these studies support the conclusions of less rigorously controlled studies. For instance, in a sample of American middle school students, self-control predicts report card grades, controlling for both general intelligence and baseline grades

(Duckworth and Seligman [2005]). Likewise, Duckworth, Tsukayama, and May [in press] use longitudinal hierarchical linear models to show that changes in self-control predict subsequent changes in report card grades. In a sample of Chinese primary school children, effortful control predicted report card grades when controlling for baseline grades (Zhou, Main and Wang [2010]).

Figure 14 shows that associations between course grades and personality and cognitive ability are generally stronger in the primary grades, a pattern consistent with the speculation of intelligence researchers e.g., Jensen [1980] that diminishing estimated predictive validity by grade arises from censoring. A notable exception to this trend is Conscientiousness, which is as strongly associated with course grades at higher grade levels. Censoring was not accounted for in Poropat [2009], presumably because norms for variance in representative samples are generally unavailable for personality measures (Duckworth [2009]). Thus, if censoring on cognitive and personality traits attenuates observed associations with course grades among students at higher grade levels, Conscientiousness might be even *more* predictive of course grades as students progress through the education system.[158] Consistent with this speculation, in a prospective study of an entire cohort of Belgium's medical students, the correlation (corrected for censoring) of Conscientiousness for GPA increased from $r = 0.18$ in the first year to $r = 0.45$ in the seventh and final year (Lievens, Dilchert and Ones [2009]).[159]

[158] Flinn and Heckman [1982].
[159] The values were corrected for truncation.

Figure 14. Correlations with Course Grades by Level of Education



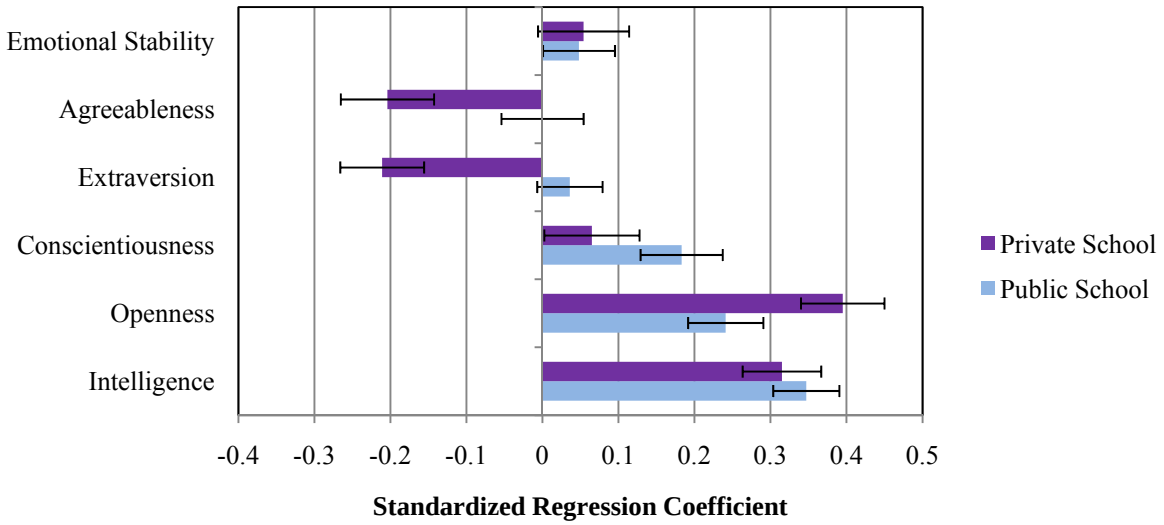
Notes: The reported values for The Big Five are partial correlations, controlled for intelligence. The meta-analysis did not address when personality was measured relative to course grades.
Source: Poropat [2009].

Overall, the empirical evidence suggests that Conscientiousness may be as consequential as cognitive ability in predicting and possibly causing higher course grades. Why? Even intelligent students might not enjoy the work (Wong and Csikszentmihalyi [1991]). To a large degree, the many tasks required of a student to earn high course grades (e.g., concentrating on difficult new concepts, attending to the teacher rather than joking with classmates, practicing skills repeatedly to the point of fluency, working on homework alone rather than socializing with friends) all yield long-term rewards at the expense of short-term comfort and pleasure. Indeed, there is evidence that the association between Conscientiousness and course grades is mediated by positive study habits and attitudes, effort, and prosocial behavior in the classroom (Credé and Kuncel [2008]; Lubbers, Van Der Werf, Kuyper et al. [2010]; Noftle and Robins [2007]; Valiente, Lemery-Chalfant and Castro [2007]; Valiente, Lemery-Chalfant, Swanson et al. [2008]).

Standardized achievement test scores

Like course grades, standardized achievement test scores reflect a student's acquired skills and knowledge. Thus, dimensions of personality that influence the acquisition of skills and knowledge should predict both outcomes. One might expect, therefore, that traits related to Conscientiousness predict achievement test scores. While studies using standardized achievement tests are less common than studies using grades, ample empirical evidence shows that aspects of personality predict both metrics of performance. As show in Section 5, two traits related to Neuroticism, locus of control and self-esteem, explain much of the variance of the Armed Forces Qualification Test (AFQT), an achievement test which is often used as a measure of pure intelligence in economics studies. Similarly, Figure 14 shows that in samples from three New York City middle schools, controlling for IQ, Openness to Experience is associated with Standardized Achievement Test Scores.

Figure 15. Associations with Standardized Achievement Test Scores



Notes: The values represent standardized regression coefficients in models including personality, IQ, gender, and ethnicity. The bars represent standard errors around the estimate. IQ is measured using Raven's Progressive Matrices. The achievement tests are based on the Comprehensive Testing Program test in the private school sample and the English/Language Arts and Mathematics standardized achievement test in the public school sample.
Source: Data collected by Authors. Author's own calculation.

Martin and colleagues were among the first to demonstrate that teacher and parent ratings of early childhood persistence, (low) distractibility, and (low) activity prospectively predict both course grades and standardized achievement test scores (see Martin [1989 for a summary]). Likewise, in a representative sample of Baltimore first graders, teacher ratings of attention span-restlessness in first grade-predicted both course grades and standardized achievement test scores four years later (Alexander, Entwisle and Dauber [1993]).

More recently, in a sample of preschool children from low-income homes, parent and teacher ratings of effortful control, a facet of Conscientiousness, accounted for unique variance in standardized achievement test scores in kindergarten, even after controlling for general intelligence (Blair and Razza [2007]). Similarly, in a sample of kindergarteners, teacher and parent ratings of effortful control predicted performance on standardized achievement tests six months later when controlling for both verbal intelligence and family socioeconomic status (Valiente, Lemery-Chalfant and Swanson [2010]). Teacher ratings of inattention at the beginning

of the school year predicted standardized achievement test scores at the end of the school year in a sample of fourth graders (Finn, Pannozzo and Voelkl [1995]).

Task measures of effortful control, a trait related to Conscientiousness, predict performance on standardized achievement tests much later in life. For instance, the number of seconds a child waits for a more preferred treat in the preschool delay of gratification paradigm predicts the SAT college admission test more than a decade later, with raw correlations of $r = 0.42$ for the verbal section and $r = 0.57$ for the quantitative section (Mischel, Shoda and Rodriguez [1989]). The Head-to-Toes and Head-Toes-Knees-Shoulders tasks requires young children to inhibit automatic responses, pay attention, and keep instructions in working memory (e.g., to touch their heads when the experimenter says "touch your toes") (Ponitz, McClelland, Jewkes et al. [2008]; Ponitz, McClelland, Matthews et al. [2009]). Performance on this brief task predicts later performance on standardized achievement tests (McClelland, Cameron, Connor et al. [2007]).

Perhaps most conclusively, Duncan and colleagues [2007] analyzed six large, longitudinal datasets and found that school-entry attention skills, measured variously by task and questionnaire measures, prospectively predict achievement test scores, even when controlling for school-entry academic skills. In contrast, internalizing behavior (e.g. depression, anxiousness, withdrawal) and externalizing behaviors (e.g. aggression, hyperactivity, antisocial behavior) at school-entry do not reliably predict standardized achievement test scores. Attention skills are related to Conscientiousness; externalizing behavior is related to Agreeableness and Conscientiousness; and internalizing behaviors are related to Neuroticism.

In sum, traits related to Conscientiousness play an important role in predicting achievement tests above and beyond cognitive ability. Nevertheless, as discussed in Section 6

time discounting and risk aversion also relate to test score performance, suggesting that both personality-related abilities and preferences are important determinants of outcomes, consistent with the economic model presented in Section 3. In contrast to educational attainment, traits related to Emotional Stability (the opposite of Neuroticism) seem less important for test performance.

Where course grades and standardized achievement test scores diverge

Course grades and standardized test scores are generally highly correlated. Each form of assessment provides reciprocal evidence on the validity of the other. Willingham, Pollack and Lewis [2002] estimate a raw correlation of $r = 0.62$ $(p < 0.01)$ between total grade average and NELS score.[160] This strong association —and the objective of each form of assessment to gauge student learning—explains why standardized achievement tests and grades are widely assumed to be "mutual surrogates; that is, measuring much the same thing, even in the face of obvious differences" Willingham, Pollack and Lewis [2002, p. 2]. What are these differences, and how might the contribution of personality to performance vary accordingly?

Standardized achievement tests are designed to enable apples-to-apples comparisons of students from diverse contexts. To this end, standardized achievement tests are uniform in subject matter, format, administration, and grading procedure across all test takers. A course grade, on the other hand, represents a "teacher's judgment as to how well a student has fulfilled the implicit local contract between teacher and student" (Willingham, Pollack and Lewis [2002, p. 28]).

[160] The correlations were even higher when the test and grades were based on similar subject matter.

The power of standardized achievement tests to predict later academic and occupational outcomes is well-established (Kuncel and Hezlett [2007]; Sackett, Borneman and Connelly [2008]; Willingham [1985]). Nevertheless, cumulative high school GPA predicts graduation from college dramatically better than SAT/ACT scores do, even without adjusting for differences in high school quality (Bowen, Chingos and McPherson [2009b]). Similarly, high school GPA more powerfully predicts college rank-in-class (Bowen, Chingos and McPherson [2009b]; Geiser and Santelices [2007]).

Perhaps more important than *which* measure of academic achievement – course grades or standardized achievement test scores – is more predictive of later outcomes is *why* these outcomes are related but not entirely interchangeable. Bowen and colleagues [2009b] speculate that aspects of Conscientiousness seem *differentially* essential to earning strong course grades because of what is required of students to earn them. Standardized achievement tests, in contrast to teacher-designed quizzes, exams, homework assignments, and long-term projects, challenge students to solve *relatively* novel problems. It is therefore not surprising that Frey and Detterman [2004] found a raw correlation of $r = 0.82 \ \ (p < 0.01)$ between SAT scores and performance on the ASVAB, an aptitude and achievement test developed for the United States Army. In a separate sample, Frey and Detterman found a correlation of $r = 0.72 \ \ (p < 0.01)$ between SAT scores and IQ when accounting for censoring. In contrast, the correlation between GPA and IQ is $r = 0.23 \ \ (p < 0.01)$ (Poropat [2009]).

In three longitudinal, prospective studies of middle school students, Duckworth, Quinn and Tsukayama [2010] compare the variance explained in year-end standardized achievement test scores and GPA by self-control (a facet of Conscientiousness) and fluid intelligence measured at the beginning of the school year. For example, in a national sample of children, 4th

grade self-control was a stronger predictor of 9th grade GPA $(\beta = 0.40, p < 0.001)$ than was 4th grade IQ $(\beta = 0.28, p < 0.001)$. In contrast, 4th grade self-control was a weaker predictor of 9th grade standardized test scores $(\beta = 0.11, p < 0.05)$ than was 4th grade IQ $(\beta = 0.64, p < 0.001)$.[161] These findings are consistent with those of Willingham, Pollack, and Lewis [2002], who show that conscientious classroom behaviors are more strongly associated with GPA than with standardized achievement test scores. Likewise, Oliver, Guerin, and Gottfried [2007] found that parent and self-report ratings of distractibility and persistence at age 16 predicted high school and college GPA, but not SAT test scores. Table 12 presents results showing that Conscientiousness and SAT scores are similarly predictive of college GPA. However, in each of the studies below, Conscientiousness was measured in college which presents problems for a causal interpretation of this evidence due to the potential for reverse causality.

[161] β denotes a standardized regression coefficient (regressors and dependent variables divided by standard deviations).

Table 12. The Predictive Power of Conscientiousness Relative and SAT Scores for College GPA

Source	Sample	Timing of Measurement and Outcome	Controls	Metric	Results	
Conard [2005]	University students in the US (N=186)	College GPA and SAT were both self-reported during college. Personality was measured in college.	Class Attendance	Standardized Regression Coefficient (β)	SAT Total Conscientiousness	0.27** 0.30**
Noftle and Robins [2007]	University students in the US (N=10,472)	College GPA and SAT were both self-reported during college. Personality was measured in college.	Gender, Other Big Five Traits	Standardized Regression Coefficient (β)	SAT Verbal SAT Math Conscientiousness	0.19*** 0.16*** 0.24***
Noftle and Robins [2007]	University students in the US (N=465)	College GPA and SAT were both self-reported during college.[1] Personality was measured in college.	Gender, Other Big Five Traits	Standardized Regression Coefficient (β)	SAT Verbal SAT Math Conscientiousness	0.28*** 0.28*** 0.18***
Noftle and Robins [2007]	University students in the US (N=444)	College GPA and SAT were both self-reported during college. Personality was measured in college.	Gender, Other Big Five Traits	Standardized Regression Coefficient (β)	SAT Verbal SAT Math Conscientiousness	0.18*** 0.25*** 0.22***
Wolfe and Johnson [1995]	University students in the US (N=201)	GPA and SAT were provided by the Colleges's Record Office. Personality was measured in college.	High School GPA	Standardized Regression Coefficient (β)	SAT Total Conscientiousness	0.23*** 0.31***

Notes: (1) Self-reported SAT scores and those obtained from college records were highly correlated $(r = 0.92)$. Self-reported GPA and that obtained from college records were highly correlated $(r = 0.89)$.
*statistically significant at the10 percent level; **statistically significant at the 5 percent level; ***statistically significant at the 1 percent level

In sum, standardized achievement tests and teacher-assigned course grades both reflect students' accumulated knowledge and skill. However, they differ in important ways. The benefits of Conscientiousness, which inclines students to more productive work habits, seem greater for course grades than for test scores. This finding might explain why girls, who are higher than boys in Conscientiousness, reliably earn higher grades than boys in every subject from primary school through college – but do not reliably outperform boys on either standardized achievement or intelligence tests (Duckworth and Seligman [2006]).

7.B. Labor Market Outcomes

"Eighty percent of success is showing up."

– Woody Allen

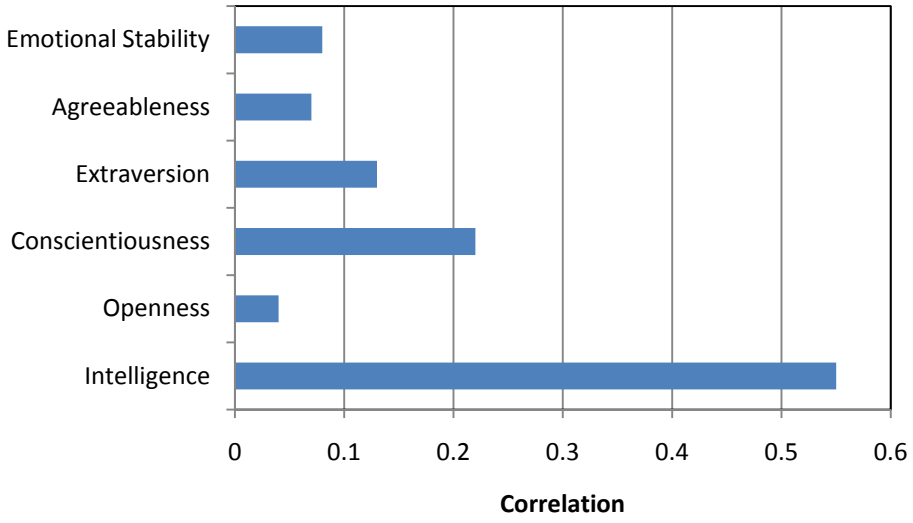
It is obvious that personality traits affect labor market outcomes. Showing up is required for completing a task. Precisely quantifying the direct effects of personality, however, is more difficult.[162] Recently, social scientists have started to tackle the problem and, in general, find that of the Big Five, Conscientiousness and traits associated with Neuroticism (locus of control and self-esteem) play a particularly important role in determining job performance and wages. The evidence suggests multiple channels of influence, including occupational matching, incentive scheme selection, absenteeism, turnover, and job search.

Aspects of job performance are related to academic performance. For example, both types of tasks require completing work on a schedule and involve intelligence to varying degrees. It is therefore not surprising, that, as with academic performance, numerous studies and meta-analyses have found that Conscientiousness is associated with job performance and wages (Nyhus and Pons [2005]; Salgado [1997]; Hogan and Holland [2003]; Barrick and Mount [1991]). Figure 16 presents correlations of the Big Five and IQ with job performance. Of the Big Five, Conscientiousness is the most associated with job performance but is about half as predictive as IQ. Conscientiousness, however, may play a more pervasive role than IQ. The importance of IQ increases with job complexity, defined as the information processing requirements of the job: cognitive skills are more important for professors, scientists, and senior managers than for semi-skilled or unskilled laborers (Schmidt and Hunter [2004]). In contrast,

[162] Even Allen admits that his estimate is partially based on the fact that "eighty" has better cadence than "seventy" (Safire [1989]). Bowles, Gintis and Osborne [2001b] discuss evidence on the association between personality traits and labor market outcomes.

the importance of Conscientiousness does not vary much with job complexity (Barrick and Mount [1991]), suggesting that it pertains to a wider spectrum of jobs. Causality remains an open question. The raw correlations presented in Figure 16 do not account for reverse-causality, and the authors do not clearly delineate when the measures of personality were taken.

Figure 16. Associations with Job Performance



Notes: The values for personality are correlations that were corrected for sampling error, censoring, and measurement error. Job performance was based on performance ratings, productivity data and training proficiency. The authors do report the timing of the measurements of personality relative to job performance. Of the Big Five, the coefficient on Conscientiousness is the only one that is statistically significant with a lower bound on the 90% credibility value of 0.10. The value for IQ is a raw correlation.
Sources: The correlations reported for personality traits come from a meta-analysis conducted by Barrick and Mount [1991]. The correlation reported for IQ and job performance come from Schmidt and Hunter [2004].

Facets related to Emotional Stability (the opposite of Neuroticism) are also important for labor market success. Accounting for reverse-causality, however, is particularly important, because strong evidence suggests labor market participation can affect traits related to Neuroticism (See the discussion of Gottschalk [2005] in Section 8). Several studies have addressed this problem by using pre-labor market measures of personality and find that locus of control and self-esteem, two facets of Emotional Stability, predict wages in the NLSY (Judge

and Hurst [2007a]; Drago [2008]; Duncan and Dunifon [1998]). Table 13 presents results from Heckman, Stixrud and Urzua [2006] suggesting that standardized adolescent measures of locus of control and self-esteem predict adult earnings to a similar degree as cognitive ability. However, the effects vary across educational labor markets. In general, noncognitive ability is relatively more predictive across all levels of education than is cognitive ability, in line with the results that Conscientiousness predicts job performance across all levels of job complexity while cognitive ability does not.

Table 13. Estimated Coefficients of Cognitive and Noncognitive Factors for Log Hourly wages

	Males		Females	
Schooling Level	Cognitive	Noncognitive	Cognitive	Noncognitive
High school dropout	.113	.424	.322	.208
	(.076)	(.092)	(.125)	(.103)
GED	.175	.357	.020	.242
	(.107)	(.117)	(.137)	(.153)
High school graduate	.259	.360	.341	.564
	(.041)	(.059)	(.049)	(.056)
Some college, no degree	.069	.401	.093	.569
	(.086)	(.110)	(.084)	(.116)
2-year-college degree	.039	.368	.206	.279
	(.138)	(.209)	(.096)	(.145)
4-year-college degree	.296	−.060	.290	.379
	(.075)	(.175)	(.066)	(.103)

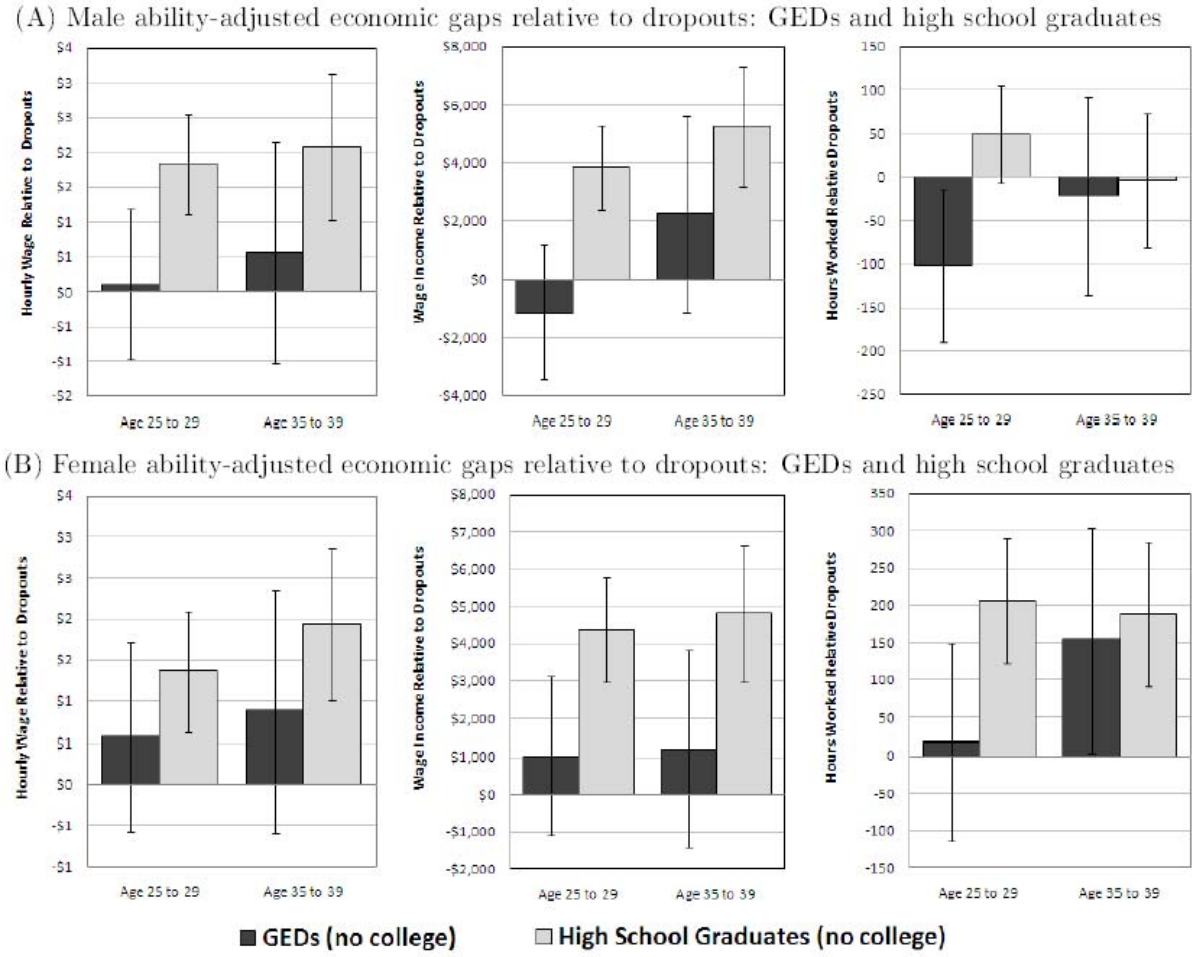
Notes: Standard errors are in parentheses. Sample from NLSY79 males and females at age 30. We exclude the oversample of blacks, Hispanics, and poor whites, the military sample, and those currently enrolled in college. The cognitive measure represents the standardized average over the raw ASVAB scores (arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge, and coding speed). The noncognitive measure is computed as a (standardized) average of the Rosenberg Self-Esteem Scale and Rotter Internal-External Locus of Control Scale. The model also includes a set of cohort dummies, local labor market conditions (unemployment rate), and the region of residence.
Source: Heckman, Stixrud and Urzua [2006].

Recent evidence, however, suggests that locus of control affects wages mostly through the channel of educational attainment. Piatek and Pinger [2010] apply statistical methods to account for reverse causality to a German sample and find that for males moving from the first to the last decile of the locus of control distribution increases earnings by 36%. The effect, however, becomes statistically insignificant when controlling for education suggesting that education might be the channel through which locus of control affects earnings. [163] Further highlighting the possible role of educational decision, Figure 17 shows that GED recipients – who have lower levels of non-cognitive skills but comparable levels of cognitive skills (see the

[163] They use pre-labor market measures of locus of control, but contemporaneous measures for education.

previous section) – have lower wages, total age income, and work fewer hours relative to high school graduates, when controlling for ability.

Figure 17. Ability-adjusted economic gaps relative to dropouts: GEDs and high school graduates for males (A) and females (B)



Notes: Regressions control for baseline AFQT scores, age, mother's highest grade completed, and dummies for urban residence at age 14, southern residence at age 14, and race. Baseline test scores are estimated using the procedure of Hansen et al. 2004 as implemented in Carneiro, Heckman and Vytlacil [2005]. The regressions use the cross-sectional subsample and minority oversamples of the NLSY79 data. The estimation sample is restricted to individuals who never attend college and who have not been incarcerated. Regressions for hourly wage and hours worked are restricted to those reporting more than $1/hour and less than $100/hour, and individuals working less than 4,000 hours in a given year. Wage income regressions are restricted to individuals reporting wage incomes between $1,000/year and $100,000/year. All monetary values are in 2005 dollars. Standard errors are clustered by individual.
Source: Data come from National Longitudinal Survey of Youth 1979 (NLSY79).

The previous studies have shown that personality is associated with wages, but do not explain *why* they are associated other than through the channel of educational attainment. Some mechanisms might be absenteeism, job turnover and unemployment. Indeed, controlling for basic demographics, employment history and health, Störmer and Fahr [2010] estimate that a standard deviation increase in Emotional Stability and Agreeableness is associated with 12% $(p < 0.01)$ and 9% $(p < 0.05)$ fewer absent days for men and a standard deviation increase in Openness to Experience is associated with 13% $(p < 0.01)$ more absent days for women. The study uses contemporaneous measures of personality and absenteeism. [164]

Personality plays a role outside of formal employer-employee relationships. Self-employed workers, with either very low or high levels of risk-aversion tend to remain self-employed for a shorter time, suggesting that they are less suited to self employment (Caliendo, Fossen and Kritikos [2008]).[165] Personality could directly affect the duration of unemployment spells. For example, Caliendo, Cobb-Clark and Uhlendorff [2010] extend standard job search models by examining whether a higher locus of control increases the perceived marginal benefit of exerting search effort, so that people with a more internal locus of control will search more intensely and will have a higher reservation wage. Supporting their theory, a one standard deviation increase in internal locus of control was associated with a 1.9% increase in the reservation wage (p<0.01) and 5.3% increase in the number of job applications submitted (p<0.01), controlling for demographic characteristics and past employment history. The associations were partially mediated when controlling for the Big Five, suggesting that locus of control overlaps with the Big Five as suggested in Section 5. While the measures were

[164] All other Big Five traits were not statistically significant at the 10% level.
[165] Caliendo, Fossen, and Kritikos use measures of risk-aversion from 2004 and employment status from 2000-2005, assuming that risk-aversion is constant during this period.

contemporaneous, the respondents became unemployed near the time that the locus of control was measured, potentially limiting the role of reverse causality.

Personality traits also affect occupational choice. From an economic perspective, some personality traits that reflect ability might be valued more highly in some occupations, and, on the supply side, people with certain personality traits that relate to preferences might value the non-pecuniary benefits associated with particular occupations. Supporting this notion, Conscientiousness (Barrick and Mount [1991]), locus of control and self-esteem (Heckman, Stixrud and Urzua [2006]) predict sorting into occupation. However, these studies use relatively broad occupational categories that might obfuscate more nuanced influences of personality. Analyzing eighteen occupational categories, Cobb-Clark and Tan [2009] find that for men a one standard deviation increase in Agreeableness is associated with a 2.8% decrease in being a manager $(p < 0.01)$ and a 2.9% decrease in being a business professional $(p < 0.01)$, and a standard deviation increase in internal locus of control is associated with 2.8% decrease in being a manager $(p < 0.01)$. In contrast, for women a one standard deviation increase in Openness to Experience is associated with a 2.5% increase in being a manager (p<0.01).[166,167]

Furthermore, the value of cognitive ability and personality differs based on occupation just as they do by education. Cattan [2010] estimates a model of comparative advantage and finds different skills are valued differently, depending on the occupation. Accounting for selection, a standard deviation increase in adolescent sociability (related to Extraversion) leads to a 9% increase in the wages of managers $(p < 0.01)$ and a 4% increase in the wages of sales workers $(p < 0.05)$, but leads to a 2% and 3% *decrease* in wages of professionals $(p < 0.10)$ and

[166] The data for occupational categories came from 2001-2006, whereas locus of control was measured in 2003-2004 and The Big Five were measured in 2005. Thus these concerns about reverse causality are valid.
[167] They find other statistically significant results at the 5% and 10% levels which we omit for brevity.

clerical workers $(p < 0.01)$, and has no significant impact on the wages of blue-collar workers. Self-esteem and locus of control are positively valued in all occupations, but the magnitudes also depend on the occupation. The effects of ability need not be uniform on wages across occupations.

Personality might affect not only the occupational selection, but also the type of compensation scheme selected within occupation. Dur, Non and Roelfsema [2010] extend the standard principal-agent model by allowing for workers to reciprocate positive attention from managers by working harder. Their theoretical model implies that promotions, rather than monetary incentives, should be more effective for eliciting effort from reciprocal workers. Workers self select into different compensation schemes. Supporting their model, they find that a one point increase on a seven point reciprocity scale for workers is associated with a 5 percentage point increase of having a job with promotion incentives $(p < 0.01)$. They use contemporaneous measures of reciprocity and job attributes, which could be problematic if pay-for-performance schemes affect reciprocity. Similarly, Dohmen, Falk, Huffman et al. [2009b] find that in a German sample self-reported positive reciprocity is associated with income, and employment, and working over time. Negative reciprocity tends to work in the opposite direction. As discussed in Section 6, these measures of social preference relate to personality.[168]

In sum, there is some evidence that personality affects labor market outcomes through channels other than just education. Conscientiousness and Neuroticism are associated with job performance and wages to a similar but lesser degree than cognitive ability. The personality traits are more important for people with lower levels of job complexity or education level, whereas

[168] Agreeableness and Conscientiousness are associated with more positive reciprocity and less negative reciprocity, whereas Neuroticism is associated with more negative reciprocity (Dohmen, Falk, Huffman et al. [2008]).

cognitive ability is more important at higher levels of job complexity. Nevertheless, some research suggests that facets related to Neuroticism might affect labor outcomes primarily through the channel of educational attainment. Other traits, such as Openness to Experience and Agreeableness affect more specific outcomes, such as selection into particular careers or type of compensation. Table A8 in Web Appendix A7 summarizes a variety of studies that associate personality with labor market outcomes.

7.C. Personality and Health[169]

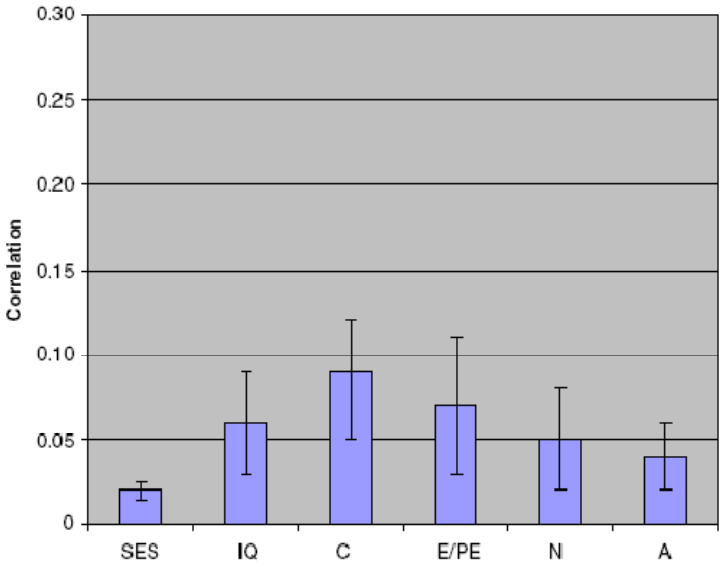
A link between personality and health has been noted for thousands of years. Hippocrates argued that an imbalance of the four temperaments would affect both personality and physical health.[170] Consistent with Hippocrates' notion, modern evidence suggests that personality predicts health. The mechanisms are relatively unexplored but some empirical evidence suggests that personality affects health-related behavior, psychological responses, and social relationships (Kern and Friedman [2010a]).

A growing body of work shows that some Big Five traits predict longevity. Friedman, Kern and Reynolds [2010] review evidence from 34 different studies on the predictive validity of Big Five personality traits, relative to that of cognitive ability and socioeconomic status, for longevity. Most studies in their meta-analysis control for relevant background factors, including gender and severity of disease. Roberts and colleagues convert the results of each study into comparable correlation coefficients. As shown in Figure 18, Conscientiousness was a stronger predictor of longevity than any other Big Five trait and a stronger predictor than either IQ or

[169] This section is a summary of Pietro Biroli's extensive discussion presented in Web Appendix A7.A.
[170] See Hampson and Friedman [2008] and Friedman [2007] for a brief historic review.

socioeconomic status.[171] In general, traits related to Conscientiousness and Openness to Experience are associated with longer lives, whereas those related to Neuroticism and Agreeableness are associated with shorter life spans (Martin, Friedman and Schwartz [2007]; Kern and Friedman [2008]; Mroczek and Spiro [2007]; Boyle, Williams, Mark et al. [2005]; Schulz, Bookwala, Knapp et al. [1996], Kubzansky, Sparrow, Vokonas et al. [2001]). The magnitudes of the relationships, however, vary across studies and not all results are replicable. While the specific channels through which personality affects longevity and health are largely unknown, several studies provide some clues.

Figure 18. Correlations of mortality with personality, IQ, and socioeconomic status (SES)



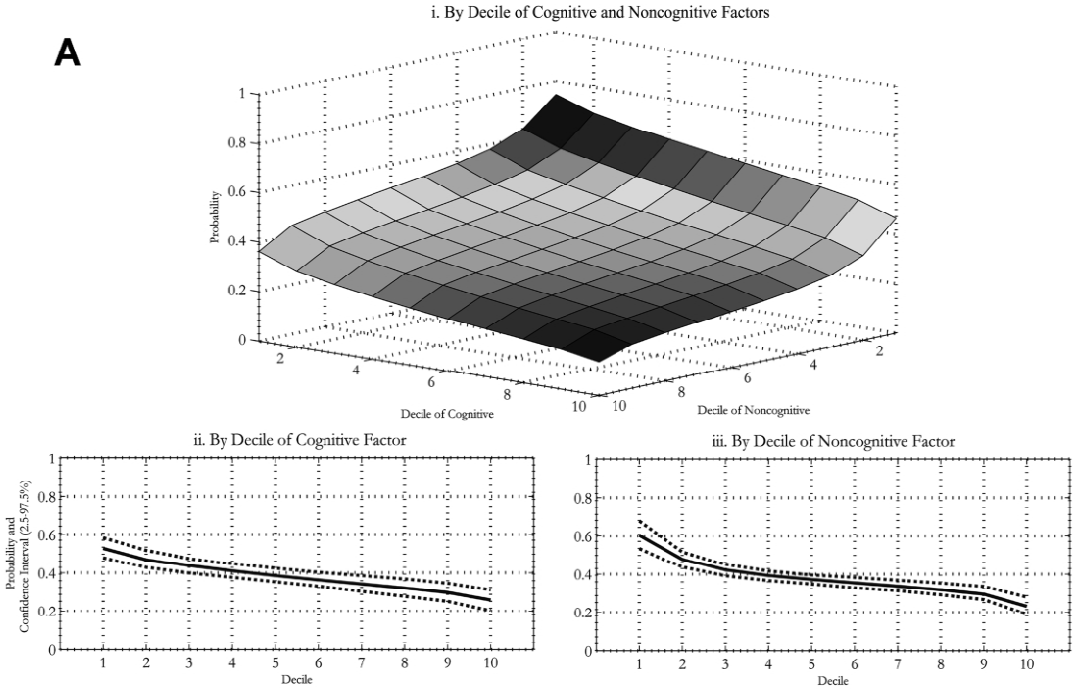
Notes: The figure represents results from a meta-analysis of 34 studies. The values represent raw correlations. The lengths of the studies represented vary from 1 year to 71 years.
Source: Roberts et al. [2007]

Personality may affect health-related behavior, such as smoking, diet, and exercise. For example, Hampson, Goldberg, Vogt et al. [2007] find that high scores of teacher-assessments of

[171] The timing of the measurements of personality relative to the outcomes varies by study.

Extraversion, Agreeableness and Conscientiousness during elementary school predict overall health status during midlife (less smoking, more exercise, better self-rated health) and indirectly affect health through educational attainment. The correlations that were statistically significant at the 5% level or less ranged from 0.06 for the effect of Extraversion on physical activity to 0.12 for the effect of Conscientiousness on self-reported health status. Both the initial level and the growth in hostility (a facet of Neuroticism) throughout elementary school predict cigarette, alcohol, and marijuana use in high school and sociability (a trait related to Extraversion) predicts drinking but not smoking (Hampson, Tildesley, Andrews et al. [2010]). As Figure 19 illustrates, Heckman, Stixrud and Urzua [2006] find that locus of control and self-esteem affect the probability of daily smoking for males. The gradient is steepest at the high and low quantiles of the locus of control and self-esteem distribution.

Figure 19. Probability of daily smoking by age 18 for males



Notes: The data are simulated from the estimates of the model and our NLSY79 sample. We use the standard convention that higher deciles are associated with higher values of the variable. The confidence intervals are

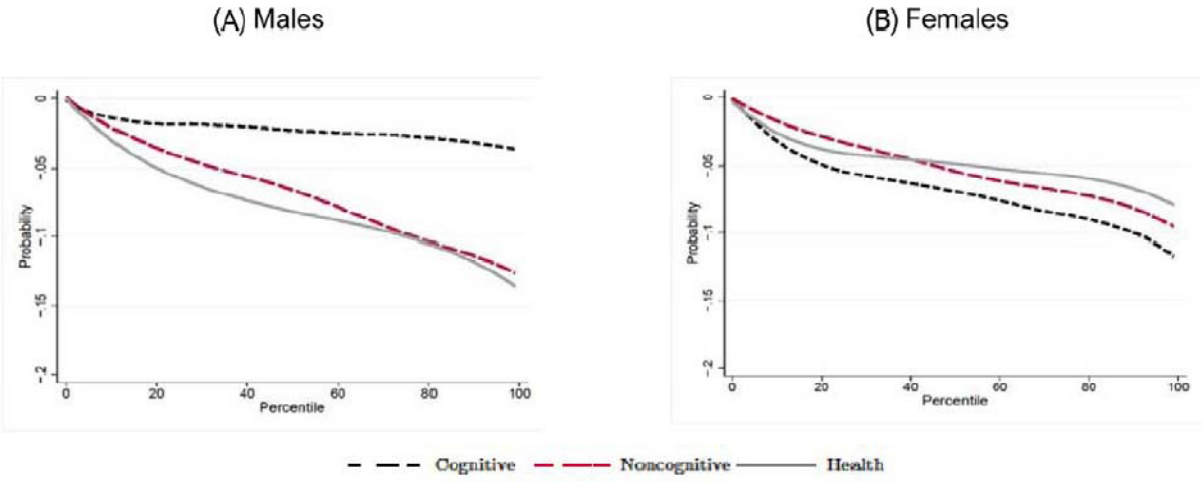
computed using bootstrapping (200 draws). Solid lines depict probability, and dashed lines, 2.5%–97.5% confidence intervals. The upper curve is the joint density. The two marginal curves (ii) and (iii) are evaluated at the mean of the trait not being varied.
Source: Heckman, Stixrud and Urzua [2006, Figure 22]

Although many studies control for socioeconomic and health factors associated with mortality, most do not explore *how* personality affects health throughout the life-cycle (Kern and Friedman [2010b]). The relationship between health and personality is complicated because health can affect personality (Pesonen, Räikkönen, Heinonen et al. [2008], Ryden, Sullivan, Torgerson et al. [2003], Sell, Tooby and Cosmides [2009] and Hoffman, Fessler, Gneezy et al. [2010]). Some studies investigate the mechanisms by considering how initial endowment of traits and health affect mid-life outcomes, such as healthy behavior and education, which in turn can influence health and longevity. For example, Gale, Batty and Deary [2008] find that a one standard deviation increase in age-10 locus of control decreases the risk of adult obesity by 8% ($p<0.05$). Similarly, Friedman, Kern and Reynolds [2010] find that in a cohort of gifted children, Conscientiousness better predicted longevity and social interactions at age 70. They find that Neuroticism is associated with worse health for women but better health for men, highlighting that personality traits do not always have a monotonic effect. These studies do not account for the possibility that health and personality exhibit dynamic complementarities over the life cycle.

Several studies have addressed the threats to causality by using structural models to estimate the life-cycle evolution of health. Using a structural model of skill expression, Conti and Heckman [2010] estimate the causal relationship between personality traits, initial health endowments and endogenous choices about schooling and post-schooling outcomes. They find that women sort into higher education based on cognitive ability, personality traits and initial health endowment. Furthermore, personality and health status measured during youth explain more than half of the difference in poor health, depression and obesity at age 30. Figure 20

shows that for males the noncognitive (personality) and health endowments are more predictive than the cognitive ones, whereas for females all three are similarly predictive.

Figure 20. Effects of Cognitive, Noncognitive, and Health Endowments on Self-rated Health



Notes: Effects of endowments on fair or poor health outcomes for males (A) and females (B). The endowments and the outcomes are simulated from the estimates of the model in each panel; when the authors compute the effect of each endowment on the outcome, they integrate out the observable characteristics and fix the other two endowments at their overall mean.
Source: Conti and Heckman [2010].

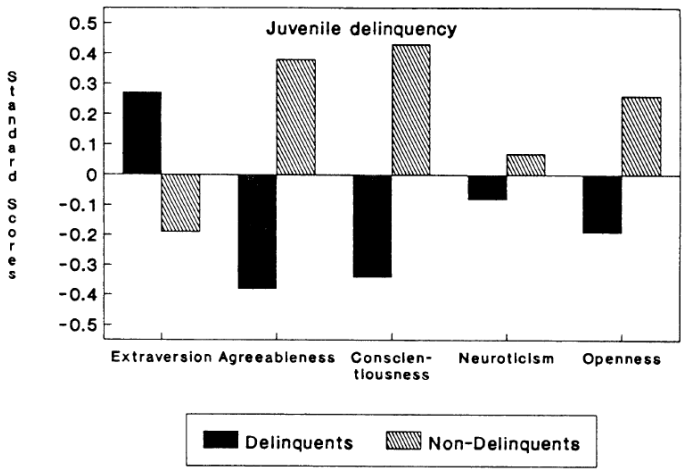
Using similar methods, Savelyev [2010] finds that both child Conscientiousness and higher education increase survival through age 80, but these traits serve as substitutes for each other so that effects of education are only strong at low levels of Conscientiousness.

In sum, Conscientiousness seems to be the most important Big Five trait in predicting health outcomes. Personality likely affects health through behaviors such as smoking, eating, and exercising. Studies that model the dynamic evolution of health over the life cycle find that personality affects health outcomes as much as cognitive measures or more so in some cases.

7.D. Crime[172]

Few studies have examined the Big Five and criminal behavior. The available evidence suggests that Big Five Conscientiousness and Agreeableness are important protective factors against criminal activity. Figure 21 illustrates that in a sample of at-risk youth, boys who had committed severe delinquent behaviors were more than ¾ of a standard deviation lower in Agreeableness and Conscientiousness – as measured by mother's reports at age 12 or 13 – than boys who had committed minor or no delinquent behaviors up to that age (John, Caspi, Robins et al. [1994]).

Figure 21. Juvenile delinquency and the Big Five



Notes: Delinquents are those who have committed at least one of the following: breaking and entering, strong-arming, or selling drugs. Non-delinquents have committed at most one of the following stealing at home, vandalism at home, or theft of something less than $5. The y-axis reports mean differences in standardized scores of the Big Five measures based on mother's reports. The measures were taken at ages 12-13 and reflect cumulative delinquent behavior.
Source: John, Caspi, Robins et al. [1994].

Much of the literature in criminology focuses on the effects of self-control on crime. People with low self-control are "impulsive, insensitive, physical (as opposed to mental), risk-taking, short-sighted, and non-verbal" (Gottfredson and Hirschi [1990], p. 90), and analytic

[172] This section summarizes the more comprehensive survey of the literature prepared by Amanda Agan. See Web Appendix Section A7.B for her survey.

measures of self-control are associated with Big Five Conscientiousness (O'Gorman and Baxter [2002]). Several studies have confirmed that self-control is associated with criminal activity. In an international sample, controlling for basic demographics, a measure of self-control explained between 10% and 16% of the variance in contemporaneously measured theft, assault, drug use, and vandalism (Vazsonyi, Pickering, Junger et al. [2001]). Self-control relates to controlling impulsive behavior so it is not surprising that sensation-seeking and impulsivity are also positively associated with crime. In a sample of college students, partial correlations between a crime factor[173] and sensation-seeking and impulsive behavior were of 0.27 and 0.13 respectively, when controlling for peer behavior and measures of risk appraisal (Horvath and Zuckerman [1993]). Self-control might not paint the entire picture. Negative emotionality – a tendency towards depression likely related Neuroticism – is associated with contemporaneously measured delinquency. Raw correlation coefficients range from $r = 0.13$ for whites $(p < 0.05)$ and $r = 0.20$ for black $(p < 0.05)$ in one sample (Caspi, Moffit, Silva et al. [1994]) to $r = 0.22$ $(p < 0.01)$ in another sample (Agnew, Brezina, Wright et al. [2002]). None of these studies control for cognitive ability nor do they address causality.

However, an emerging literature investigates the effect of education on crime, providing some evidence for causal change. Heckman, Stixrud and Urzua [2006] find that both cognitive traits and noncognitive traits, as captured by locus of control and self esteem, are equally predictive of criminal activity. Their measure of prediction is the effect of decile improvements of cognition and personality traits on the probability of being in jail. Using changes in compulsory schooling laws as an instrument, Lochner and Moretti [2004] and Machin, Marie

[173] The crime factor is based on arrest for selling or buying drugs, shoplifting, driving while drunk, perjury, forging checks, and vandalizing.

and Vujić [2010] find that years of education are negatively associated with criminal activities in the US and UK. Cunha, Heckman and Schennach [2010] show that personality traits are relatively more important in predicting criminal activity than are cognitive traits.

8. Stability and Change in Personality Traits and Preferences

Despite its predictive power, personality would have little policy relevance if it is strictly genetically determined. In this section, we review evidence that shows that personality and IQ change over the life cycle. We explore three channels through which personality can change. First, we discuss the contribution of genetics and show how aspects of personality, such as sensation-seeking, evolve as the brain develops. Second, we show how personality can change through external forces that operate through biology, such as brain lesions and chemical interventions. Third, and most relevant for policy, we show that education, interventions, and parental investment can affect personality. Finally, we discuss the less abundant evidence on the malleability of preferences. In each case, we critically discuss causality.

8.A. Broad Evidence on Changes in Traits over the Life Cycle

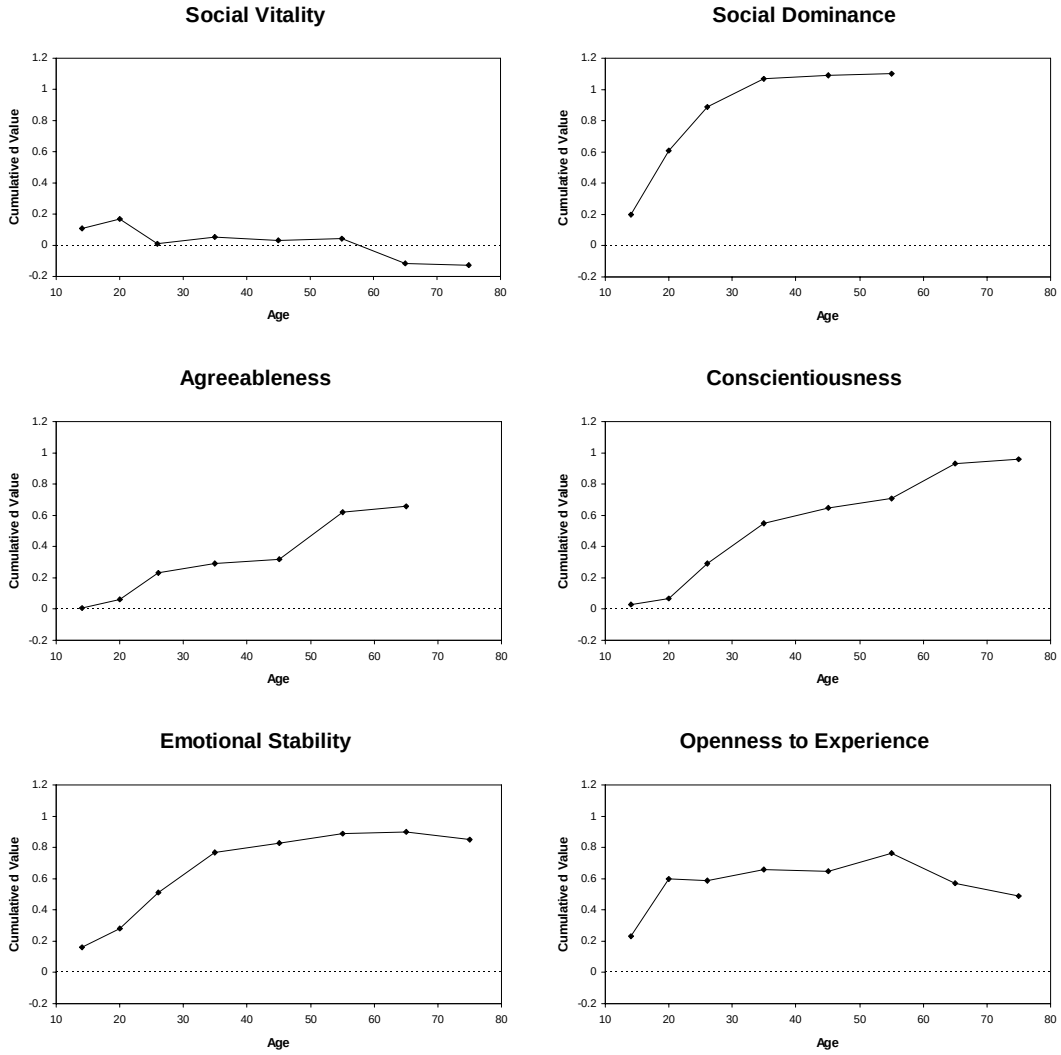
The malleability of personality can be defined and measured in several ways: *Mean-level change* refers to change over time in absolute levels of a trait and is measured by changes in scores over time. *Rank-order change*, in contrast, refers to changes in the ordinal ranking of a trait in a population and is measured by test-retest rank correlations. One commonly held view is that rank-order or mean-level change in personality is nearly impossible after early adulthood. The speculation of James [1890] that "in most of us, by the age of thirty, the character has set like plaster, and will never soften again" (pp. 125-126) is widely touted (see Costa and McCrae [1994], McCrae and Costa [1990; [1994; [1996; [2003], Costa, McCrae and Siegler [1999]). However, mounting evidence suggests that the personality-as-plaster view is not correct (Roberts, Walton and Viechtbauer [2006], Roberts and Mroczek [2008]).

During the early years of life, mean-level changes in traits are obvious and dramatic. For example, children become much more capable of self-control as they move from infancy into

toddler and preschool years (McCabe, Cunnington and Brooks-Gunn [2004], Mischel and Metzner [1962], Posner and Rothbart [2000], Vaughn, Kopp and Krakow [1984]). But mean-level changes in personality are also apparent, albeit less extreme, later in life. In a 2006 meta-analysis of longitudinal studies, Roberts, Walton and Viechtbauer [2006] examine cumulative lifetime change in Big Five Openness to Experience, Conscientiousness, Extraversion, and Agreeableness. They disaggregate Big Five "Extraversion" into social dominance (assertiveness, dominance) and social vitality (talkativeness, gregariousness, and sociability). Figure 22 shows that people typically become more socially dominant, conscientious, and emotionally stable (opposite of neurotic) across the life cycle, whereas social vitality and Openness to Experience rise early in life and then fall in old age.[174] Surprisingly, after childhood, the greatest mean-level change in most personality traits takes place not during adolescence, but rather in young adulthood. An exception is the trait of sensation seeking, discussed earlier as a trait related to risk aversion. During adolescence sensation seeking reaches a dramatic peak that has been associated with synchronous changes in brain structure and function (Spear [2000a; b]).

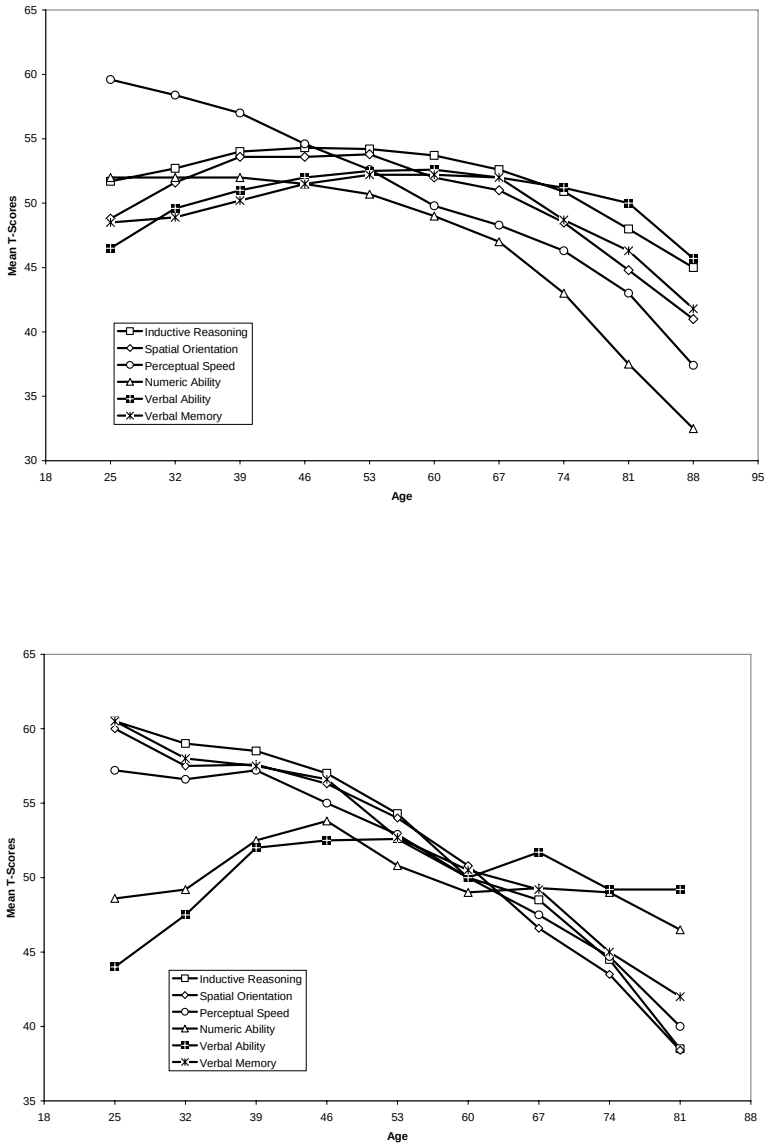
[174] Figure A3 in Section A8 of the Web Appendix presents results for a variety of cognitive, personality and preference parameters from a cross-sectional study based on the GSOEP data. Samples are small and standard errors are large. Many preference parameters show a surprising stability over the life cycle.

Figure 22. Cumulative mean-level changes in personality across the life cycle



Note: Social vitality and social dominance are aspects of Big Five Extraversion. Cumulative *d* values represent total lifetime change in standard deviations.
Source: Figure taken from Roberts, Walton and Viechtbauer [2006] and Roberts and Mroczek [2008]. Reprinted with permission of the authors.

Figure 23. Longitudinal analysis (top panel) and cross-sectional analysis (bottom panel) of mean-level change in cognitive skills over the lifespan

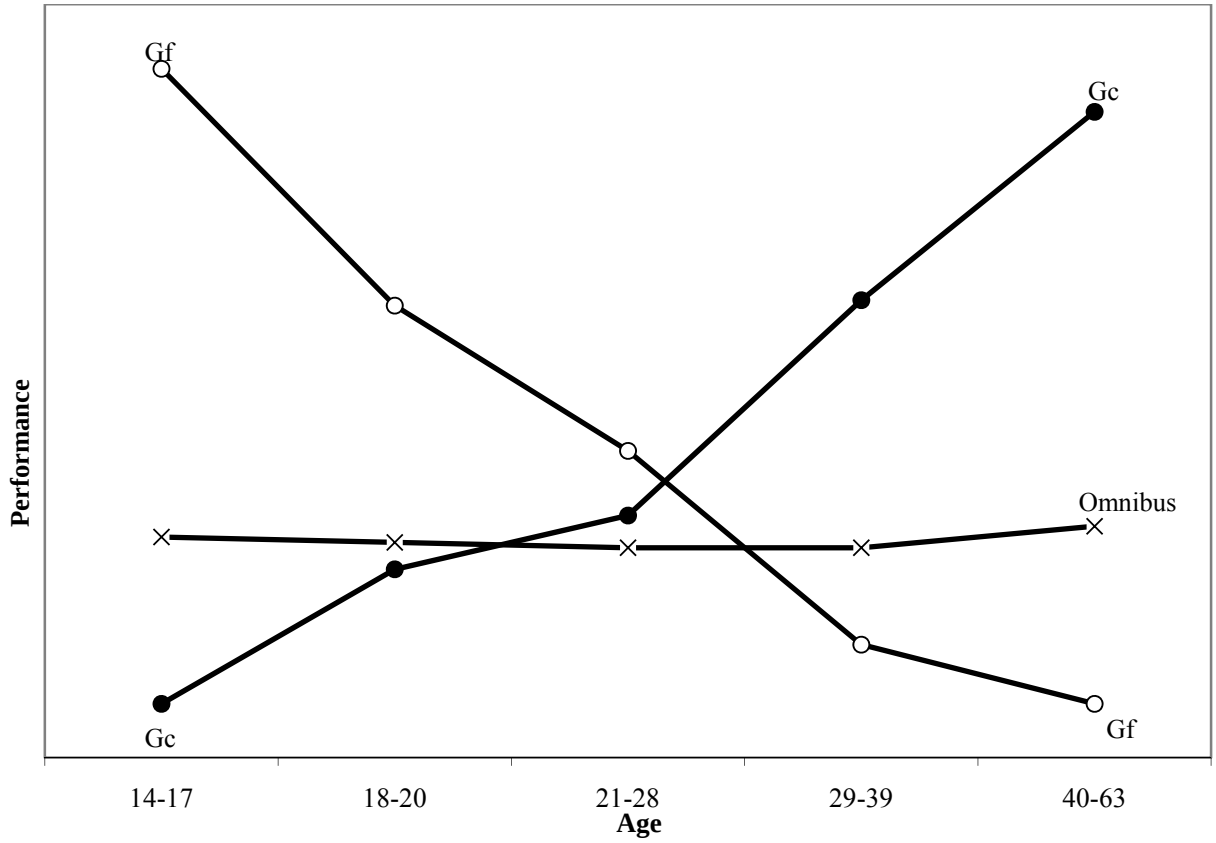


Notes: T-scores on the y-axis are standardized scores with a mean of 50 and a standard deviation of ten.
Source: Figures taken from Schaie [1994]. Used with permission of the publisher.

In contrast, a longitudinal study of adult intellectual development shows mean-level declines in cognitive skills, particularly cognitive processing speed, after age 55 or so (Schaie [1994]). The top panel of Figure 23 shows mean-level changes in cognitive skills using a longitudinal analysis, and the bottom panel of Figure 23 shows mean-level changes using a

cross-sectional analysis.[175] As schematically illustrated in Figure 24, fluid intelligence decreases and crystallized intelligence rises over the life cycle (Horn [1970]). Accumulated skills and knowledge are important: most of us would rather use an experienced cardiac surgeon who has seen hundreds of cases just like ours to perform our surgery, rather than an exceptionally bright young surgeon with minimal experience.

Figure 24. Fluid intelligence decreases and crystallized intelligence increases across the lifespan

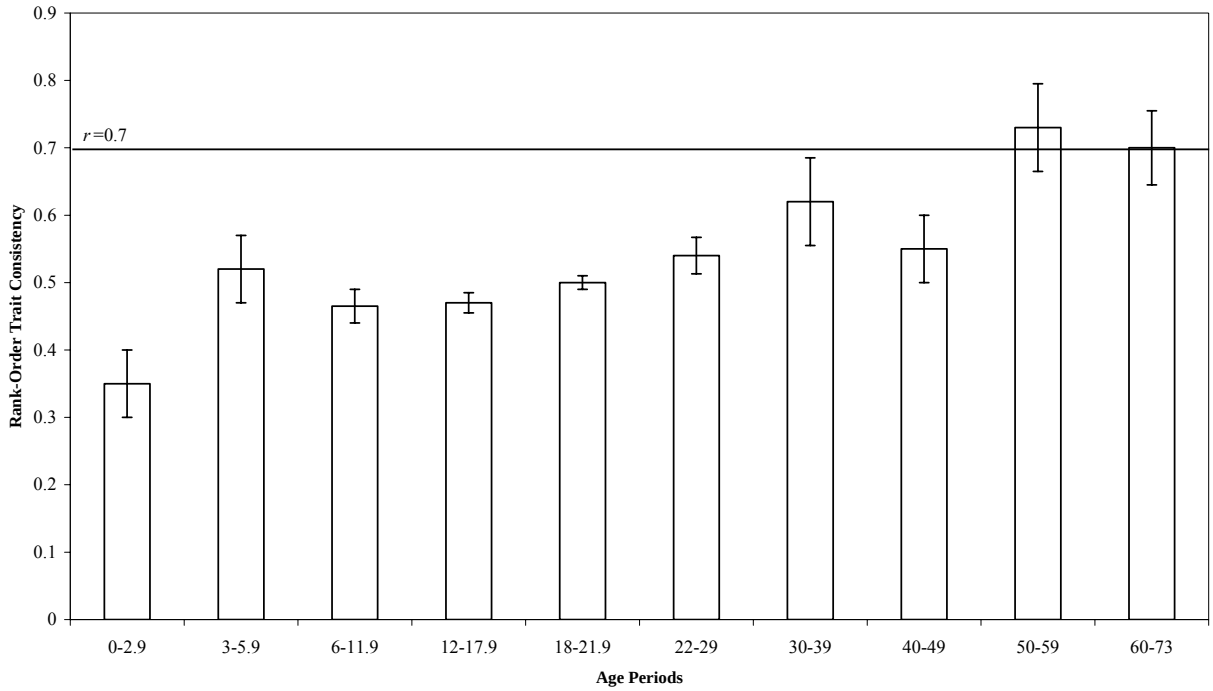


Source: Figure from Horn [1970]. Used with permission of Elsevier.

[175] Cross-sectional estimates of mean-level change are biased by cohort effects (for example, the Flynn effect) whereas longitudinal estimates are biased by test-retest learning (when the same IQ tests are administered repeatedly to the same subjects) and by selective attrition. Thus, both estimates must be considered in conjunction as evidence for mean-level change.

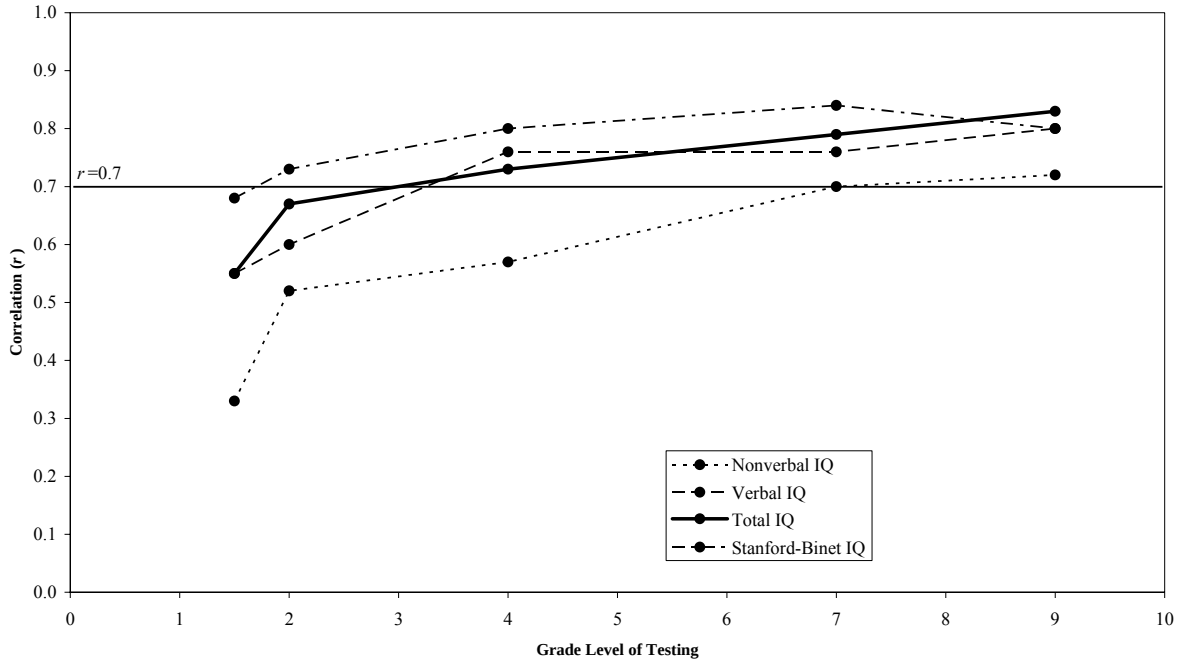
Rank-order stability in personality increases steadily over the lifespan. Figure 25 shows that seven-year test-retest stability estimates for personality plateau far from unity, at $r = .74$, about the same level as terminal stability estimates for IQ Roberts and DelVecchio [2000]. However, personality does not reach this plateau until at least age 50; whereas IQ reaches this plateau by age six or eight (Hopkins and Bracht [1975], Schuerger and Witt [1989]). Figure 26 shows rank order stability of IQ over broad age ranges.

Figure 25. Rank-order stability of personality over the life cycle



Notes: The meta-analysis and reflects test-retest correlations over, on average, 6.7-year periods.
Source: Figure taken from Roberts and DelVecchio [2000]. Reprinted with permission of the authors.

Figure 26. Rank-order stability of IQ across the life span



Notes: The points represent ten-year, test-retest correlations over ten-year intervals. Grade level, not age, is on the x-axis.
Source: Figure reproduced from Hopkins and Bracht [1975]. Used with permission of the publisher.

8.B. Evidence on Genetic Determination of Traits and Ontogenic Change

A second useful dichotomy contrasts *normative change*, defined as changes that are typical of the average individual in a given population and caused either by biological programming (ontogenic) or by predictable changes in social roles (sociogenic), and *non-normative change*, encompassing both intentional change, caused by deliberate, self-directed efforts, deliberately chosen changes in social roles and atypical life events (trauma, for example).[176] More relevant to the current review is intentional change – can investments be made by the person or the parent that change ability or preferences?

[176] Normative here refers to what most people or the average person experiences. If most people deliberately do something that causes change, it would be normative. But that seems unlikely. Therefore, most deliberative change is non-normative, but logically this is not necessarily true.

What mechanisms underlie stability and change in personality? Why do most of us mature into more responsible, agreeable, confident, and emotionally stable individuals as we age? Why do our personality traits, relative to others in our age-group, steadily stabilize as we approach our fifties? If, as McCrae and colleagues have claimed, normative changes reflect genetically programmed processes, investment should not affect change. The current literature in psychology claims that genetic factors are largely responsible for *stability* in personality in adulthood whereas environmental factors are mostly responsible for *change* (Blonigen, Hicks, Krueger et al. [2006]; Plomin and Nesselroade [1990]).[177] In a longitudinal study of twins surveyed at age 20 and then again at age 30, about 80 percent of the variance of the stable component of personality was attributed to genetic factors (McGue, Bacon and Lykken [1993]). In the same study, change in personality was mostly attributed to environmental factors. Helson, Kwan, John et al. [2002], for example, document the substantial influence that social roles and cultural milieu can have on personality development. This conclusion is consistent with an economic model of investment and the response of measured traits to incentives. However, recent evidence suggests that environmental factors, and in particular stable social roles, also contribute to stability in personality and that genetic factors can contribute to change (see Roberts, Wood and Caspi [2008] for a review).

Research on IQ also points to the enduring effects of genes, which are with us all of our lives, in contrast to more transient effects of environmental influences, which depend on a multitude of unstable variables, including social roles, levels of physical maturity and decline,

[177] Plomin and the essays in the December issue of *Monographs for the Society for Research in Child Development* (Overton [2007]) extend this analysis to childhood.

and historical and cultural milieu.[178] Increases in the heritability of IQ from childhood (about 40 percent) to adulthood (estimates range from 60 percent to 80 percent) are well-documented in studies of behavioral genetics and possibly reflect increasing control of the individual (vs. parents) over environment (Bergen, Gardner and Kendler [2007]; McGue, Bouchard, Iacono et al. [1993]; Plomin, DeFries, Craig et al. [2002]).[179] Heritability estimates for Big Five traits are relatively stable across the life cycle at about 40 to 60 percent (Bouchard and Loehlin [2001]).[180] Behavioral genetics studies typically estimate the effect of parental environments to be near zero, but Turkheimer, Haley, Waldron et al. [2003] find estimates from such studies to be biased downward by the over-representation of middle- and upper-class families. Among poor families, Turkheimer et al. find that 60 percent of the variance in IQ is accounted for by shared environment and heritability estimates are almost nil, whereas among affluent families, the result is reversed. Krueger and colleagues have recently demonstrated that other moderators may influence the heritability of traits (see Krueger, South, Johnson et al. [2008]).

Genes exert their influence in part through the selection and evocation of environments that are compatible with one's genotype—a phenomenon sometimes referred to as "gene-environment correlation" or "nature via nurture" (see Rutter [2006b]). As individuals move from childhood to adulthood, they have more control over their environments, and thus gene-environment correlation becomes more important because shared environments become less

[178] We note here that while genes remain constant through the life cycle, the expression of genes is determined, in part, by experience.
[179] Devlin, Daniels and Roeder [1997] suggest that traditional estimates of the heritability of IQ may be inflated because they fail to take into account the effect of the environment of the maternal womb. See also Rutter [2006b] and an emerging literature on epigenetics.
[180] Lykken [2007] suggests that heritability estimates for personality are substantially higher when situational influence and measurement error are minimized by taking multiple measures at least a few months apart.

common.[181]

Substantial but temporary influence of environment is a basic assumption of the Dickens-Flynn model reconciling the high heritability of IQ and massive gains of IQ between generations (Dickens and Flynn [2001]).[182] The relatively short half-life of common environmental influences may also explain why adopted children resemble their biological parents more and more and their adopted parents less and less as they grow older (Scarr, Weinberg and Waldman [1993]).[183] Lizzeri and Siniscalchi [2008] develop an economic model of parenting that explains the evidence on twins raised apart that shows conditions under which parents will differentially treat their children.

It is important to note that the family studies of genetic influence measure only the effects of *shared* environments, which become less similar as children age. Thus even identical twins may be motivated to seek out different environments over time (Rutter [2006b]). Recent evidence that first born children grow up, on average, to have three points higher IQ than their younger siblings reinforces the point that parents do not necessarily provide identical environments in childhood (Kristensen and Bjerkedal [2007]). Conti, Heckman, Yi et al. [2010] demonstrate how parents differentially respond to health shocks of individual identical twins.

---

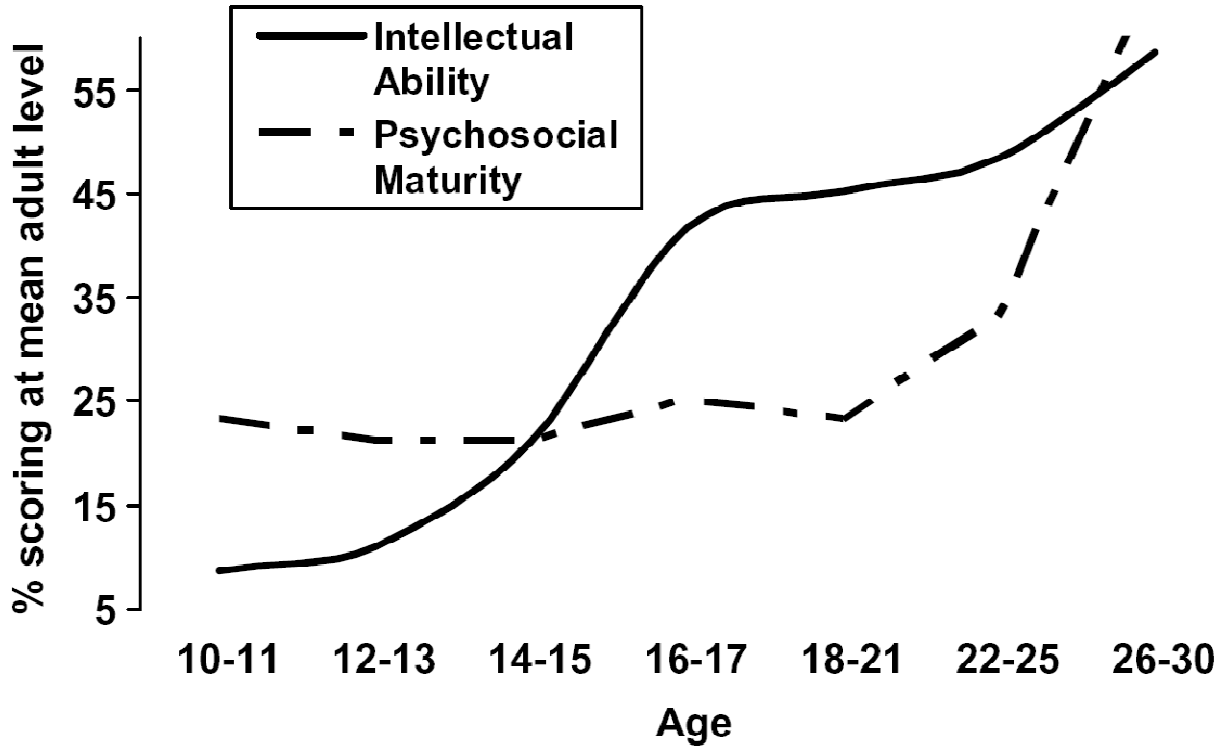
[181] Gene-environment interactions are another means by which genes and environment jointly influence traits. The effects of the environment depend on the genes and vice versa (see Caspi, Sugden, Moffitt et al. [2003]; Moffitt, Caspi and Rutter [2005]; and Caspi, McClay, Moffitt et al. [2002]).

[182] A second crucial assumption is that environmental influence can be amplified by a "social multiplier" effect: smarter individuals create for one another an enriched environment, which in turn increases intelligence, and so on. Some caution must be taken in relying on the claims in this literature. Blair, Gamson, Thorne et al. [2005] attribute the Flynn effect to increasing access to formal schooling early in the twentieth century and, from the mid-century onward, to increasing fluid cognitive demand of mathematics curricula. Flynn [2007] concurs about the former but believes that the latter had negligible impact.

[183] The literature establishes that shared environments become less important as children age. This literature does *not* say that environments do not matter. This effect can arise because genetically similar children (or their parents) choose different environments to distinguish themselves or because of parental investment (Lizzeri and Siniscalchi [2008]).

As mentioned earlier, genes could affect not only the base level of personality but also how personality changes over the life cycle. Just as people grow taller throughout childhood, people's personalities might naturally develop, even without investment. Steinberg [2008] provides one such example, arguing that typical biological (ontogenic) development explains the surge of risk-taking in adolescence followed by the decline in adulthood. Figure 27 illustrates how basic intellectual ability and psychosocial maturity (e.g. impulsivity, risk perception, sensation-seeking, future orientation) evolve over the life cycle. Intellectual ability peaks at around age 16, whereas psychosocial maturity develops during late adulthood. In his model, the increase in adolescent risk taking due to a restructuring of the brain's dopaminergic system (responsible for the brain's reward processing) in such a way that immediate or novel experiences yield higher rewards, especially in the presence of peers. He attributes declines in risk-taking due to development of the brain's cognitive control system, specifically improvements in the prefrontal cortex which promote aspects of executive function such as response inhibition, planning ahead, weighing risks and rewards, and the simultaneous consideration of multiple information sources. Interestingly, even in his model, sensation-seeking partly depends on the presence of peers, which corresponds to aspects of the situation (S_h) in the framework of Section 3. This example highlights the difficulty in disentangling situational and biological changes in personality.

Figure 27. Proportion of individuals in each age group scoring at or above the mean for 26- to 30-year-olds on indices of intellectual and psychosocial maturity.



Source: From Steinberg, Graham, O'Brien et al. [2009] submitted for publication.

What other than preprogrammed genetic influences might account for mean-level changes in personality? Personality change in adulthood may be precipitated by major shifts in social roles (for example, getting a job for the first time, becoming a parent). If social role changes are experienced by most people in a population at the same time, we will observe the effects as mean-level changes in personality. If, on the other hand, these social roles are not assumed synchronously, we will observe rank-order changes. One difficulty with many of these studies is the problem of reverse causality discussed in Section 3 and addressed in Hansen, Heckman and Mullen [2004] and Heckman, Stixrud and Urzua [2006]. Changes in personality may drive social role changes rather than the other way around.

8.C. External Changes to Biology

The previous subsection provides examples of how difficult it is to disentangle biological changes in personality from environmental or situational effects. In this subsection we provide some clearer evidence for true causal changes in personality due to external forces that either damage parts of the brain or abruptly alter the chemistry of the brain.

Brain Lesion Studies

Brain lesion studies provide the most dramatic and convincing evidence that personality can change. The most famous example is Phineas Gage, a construction foreman whose head was impaled by a metal spike and who subsequently changed from being polite and dependable to rude and unreliable (Damasio, Grabowski, Frank et al. [2005]). Since then, there have been many more case studies of patients with brain damage. For example, Mataro, Jurado, Garcia-Sanchez et al. [2001] describe the behavior of a Spanish patient whose head was impaled by an iron spike, injuring both frontal lobes. Like Phineas Gage, his behavior changed. After the accident, he had difficulty planning, became more irritable, and had problems regulating emotions. Unlike Phineas, he was cheerful and did not display anti-social behavior, suggesting the personality is malleable in different dimensions, even through brain damage. Furthermore, the effects of brain damage are persistent. After five years, patients who suffered traumatic head injuries have social impairments, such as anger control, even when their performance on cognitive tasks returns to the normal range (Lezak [1987]).

Using more advanced methods, neuroscientists have delved deeper into the inner workings of the brain. Some recent studies have investigated how two parts of the brain, the amygdala and ventromedial prefrontal cortex (VMPC), affect personality by regulating emotion. Bechara [2005] discusses how emotion is one way that people might assign and store value to

particular outcomes in a way that is useful for decision-making. The amygdala is thought to signal "impulsive" emotional responses to immediate environmental stimuli, such as reacting quickly to a snake. In contrast, the VMPC is thought to signal "reflective" emotional responses to memories and knowledge. These two parts of the brain conflict with each other when people make decisions: signals from the amygdala induce behavior that implicitly values immediate outcomes, whereas signals from the VMPC reflect long-run considerations. The stronger signal dictates the resultant behavior. People with damage to these parts of the brain, exhibit changes in personality. For example, people with damage to the VMPC, the part that regulates "reflective" emotion, tend to act impulsively and seem to overvalue short-term outcomes in a way that leads to long-term financial loss and loss of friendships, despite having relatively normal levels of intellectual capacity. McClure, Laibson, Loewenstein, and Cohen's $\beta - \delta$ system (McClure, Laibson, Loewenstein et al. [2004]) claims to establish a link. However, the recent research in neuroscience challenges this claim (Monterosso and Luo [2010]).

Further experiments involving these parts of the brain highlight why separating cognitive and noncognitive traits might be a red herring. For example, Bechara and Damasio [2005] compare patients who have lesions in the ventromedial prefrontal cortex (VMPC) or amygdala to people with no lesions. The participants were given the Iowa Gambling Task, in which they repeatedly chose between four decks of cards that represented lotteries of different value. Throughout the experiment, the authors also measured skin conductance responses (SCRs), a known physiological reflection of emotion. People without lesions learned to choose the "better" decks of cards with lower short-term payoffs but higher average payoffs. They also showed emotional activity both when picking their card and when receiving the rewards or penalties. Amazingly, they began picking the better decks of cards even before they were consciously

aware of the difference between the decks; their emotions guided their behavior. In contrast, people with lesions never learned to pick the better decks, seemingly because they could not develop emotional responses. Patients with damage to the amygdala never showed emotional response to rewards or penalties, suggesting they never learned to value the outcomes at all. Patients with damage to the VMPC showed emotional response only when receiving the reward or penalty but not when selecting decks, suggesting that they might not have reflective emotional responses crucial in considering long-term consequences. Interestingly, in a follow-up experiment, the VMPC patients continued to choose from the bad decks even after forming knowledge of the consequences.

These results might also apply to addiction and highlight the importance of the "situation" discussed in the framework of Section 3. Drug addicts have abnormalities in the VMPC[184] or perform similarly to VMPC patients on the Iowa Gambling Task. Environmental cues, such as seeing alcohol, might trigger an impulsive response from the amygdala that is not checked by the VMPC (Bechara [2005]). Numerous other studies show the role of the amygdala in signaling emotions and its relationship to cognition and behavior (Phelps [2006]).

Chemical and Laboratory Interventions

A few recent studies show that it is possible to alter preferences and personality through experiments that change the brain's chemistry. For example, magnetic disruption of the left lateral prefrontal cortex can increase experimentally elicited discount rates (Figner, Knoch, Johnson et al. [2010]). Similarly, nasal sprays of oxytocin increase trust (distinct from altruism or ability to assess probabilities) in a game-theoretic experiment (Kosfeld, Heinrichs, Zak et al. [2005]). As discussed in Section 5, the Big Five traits are linked to personality disorders.

[184] Though, it is not known whether the abnormalities precede addiction.

Therefore, it is not surprising that administering paroxetine, a drug for treating depression, decreases Neuroticism and increases Extraversion. More surprising is that the drug affects personality above and beyond its direct effects on depression. Furthermore, patients who become less neurotic are also less likely to relapse even after treatment, suggesting that the paroxetine might have a long-lasting impact through a bio-chemical change in the brain (Tang, DeRubeis, Hollon et al. [2009]). Similarly, Knutson, Wolkowitz, Cole et al. [1998] find evidence that paroxetine can diminish hostile behavior through a decrease in general negative effect.

8.D. A Theoretical and Empirical Framework for Analyzing Changes through Education, Interventions, and Investment

Even though brain lesion studies and laboratory experiments provide convincing causal evidence that personality can be changed, they are not viable mechanisms for large scale policy interventions. However, a growing body of evidence suggests that education, parental investment, and interventions can affect personality. In this section, we elaborate the theoretical and empirical framework introduced in Section 3, highlighting the dynamics of capability development. In the next section, we discuss the empirical evidence through the lens of the framework presented in this section.

In order to match the empirical literature more closely, we will work with a particular case of the general framework introduced in Section 3. Rather than considering all traits, including preferences and effort endowments, we will simplify the trait vector to contain only traits. Following the notation in Section 3, denote the vector of age τ traits as θ^{τ} where we divide θ^{τ} into "mental" (μ) and "personality" (π) traits, θ_{μ}^{τ} and θ_{π}^{τ}. Solving out for the actions in terms of their determinants, the outcome from task j at age τ is P_{J}^{τ}:

(24) $$P_j^\tau = \phi_j(\theta_\mu^\tau, \theta_\pi^\tau, e_j^\tau)$$

where e_j^τ is effort devoted to task j at time τ. We also adopt a modified version of the effort supply function introduced in Section 3 that depends on rewards and endowments:

(25) $$e_j^\tau = e_j^\tau(R_j^\tau, A^\tau)$$

where R_j^τ is the reward per unit effort in activity j and A^τ represents other determinants of effort which might include some or all of the components of θ^τ.

This vector of traits evolves by the following simplified version of Equation (15):

(26) $$\theta^{\tau+1} = f^\tau(\theta^\tau, IN^\tau, \theta_E^\tau), \tau = 0, \ldots, \mathcal{T}$$

where θ_E^τ represents "parental environment" that could be thought of as part of situation S_{h^τ}. We interpret IN^τ very generally so that it includes parental investment, education, and investments from interventions. θ^0 is the vector of initial endowments determined at birth or at conception.[185]

As discussed in Section 3, both cognitive and personality skills can be affected by parental investment and schooling, components of IN^τ. The returns to investment, however, might depend on the age of investment or the traits. A crucial feature of the technology that helps to explain many findings in the literature on skill formation is *complementarity of traits with investment*:

(27) $$\frac{\partial^2 f^\tau(\theta^\tau, IN^\tau, \theta_E^\tau)}{\partial \theta^\tau \partial IN^{\tau'}} \geq 0.$$

Technology (26) is characterized by *static complementarity* between period τ traits and period τ investment. The higher θ^τ, the higher the productivity of the investment.

[185] Modifications for adult technologies of skill formation are straightforward.

There might also be *dynamic complementarity* if technology determines period $\tau+1$ traits $(\theta^{\tau+1})$. This generates complementarity between investment in period $\tau+1$ and investment in period s, $s>\tau$. Higher investment in period τ raises $\theta^{\tau+1}$ because technology is increasing in IN^{τ}, which in turn raises θ^{s} because the technology is increasing in θ^{τ}, for τ between τ and s. This, in turn, raises $\frac{\partial f^{s}(\cdot)}{\partial IN^{s}}$ because θ^{s} and IN^{s} are complements, as a consequence of (27).

Dynamic complementarity explains the evidence that early nurturing environments affect the ability of animals and humans to learn. (See Knudsen, Heckman, Cameron et al. [2006].) It explains why investments in disadvantaged young children are so productive. Early investments enhance the productivity of later investments. Dynamic complementarity also explains why investment in low ability adults often has such low returns—because the stock of θ^{τ} is low. (See the evidence in Cunha and Heckman [2007], Heckman [2007], Heckman [2008] and in Cunha, Heckman, Lochner et al. [2006].) Using dynamic complementarity, one can define *critical* and *sensitive* periods for investment. If $\frac{\partial f^{\tau}(\cdot)}{\partial IN^{\tau}}=0$ for $\tau\neq\tau^{*}$, τ^{*} is a critical period for that investment. If $\frac{\partial f^{\tau}(\cdot)}{\partial IN^{\tau}}>\frac{\partial f^{\tau'}(\cdot)}{\partial IN^{\tau'}}$ for all $\tau\neq\tau^{*}$, τ is a sensitive period. The technology of skill formation is consistent with a body of evidence on critical and sensitive periods.

Figure 28 shows how adult choices and outcomes are shaped by *sequences* of investments over the life cycle of the child. The importance of the early years depends on how easy it is to reverse adverse early effects with later investment. Resilience and remediation are possible, but are more costly later on. The accumulation of investments over the life cycle of the

child determines adult outcomes and the choices people will make when they become adults. To capture these interactive effects requires nonlinear models.

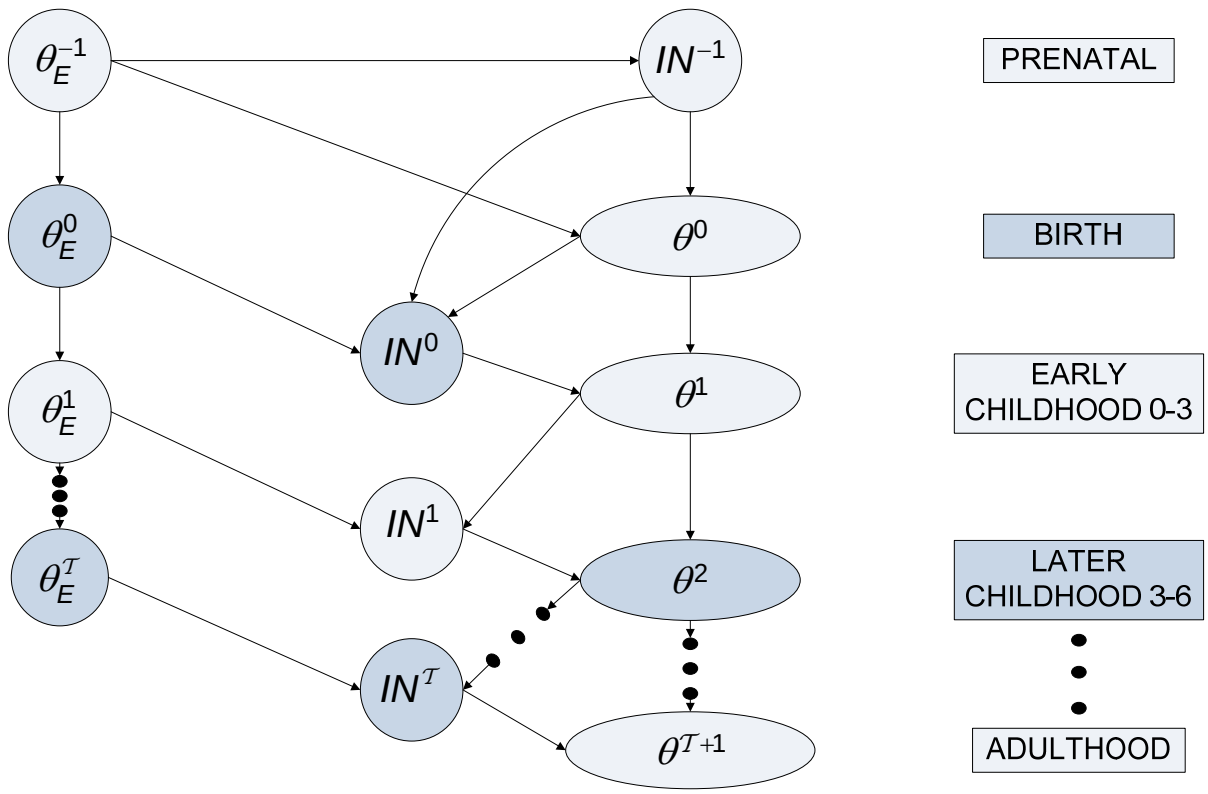
Figure 28. A Life Cycle Framework for Organizing Studies and Integrating Evidence: $\mathcal{T}+1$ Periods of Life Cycle

θ^{τ} capacities at τ

IN^{τ}: investment at τ

θ_E^{τ}: environment at time τ

$$\theta^{\tau+1}=f^{\tau}\left(\theta^{\tau}, IN^{\tau}, \theta_E^{\tau}\right)$$



It is important for policy purposes to know at which stage of the life cycle interventions are the most effective and to move beyond the correlations between early life and later life events—to understand the mechanisms of skill formation. Cunha and Heckman [2008] and

Cunha, Heckman and Schennach [2010] estimate technologies of skill formation to understand how the skills of children evolve in response to the stock of skills children have already accumulated; the investments made by their parents; and the stock of skills accumulated by the parents themselves.

They allow for $\mathcal{L}$ different developmental stages in the life of the child: $\ell \in \{1,\ldots,\mathcal{L}\}$. Developmental stages may be defined over specific ranges of ages, $\tau \in \{1,\ldots,\mathcal{T}\}$, so $\mathcal{L} \leq \mathcal{T}$. They assume that each component of θ^{τ} and IN^{τ} is a scalar as is parental environment θ_E^{τ}. Let IN_k^{τ} be investment in capability k at age τ. The technology for producing capability k at stage ℓ is

$$\theta_k^{\tau+1} = \left[\gamma_{\mu,k}^{\ell}\left(\theta_{\mu}^{\tau}\right)^{\sigma_k^{\ell}} + \gamma_{\pi,k}^{\ell}\left(\theta_{\pi}^{\tau}\right)^{\sigma_k^{\ell}} + \gamma_{IN,k}^{\ell}\left(IN_k^{\tau}\right)^{\sigma_k^{\ell}} + \gamma_{E,k}^{\ell}\left(\theta_E^{\tau}\right)^{\sigma_k^{\ell}}\right]^{\frac{1}{\sigma_k^{\ell}}},$$

$$\gamma_{m,k}^{\ell} \geq 0, \quad \sum_m \gamma_{m,k}^{\ell} = 1 \text{ for all } k \in \{\mu,\pi\} \text{ and } \ell \in \{1,\ldots,L\}.$$

A main finding of Cunha, Heckman and Schennach [2010] is that the elasticity of substitution σ_{μ}^{ℓ} decreases with ℓ. This is consistent with evidence on the declining malleability of IQ with age, i.e., that cognitive deficits are easier to remedy at early ages than at later ages. At the same time, σ_{π}^{ℓ} stays roughly constant over ℓ. This is consistent with evidence on the emergence of psychological maturity, as shown in Figure 27.

Cunha and Heckman [2008] estimate the model

(28) $$\theta^{\tau+1} = A^{\tau}\theta^{\tau} + B^{\tau}IN^{\tau} + \eta^{\tau},$$

where η^{τ} is an unobserved shock. The main problem that arises in estimating the technology is that vector $(\theta^{\tau}, IN^{\tau})$ is not directly observed, i.e. it is observed with error. Cunha and Heckman

[2008] treat $(\theta^{\tau}, IN^{\tau})$ as a vector of unobserved factors and use a variety of measurements of the latent constructs to proxy these factors.

There is a large body of econometric work on linear factor models (see, e.g., Aigner, Hsiao, Kapteyn et al. [1984]). There is parallel work in other fields ("structural equation modeling," see Bollen [1989], or "state-space Kalman filter models"). These models account for measurement errors in the proxies. Cunha and Heckman [2008] find these errors to be quantitatively large. If they are not accounted for, estimates of technology parameters are substantially biased.

In a linear setting, it is assumed that multiple measurements on inputs and outputs can be represented by a linear factor setup:

(29) $$M_{k,j}^{\tau} = \mu_{k,j}^{\tau} + \alpha_{k,l}^{\tau}\theta_{k}^{\tau} + \varepsilon_{k,j}^{\tau}, \quad \text{for } j \in \{1,\ldots,M_{k}^{\tau}\}, k \in \{\pi, \mu, IN\},$$

where M_{k}^{τ} is the number of measurements on latent factor k, and θ_{IN}^{τ} is latent investment at age τ. They anchor the scales of θ^{τ} using outcome equations. Test scores are intrinsically meaningless. Cunha and Heckman anchor test scores in labor market and educational outcomes. This approach generalizes to a nonlinear-nonparametric framework both for the measurement equations and the state-space equations (Cunha, Heckman and Schennach [2010]).

Low dimensional $(\theta^{\tau}, IN^{\tau})$ (abilities and investment) are proxied by numerous measurements for each component, which accords with standard economic models where time preference, risk aversion, ambiguity aversion and leisure preferences are low-dimensional unobserved factors affecting numerous outcomes.

8.E. The Evidence on the Causal Effects of Parental Investment, Education, and Interventions

Evidence of Change in Traits from Studies of Parental Investment

Cunha, Heckman, Lochner et al. [2006] summarize a large literature on child development. Evidence from a substantial literature suggests that the enduring effects of environment are greater earlier in life for intelligence. Duyme, Dumaret and Tomkiewicz [1999] studied children with IQs below 86 who were adopted between the ages of four and six into stable homes. In adolescence, children adopted into high-SES homes gained an average of 19.5 IQ points; children adopted into low-SES homes showed an average gain of 7.7 IQ points. In studies of Romanian children taken from impoverished orphanages and placed into middle-class British homes, the long-term salutary effects of adoption on cognitive ability were dramatic when infants were placed before they reached six months, and markedly less so when adoption was delayed until later ages (Beckett, Maughan, Rutter et al. [2006]). Notably, children adopted at different ages between six to 42 months did not differ at age 11 from each other in the terms of cognitive ability, with all children demonstrating an average deficit of 15 IQ points relative to children who had been adopted earlier in life. The effects of low nutrition had no effect on cognitive outcomes at age 11, suggesting a prominent role for psychological deprivation. As Beckett and colleagues point out, these findings are consistent with the existence of a very early critical or sensitive period for intellectual development in which particular environmental stimuli are necessary for normative axonal rewiring (see Uylings [2006] and Rutter [2006b] for reviews).[186]

[186] However, the data are also consistent with alternative explanations such as extreme stress permanently damaging brain structures.

Using the framework introduced in the previous subsection, Cunha, Heckman and Schennach [2010] estimate a dynamic model of skill formation with a two-stage model of childhood $(L = 2)$. Stage 1 is birth through age 4. Stage 2 corresponds to age 5 through 14. In contrast to linear models that assume perfect substitutability among inputs in the scale in which investment is measured, Cunha, Heckman and Schennach [2010] estimate nonlinear technologies to identify key substitution parameters. These parameters address the question of whether or not it is possible to remediate from early disadvantage and the cost of waiting.

Their findings shed light on the dynamic process of capability formation in a way that raw correlations across time do not. They find that self-productivity becomes stronger as children become older, for both cognitive and noncognitive capability formation. The elasticity of substitution for cognitive inputs is *smaller* in second stage production, so that it is more difficult to compensate for the effects of adverse environments on cognitive endowments at later ages than it is at earlier ages. This finding helps to explain the evidence on ineffective cognitive remediation strategies for disadvantaged adolescents documented in Cunha, Heckman, Lochner et al. [2006], Knudsen, Heckman, Cameron et al. [2006] and Cunha and Heckman [2007]. Personality traits foster the development of cognition but not vice versa. It is equally easy to substitute at both stages for socioemotional skills over the life cycle. Overall, 16% of the variation in educational attainment is explained by adolescent cognitive traits, 12% is due to adolescent personality (socioemotional traits), and 15% is due to measured parental investments.

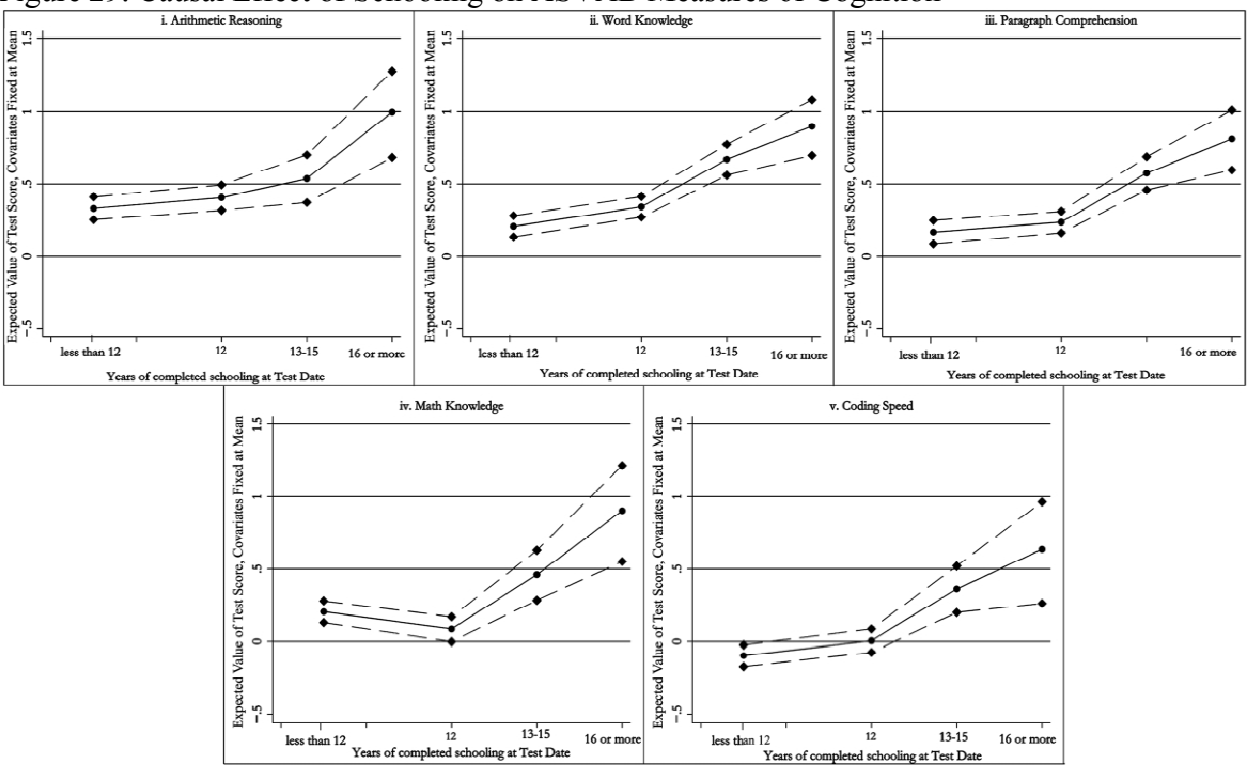
The Effects of Schooling on Cognitive and Personality Traits

Despite a large literature on the effects of schooling on shaping preferences (see Bowles and Gintis [1976] and the literature it spawned), there is surprisingly little direct evidence on this

question. An exception is the analysis of Heckman, Stixrud and Urzua [2006]. The authors formulate and estimate an economic model that identifies the effect of cognitive and personality traits on schooling and a variety of other outcomes. The model controls for the effect of schooling in boosting both cognitive and personality measures and thus controls for reverse causality. They estimate their model on the National Longitudinal Survey of Youth 1979, which has measures on the components of the Armed Services Vocational Battery (ASVAB) that are used to create the Armed Forces Qualifying Test (AFQT) that is widely used as a measure of cognition. In addition, the NLSY79 has two measures of personality. The Rotter Locus of Control Scale, discussed in Section 5, is designed to capture the extent to which individuals believe that they have control over their lives through self-motivation or self-determination as opposed to the extent that the environment controls their lives (Rotter [1966]). The NLSY79 data also contain the Rosenberg Self-Esteem Scale, which attempts to assess the degree of approval or disapproval of one's self (Rosenberg [1965]). The relationship between these measures and the Big Five traits of Neuroticism was discussed in Section 5.

As suggested in the economic framework introduced in the previous subsection, different abilities might be more responsive to investment at different ages. Figure 29 shows the causal effects of years of schooling attained on five components of ASVAB. Schooling in the high school years has moderate but positive effects on the measures of cognition, consistent with previous research by Hansen, Heckman and Mullen [2004], Neal and Johnson [1996], and Winship and Korenman [1997]. The most dramatic causal effects on cognition arise from college attendance. In contrast, locus of control is primarily affected by high school attendance but not college attendance. On measures of self-esteem, an additional year of high school and college play powerful roles.

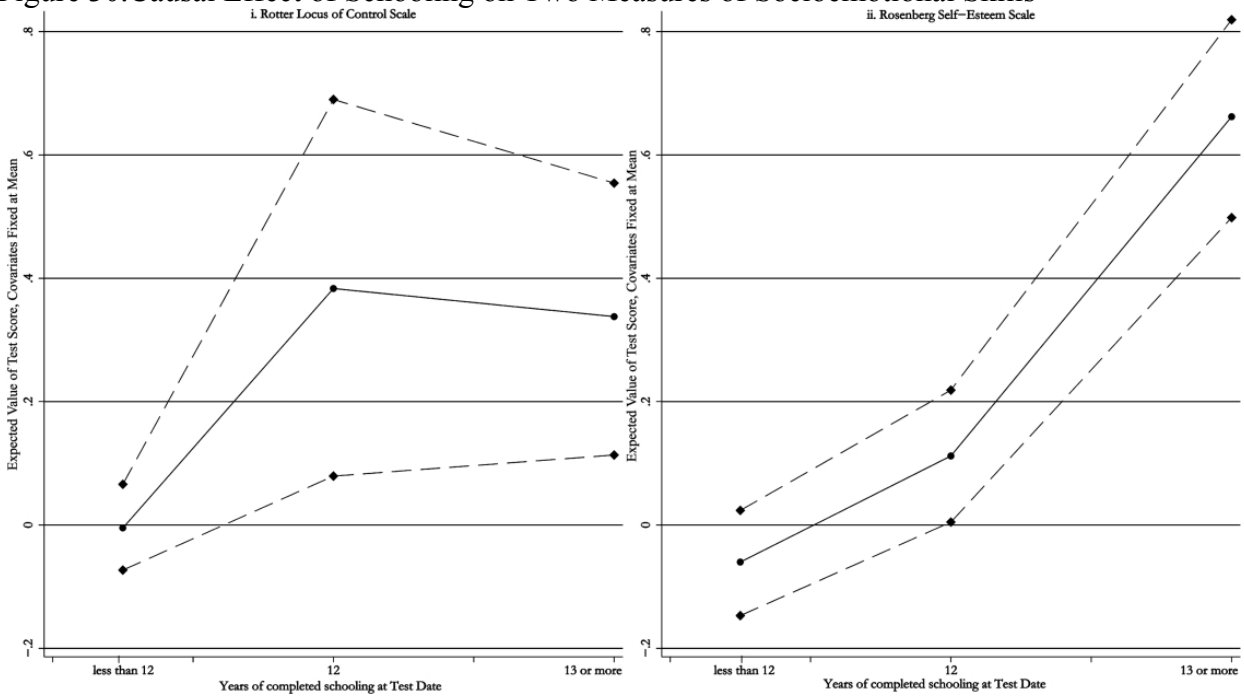
Figure 29. Causal Effect of Schooling on ASVAB Measures of Cognition



Notes: Effect of schooling on components of the ASVAB. The first four components are averaged to create male's with average ability. We standardize the test scores to have within-sample mean zero, variance one. The model is estimated using the NLSY79 sample. Solid lines depict average test scores, and dashed lines, 2.5%–97.5% confidence intervals.

Source: Heckman, Stixrud and Urzua [2006, Figure 4].

Figure 30.Causal Effect of Schooling on Two Measures of Socioemotional Skills



Notes: Effect of schooling on socioemotional scales for males with average ability, with 95% confidence bands. The locus of control scale is based on the four-item abbreviated version of the Rotter Internal-External Locus of Control Scale. This scale is designed to measure the extent to which individuals believe that they have control over their lives through self-motivation or self-determination (internal control) as opposed to the extent to which individuals believe that the environment controls their lives (external control). The self-esteem scale is based on the 10-item Rosenberg Self-Esteem Scale. This scale describes a degree of approval or disapproval toward oneself. In both cases, we standardize the test scores to have within-sample mean zero and variance one, after taking averages over the respective sets of scales. The model is estimated using the NLSY79 sample. Solid lines depict average test scores, and dashed lines, 2.5%–97.5% confidence intervals.
Source: Heckman, Stixrud and Urzua [2006, Figure 5]

Some other evidence supports the possibility that school can affect measures of intelligence. Cahan and Cohen [1989] use a quasi-experimental paradigm comparing children who differ in both age and schooling to show that schooling increases intelligence test scores independently of age.[187] Schooler and colleagues show that complex (that is, cognitively demanding) work increases intellectual functioning among adults and vice versa (Schooler, Mulatu and Oates [1999]; Kohn and Schooler [1978]).

[187] See Hansen, Heckman and Mullen [2004] for additional estimates of the causal effect of schooling on AFQT.

Evidence from Interventions

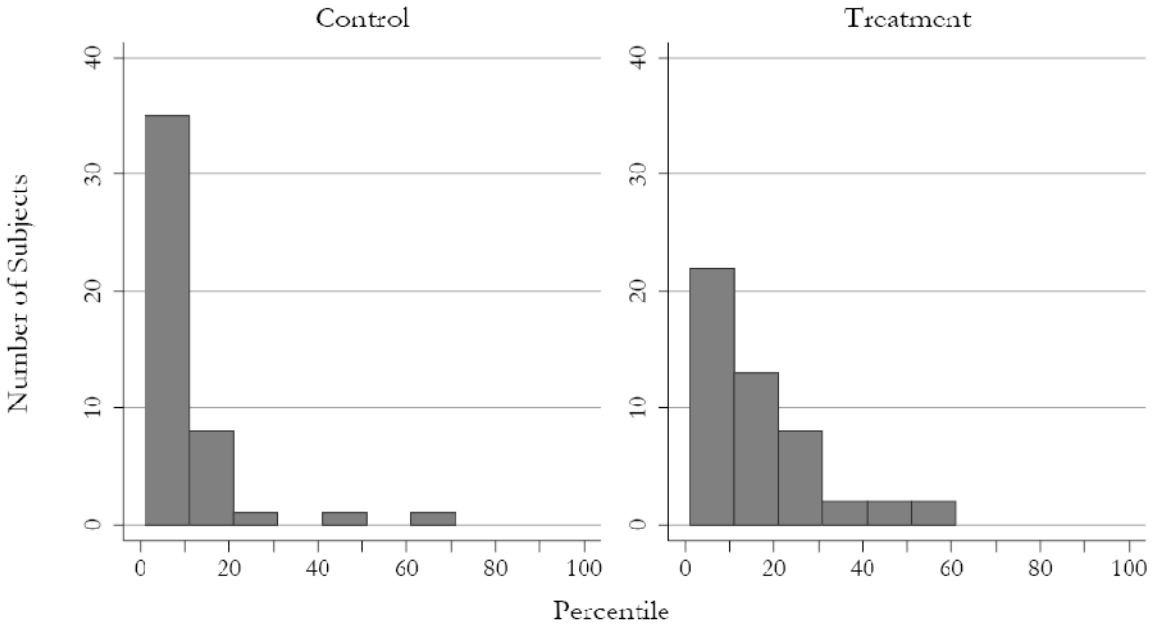
As noted in the introduction, the Perry Preschool Program, did not have a lasting improvement on cognitive ability, but did improve important later-life outcomes through personality (Heckman, Malofeeva, Pinto et al. [2010]). The Perry preschool program enriched the lives of low-income black children with initial IQs below 85 at age 3. In addition, there were home visits to promote parent-child interactions. The program stopped after two years. Participants were taught social skills in a "plan-do-review" sequence where students planned a task, executed it and then reviewed it with teachers and fellow students. They learned to work with others when problems arose.[188] The program was evaluated by the method of random assignment.

The program had strong effects for both boys and girls, although the effects differ by age and outcomes. As noted in the introduction, the program has a statistically significant rate of return of around 6-10% per annum—for both boys and girls—above the post World War II stock market returns to equity in U.S. labor market estimated to be 5.8%. The Perry Preschool Program worked primarily through socioemotional channels. Figure 31 shows that the program improved scores on the California Achievement Test (CAT). The program, however, did not have a lasting effect on IQ scores, consistent with the discussions in Sections 5 and 7 that show that achievement test results are strongly dependent on personality traits (Borghans, Duckworth, Heckman et al. [2008]; Borghans, Golsteyn, Heckman et al. [2009]). Indeed their personalities improved. Participants had better direct measures of personal behavior (a weighted average of "absences and truancies," "lying and cheating," "stealing," and "swears or uses obscene words"

[188] Sylva [1997] describes the Perry program as a Vygotskian program fostering personality traits.

measured by teachers in the elementary school years). Participants improved their internalizing behavior, which, as noted in Section 5, is related to Neuroticism. Heckman, Malofeeva, Pinto et al. [2010] decompose the treatment effects of the Perry Program into components due to experientially-induced changes of cognition (IQ) and the measures of personality at their disposal, and to residual factors. Personality played a more important role in outcomes, particularly in later life.

Figure 31. Perry Age 14 Total CAT Scores, by Treatment Group



CAT = California Achievement Test
Treatment: $N = 49$; Control: $N = 46$
Statistically Significant Effect for Males and Females (p-values 0.009, 0.021 respectively)

Source: Heckman, Malofeeva, Pinto et al. [2010].

Analyses of data Project STAR, a program that randomly assigned kindergarteners and teachers to classes of different sizes, yields similar results to the Perry Program. Using data from

Project STAR, Dee and West [2008] find that assignment to a small class is associated with positive changes in personality. In a follow-up reanalysis, Chetty, Friedman, Hilger et al. [2010] examine the Project STAR program and find that students placed in higher quality kindergarten classes – as measured by their peer's average performance on a Stanford Achievement Test – tend to have higher test scores at the end of kindergarten. The effect fades out over time; by eighth grade, students in better kindergarten classes perform no differently on tests. However, as with the Perry Program, the benefits reemerge later in life. People in better kindergarten classrooms had significantly higher earnings in early adulthood. Furthermore, kindergarten classroom quality also predicted better 4th and 8th grade behavior as measured by teacher-assessed effort, initiative, interest in the class, and disruptive behavior.[189] In turn, behavior predicted earnings in adulthood, suggesting that personality is the channel through which better kindergarten classrooms improve earnings.

The Perry Program and Project STAR did not primarily focus on improving personality, but a few other programs have. The Promoting Alternative Thinking Strategies (PATHS) curriculum teaches self-control, emotional awareness, and social problem-solving skills and is aimed at elementary school children (see Bierman, Coie, Dodge et al. [2010]). A recent random-assignment, longitudinal study demonstrates that the PATHS curriculum reduces teacher and peer ratings of aggression, improves teacher and peer ratings of prosocial behavior, and improves teacher ratings of academic engagement.[190] PATHS is an exemplar of school-based social and emotional learning (SEL) programs, whose impact on both course grades ($d = 0.33$) and standardized achievement tests scores ($d = 0.27$) was recently documented in a meta-analysis of

[189] These scales are based on more detailed questionnaires. Only a subset of the sample has the behavioral measures.
[190] Bierman, Coie, Dodge et al. [2010]

controlled studies involving over 270,000 children in kindergarten through college (Durlak and Weissberg [in press]).[191,192] Likewise, a random assignment evaluation of *Tools of the Mind,* a preschool and early primary school curricula, show that in short-term follow-ups it can improve classroom behavior as well as executive function, defined as higher-level cognitive skills including inhibitory control, working memory, and cognitive flexibility (Barnett, Jung, Yarosz et al. [2008]; Barnett, Yarosz, Thomas et al. [2006]; Bodrova and Leong [2001]; Bodrova and Leong [2007]; Diamond, Barnett, Thomas et al. [2007]). Similar findings are reported for the Montessori preschool curriculum (Lillard and Else-Quest [2006]).

There is also evidence that targeted intervention efforts delivered to individual children can improve aspects of Conscientiousness. These studies are typically more short-term and, in contrast to the multi-faceted curricula described above, are designed to isolate a particular mechanism for behavior change. For instance, Rueda and colleagues [2005] designed a set of computer exercises to train attention in children between four and six years of age. Children in the intervention group improved in performance on computer tasks of attention relative to children who instead watched interactive videos for a comparable amount of time. Similarly, Stevens and colleagues [2008] designed a six-week computerized intervention and showed that it can improve selective auditory attention (i.e., the ability to attend to a target auditory signal in the face of an irrelevant, distracting auditory signal).

Several studies suggest that personality can be remediated in adolescence. Martins [2010] analyzes data from EPSIS, a program developed to improve student achievement of 13-15 year-olds in Portugal by increasing motivation, self-esteem, and study skills. The program consists of

---

[191] "*d*" is the difference between the means divided by the standard deviation of the outcome.
[192] Note however that the largest federal study to date on character education programs, including PATHS, failed to find evidence for improvements in behavior or academic performance (see Social and Character Development Research Consortium [2010]).

one-on-one meetings with a trained staff member or meetings in small groups. The intervention was tailored to each participant's individual skill deficit. Overall, the program was successful, cost-effectively decreasing grade retention by 10 percentage points. Bloom, Gardenhire-Crooks and Mandsager [2009] analyze the data from the National Guard Youth ChalleNGe program, a 17-month intervention for youth who have dropped out of high school. While the program does not require military enrollment, it stresses aspects of military discipline. The program features a two-week assessment period, a 20-week residential program often conducted at a military base, and a one-year mentoring program. 9 months after entry, participants in the program were 12% more likely to obtain a high school, 9% more likely to be working full time, and less likely to be arrested. Furthermore, participants had higher levels of self-efficacy (a trait related to Emotional Stability), suggesting that personality change might have helped with the improvements. However, the 9-month follow-up period might be too short to know whether the program has long-lasting effects. While these studies show that personality can be improved through intervention, a couple of other studies show less promising results when the interventions were targeted at adolescents (Rodríguez-Planas [2010], Holmlund and Silva [2009]).

The above interventions consider long-term programs. Behncke [2009] provides some experimental evidence that short-term exogenous shocks to non-cognitive skills affect test performance. She finds giving words of encouragement, an intervention that might boost short-term self-efficacy or self-esteem, before a diagnostic math test was associated with 2.5% higher scores amongst all students $(p < 0.05)$ and 8% higher scores amongst those with self-reported difficulties with math $(p < 0.01)$. The result suggests that non-cognitive skills can be shaped, even in the very short-term.

The evidence for adults corroborates the finding of Cunha, Heckman and Schennach

[2010] for children. Personality is malleable throughout the life cycle. For example, Gottschalk [2005] shows evidence from a randomized control trait that working at a job can improve locus of control. He uses data from the Self-Sufficiency Project (SSP) in which some welfare recipients were randomly offered substantial subsidies to work. The subsidy more than doubled the earnings of a minimum wage worker, and people in the experiment group worked about 1/3 more hours than those in the control group. After 36 months, those who received the subsidy were more likely to have an improved locus of control.

Several other studies find similar results. Clausen and Gilens [1990] claim that female labor force participation increases self-confidence. Roberts [1997] reports an increase of social dominance and Roberts and Chapman [2000] a decrease in Neuroticism for working women. Others show that marital and family experiences shape personality (Helson and Picano [1990], Roberts, Helson and Klohnen [2002]).[193] However, none of these studies have the random assignment factors of the Gottschalk study.

Personality may even be malleable at the end of life. Jackson, Hill, Payne et al. [2010] investigate causal mechanisms behind the association between Openness to Experience and IQ, using data from a 16-week intervention designed to boost inductive reasoning for elderly people. The intervention consisted of lab training for how to recognize novel patterns and around 10 hours a week of solving crossword, Sudoku, and logic puzzles. Controlling for inductive reasoning, self-reported Openness to Experience increased for participants during the training program relative to those in a waitlisted control group. However, the elderly people were not

[193] We note that there is controversy in the literature about the validity of conventional estimates of heritability. It centers on the linearity and additivity assumptions, the assumed absence of interactions between genes and environment, and the assumption that genes do not select environments.

followed after the program to determine whether the change was long-lasting nor were other important outcomes, like life expectancy, tracked.

Table 14. The Effect of Interventions on Personality

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Barnett, Jung, Yarosz et al. [2008]	Outcome(s): *internalizing and externalizing behavior* – teacher-assessed Problem Behaviors Scale of the Social Skills Rating System (SSRS) Intervention: participation in a year-long Tools of the Mind preschool program compared to a generic curriculum	Data: collected by authors; 210 children aged 3 and 4 Methods: Students were randomly assigned to classrooms within the same school after parental consent was obtained. Teachers were randomly assigned to control and treatment classrooms.	Control Variables: n/a Timing of Measurements: *Baseline* – Behavior measures were taken prior to the program in October-November of 2002. *Post-treatment* – Behavioral measures taken immediately after the program in May-June of 2003.	Participants in the program had a 0.47 standard deviation lower score for the behavioral problems index ($p<0.05$).
Behncke [2009]	Outcome(s): *cognitive ability* – performance on a diagnostic math test for a college economics class Intervention: verbal encouragement before the test	Data: Collected by author; 440 students from a Swiss University Methods: The treatment was randomly assigned to already-established classroom sections. Students were unaware they were in an experiment.	Control Variables: n/a Timing of Measurements: *Post treatment* – The diagnostic test was given immediately after the treatment.	Verbal encouragement raised test scores by 2.5% amongst all students ($p<0.05$) and by 8.0% amongst students who reported difficulties with math ($p<0.01$).
Bierman, Coie, Dodge et al. [2010]	Outcome(s): *teacher-assessed behavior* –Social Health Profile (SHP) including authority acceptance, cognitive concentration, and social competence; *peer-assessed behavior* – survey questions about behavior labeled as aggressive, prosocial, and hyperactive Intervention: – participation in a 3-year-long Fast Track PATHS program focused on improving self-control and positive social behavior	Data: 2,937 children (grades 1-3) Methods: School administrators were offered participation in the experiment, knowing the school would receive treatment with a 50% probability.	Control Variables: time, time squared, individual baseline, school baseline, city fixed effects, poverty level, interactions of intervention with time, time squared, individual baseline, poverty, and poverty and time Timing of Measurements: *Baseline* – Behavioral measures were taken prior to the program in the fall of 1st grade. *Post treatment* – Behavioral measures were taken again in the spring of 3rd grade around the end of the program.	Immediately after the 3-year program, participation was associated with a 0.24 standard deviation increase in authority acceptance ($p<0.001$), a 0.12 standard deviation increase in cognitive concentration ($p<0.001$), and a 0.34 standard deviation increase in social competence ($p<0.0001$) compared to the control group. The effects were stronger in more disadvantaged schools. Similar but weaker results apply for the peer-assessed measures.

Bloom, Gardenhire-Crooks and Mandsager [2009]	Outcome(s): *educational attainment* – high school diploma, *labor force participation* – whether working at a job, *personality* – self-efficacy and social adjustment Intervention: participation in the ChalleNGe program consisting of a 2-week assessment period, 20-week residential program often conducted at a military base, and a 1-year mentoring program.	Data: 1,018 young people between the ages 16 and 18 who have dropped out of school Methods: The control group was constructed out of applicants who qualified for the program but were not taken due to lack of space.	Control Variables: sample member characteristics Timing of Measurements: *During treatment* – Outcomes were measured approximately 9 months after entering the study.	Participants in the program were 12.0 percentage points more likely to earn a high school diploma ($p<0.01$), 9.1 percentage points more likely to be working ($p<0.01$), and 9.6 percentage points less likely to report a self-efficacy and social adjustment score one standard deviation below the mean ($p<0.01$). The program also improved measures of criminality and health.
Chetty, Friedman, Hilger et al. [2010]	Outcome(s): *non-cognitive skills* – an index based on the teacher's observations of the students Intervention: randomly assigned kindergarten class quality as measured by difference in percentiles of the mean end-of-year test scores of the students' classmates and the scores of the other kindergarteners at the same school	Data: Project STAR; 1,671 4th grade students and 1,780 8th grade students Methods: Students and teachers were randomly assigned in kindergarten to classrooms of different sizes. The students were assigned to the same size classroom through 3rd grade.	Control Variables: wave fixed effects, student gender, free-lunch status, age, race, a quartic in the claiming parent's household income interacted with parent's marital status, mother's age at child's birth, whether the parents own a home, and whether the parents made a 401 (k) contribution between 1996 and 2008 Timing of Measurements: *During treatment* – Age-relevant SAT tests were administered in kindergarten and grades 1-3. *Post treatment* – Age-relevant SAT test were and behavioral surveys were given in 4th and 8th grade. College quality and attendance was at age 19. Earnings were at age 27.	A 1 percentile improvement in kindergarten class quality increases an index of non-cognitive skills by 0.15 percentiles in 4th grade ($p<0.05$) and 0.13 percentiles in 8th grade ($p<0.05$). Better classrooms were also associated with better life outcomes.
Diamond, Barnett, Thomas et al. [2007]	Outcome(s): *Executive Function* – Dots-Mixed task, Reverse-Flanker task Intervention: participation in a Tools of the Mind program instead of the regular school district's balanced literacy program	Data: 147 preschoolers Methods: Teachers and students were randomly assigned to classrooms within the same school.	Control Variables: age, gender, years in program Timing of Measurements: *Post-treatment* – The tasks were given at the end of the program 2nd year of the program.	84% of students in Tools were successful in the Reverse Flanker task compared to 65% in the control group. Almost twice as many students in the Tools program achieved greater than 75% accuracy on the Dots-Mixed task compared to the control group.

Durlak and Weissberg [in press]	Outcome(s): social and emotional learning skills, attitudes, positive social behavior, conduct problems, emotional distress, academic performance Intervention: Meta-analysis of school-based, universal social and emotional learning program.	Data: 270,034 kindergarten through high school students Methods: All studies include a control group.	Control Variables: n/a Timing of Measurements: All studies contained follow-up data at least 6 months after the intervention.	The mean difference in standard deviations between the treatment and control groups are as follows: social and emotional learning skills = 0.57 ($p<0.05$); attitudes = 0.23 ($p<0.05$); positive social behavior = 0.24 ($p<0.05$); conduct problems = 0.22 ($p<0.05$); emotional distress = 0.24 ($p<0.05$); academic performance = 0.27 ($p<0.05$). All variables are coded so that positive numbers reflect better outcomes.
Gottschalk [2005]	Outcome(s): *Personality* –four measures of locus of control based on whether the respondent agrees strongly, agrees, disagrees, or strongly disagrees with statements Intervention: A subsidy for full-time work during a 36-month period	Data: Self-Sufficiency Project; 4,958 single parents over the age of 19 in New Brunswick and British Columbia Methods: The subsidy was randomly offered to a population of people receiving Income Assistance (IA)	Control Variables: age, age squared, region, gender, speaks French, number of children Timing of Measurements: *Baseline* – Locus of control was measured before the program. *During treatment* – Locus of control was measured again 18 and 36 months after the baseline.	Using whether the participant received the subsidy as an instrument for hours worked, the authors find that working tends to improve locus of control by the 36 month re-interview.
Heckman, Malofeeva, Pinto et al. [2010]	Outcome(s):*externalizing behavior, internalizing behavior* – measured using Pupil Behavior Inventory (PBI) of teacher reports Intervention: participation in the Perry Preschool Program, an intervention that lasted 2 years and enriched the lives of low income black children	Data: Perry Preschool Program; 123 preschool students Methods: The students were randomly assigned to treatment through a complex procedure.	Control Variables: n/a Timing of Measurements: *Post treatment* – The measure of externalizing and internalizing behavior are taken ages 7-9 (2-4 years after treatment). Other life outcomes were measured at ages 19, 27, and 40.	The intervention improved mean externalizing behavior for both males and females ($p<0.05$). It had a borderline statistically significant impact on internalizing behavior. The program also benefited a wide range of later life outcomes primarily through non-cognitive skills.

Holmlund and Silva [2009]	Outcome(s): *academic performance* – average of standardized test scores in English, Math, and Science Intervention: participation in the "xl programme" targeting the non-cognitive skills of secondary school students aged 14	Data: "xl club programme," National Pupil Database (NPD), Pupil Level Annual Schools Census (PLASC) ; 2,333 and 259,189 treated and control students aged 14 in England (2004) Methods: logit, propensity score matching, OLS, difference-in-difference, double differences, random-growth model	Control Variables: sex, language, eligibility for school meals, special needs status, and race Timing of Measurements: *Baseline* – Standardized exams were taken at age 11 and age 14 before the start of the program. *Post treatment* – Standardized national exams were taken again at age 16 at the end of the program (2 years after the beginning of the program).	Unconditional on observables, the performance of the students in the xl club is 1.2 to 1.4 standard deviations lower than the control subjects ($p<0.01$). Using OLS , the effect is -0.17. The propensity score estimates are -0.13 and -0.15. For the difference-in-difference models estimated using OLS and propensity score matching, there is no longer a significant effect of the program in either direction. Overall the program had little effect.
Social and Character Development Research Consortium [2010]	Outcome(s): *Social and Emotional Competence* – self-efficacy for peer interaction, normative beliefs about aggression, empathy; *Behavior* – altruistic behavior, positive social behavior, problem behavior, ADHD-related behavior; *Academics* – engagement with learning, academic competence and motivation; *Perceptions of School Climate* – positive school orientation, negative school orientation, student afraid at school, victimization at school, feelings of safety, student support for teachers Intervention: 7 different programs (ABC, CSP, LBW, PA, PATHS, 4Rs, SS) aimed to build Social and Character Development (SACD) compared to the "standard practice" programs at non-treated schools	Data: Social and Character Development (SACD) Research Program; around 6,000 elementary school students Methods: Schools were first asked to participate in the program and were then randomly assigned one of the 7 SACD programs or left with their traditional curriculum. The data were analyzed using HLM.	Control Variables: gender, race, parental education, family structure, household income, measures of poverty, parental labor force participation, teacher race, teacher experience. (Note: the specific set depended on the outcome of interest.) Timing of Measurements: *Baseline* – Initial measures were collected near the start of the program in the fall of 2004. *During treatment* – Data were collected in the spring of 2005, the fall of 2005, and the spring of 2006. *Post treatment* – Data were collected near the end of the program in the spring of 2007.	Fall 2003 to Spring 2005: Of the 20 outcomes, the only significant effects were that participation in any program was associated with a 0.07 standard deviation higher primary caregiver-reported altruistic behavior ($p<0.10$), a 0.06 standard deviation lower child-reported altruistic behavior ($p<0.10$), and a 0.12 standard deviation higher teacher-reported student support for teachers ($p<0.05$). Fall 2003 to Spring 2006: Of the 20 outcomes, the only significant effects were that participation in any program was associated with a 0.07 standard deviation lower child-reported self-efficacy for peer interactions ($p<0.10$) and 0.16 standard deviation higher teacher-reported student support for teachers ($p<0.05$). Fall 2003 to Spring 2007: There were no statistically different effects of participating in any program. Other Analyses: The results were similar when analyzing each of the programs separately and when using growth curves. There is some evidence that programs were beneficial for high-risk students.

Jackson, Hill, Payne et al. [2010]	Outcome(s): *Personality* – Openness to Experience Intervention: participation in a 16-week inductive reasoning training program coupled with 10 hours of puzzle solving per week	Data: collected by the authors; 183 adults aged 60 to 94 Methods: Participants were randomly assigned to treatment and control groups after deciding to participate in the experiment.	Control Variables: n/a Timing of Measurements: *Baseline* – Openness to Experience was measured pre-treatment. *During treatment* – Openness to Experience was measured at week 5and at week 10. *Post treatment* – Openness to Experience was measured at the end of the program in week 16.	On average, participants in the program were 0.39 standard deviations higher in Openness to Experience after the program relative to people in the control group ($p<0.05$).
Martins [2010]	Outcome(s): *Educational attainment* – grade retention Intervention: participation in the EPIS program that boosts non-cognitive skills including motivation, self-esteem, and study skills	Data: EPIS database; 15,307 students in 7th -9th grade in Portugal Methods: linear probability model, quasi-randomization	Control Variables: student fixed effects, time fixed effects Timing of Measurements: *Baseline* – Measures of academic achievement were taken before the intervention in 7th and 8th grade. *During treatment* – Measures were taken each quarter that the students participate in the program through 7 academic quarters after the beginning of the program (students entered the program at different times and remained in treatment for different lengths of time but were followed if they left treatment).	The program reduced annual grade retention by at least 10.1 percentage points ($p<0.001$).

Rodríguez-Planas [2010]	Outcome(s): *educational attainment* – high-school completion and post-secondary education; *academic achievement* – math test score percentile, reading test score percentile , GPA; *labor market success* – earnings during the last year of the program, 3 years after the program, and five years after the program Intervention: – participation in the Quantum Opportunity Program (QOP) that was available for 5 years, centered around mentoring, developing social skills, community service, and providing incentives for academic success for 9th graders	Data: Quantum Opportunity Program (QOP); 1,069 students from seven large US cities Methods: Students in schools participating in the program were randomly assigned to treatment or control groups.	Control Variables: n/a Timing of Measurements: *Post treatment* – Interviews were conducted during the last year of the program, 3 years after the program, and 5 years after the program.	During last year of the program: Participation in the program was associated with a 7 percentage point increase in the probability of graduating high school ($p<0.10$) and 6 percentage point increase in the probability of attending college ($p<0.10$). There were no differences in academic achievement. 3 years after the program: Participation in the program was associated with a 7 percentage point increase in the probability of ever attending college ($p<0.10$), 9 percentage point increase in the probability of attending college ($p<0.05$), and a 7 percentage point decrease in the probability of having a job ($p<0.10$). 5 years after the program: There are no significant differences 5 years after the program. Findings for sub-populations: The program benefited people who were 14 or less upon entering high school significantly more than older students. It also tended to benefit girls more than boys.
Stevens, Fanning, Coch et al. [2008]	Outcome(s): *attention* – ERP index of selective auditory attention; *language skills* – Clinical Evaluation of Language Fundamentals-3 Intervention: Participation in a six-week (100 min/day) computerized training program for boosting language skills (Fast ForWord program)	Data: collected by the authors; 33 children aged 7 on average Methods: The students who received treatment were compared to a control group who did not.	Control Variables: Test scores were normalized by age Timing of Measurements: *Baseline* – Measures were taken right before the start of the program. *Post treatment* – Measures were taken again at the end of the program (6 weeks after the start).	The increase in the attention was 0.81 standard deviations higher for the participants than for the non-participants ($p<0.01$). The increase in the receptive language scores was 0.91 standard deviations higher in the participants than for the control group ($p<0.01$). There was no significant effect on expressive language scores between the participants and the control group.

Evidence from Psychotherapy

The accomplishments of psychotherapy also support the possibility of intentional, mean-level, and rank-order change. In a 1980 meta-analysis, Smith, Glass and Miller summarized 475 controlled studies, concluding that individuals who undergo psychotherapy are about 0.85 standard deviations better on outcome measures than those who do not. The large benefits of therapy are not permanent, however: the effect of psychotherapy over control conditions falls to about half a standard deviation two years after therapy is concluded. Moreover, it is not clear that the effects of psychotherapy on individuals who seek change generalize to individuals who are not actively seeking treatment for a condition that causes them distress.[194]

More evidence on the possibility of intentional change comes from the psychological literature on expertise. Ericsson, Krampe and Tesch-Romer [1993] demonstrate across domains as diverse as chess, musical performance, and digit span memory, that thousands of hours of sustained, deliberate practice lead to dramatic improvements in skill. Ericsson points out that the top performers in nearly every field do not reach world-class levels of skill until at least ten years of deliberate practice.[195]

8.F. Stability of Economic Preference Parameters

Less is known about the stability of economic preferences. To our knowledge, no longitudinal study has measured the mean-level or rank-order stability of time preference over the life cycle (Frederick, Loewenstein and O'Donoghue [2002]). A handful of cross-sectional studies using relatively small samples have examined mean-level stability, and their findings are

[194] Some evidence that further intervention can produce enduring change in non-clinical populations comes from Gillham and Reivich [1999] who show that children taught to make more optimistic causal attributions about negative events maintain this optimistic outlook two years post-intervention.
[195] See Ericsson and Ward [2007] for a recent review of the evidence.

mixed. Green, Fry and Myerson [1994a] and Harrison, Lau and Williams [2002] find that discount rates are lower among older individuals. On the other hand, Chesson and Viscusi [2000] claim to find that older adults have *higher* discount rates than younger adults. Chao, Szrek, Sousa Pereira et al. [2007], de Wit, Flory, Acheson et al. [2007], and Coller and Williams [1999] find no relationship between age and discount rate. Finally, Read and Read [2004] find a curvilinear relationship in which older people discount more than younger people, and middle-aged people discount less than either group. Sahm [2007] shows that risk aversion increases with age. Table 15 below summarizes the findings for a variety of recent economic studies on the heritability, malleability and stability of preferences and personality.

In summary, the answer to the question of whether change in personality is possible must be a definitive yes, both in terms of mean-level and rank-order change. However, change may be more difficult later in the life cycle, change may be more enduring for some (such as more emotionally stable individuals) than for others, change may require persistent and consistent environmental pressure (as opposed to transient pressure from short-term interventions), and there are powerful forces for stability (such as genes and habit) which make change difficult.

Table 15. The Heritability, Malleability, and Stability of Preferences and Personality

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Booth and Nolen [2009]	Outcome(s): *risk aversion* – choice whether to accept a real-stakes lottery vs. a certain payment Explanatory Variable(s): *short-term gender environment* – whether the student was assigned to a coed or single-sex group during the experiment; *long-term gender environment* – whether the student attends a coed or single-sex school	Data: Collected by the authors; 260 students in grades 10 and 11 from eight publicly funded schools in England (2007) Methods: probit, IV, propensity score matching	Controls: n/a Timing of Measurements:: The measures are contemporaneous. Theory: Growing up in an environment with males might cause girls to act more "feminine" and take fewer risks. Similarly, boys in coed environments might exhibit more risk-taking in coed environments to try to impress girls.	Girls from coed high school in England were 36% ($p<0.01$) less likely to accept a real-stakes lottery. Girls assigned to experimental group with all girls were 12% ($p<0.10$) more likely to accept the lottery than girls in coed experimental groups.
Burks, Carpenter, Goette et al. [2010]	Outcome(s): *demand for information* – whether people request the results of their IQ and numeracy tests; *overconfidence* – the difference between ex-ante estimate of quintile in the IQ distribution and the true quintile in the IQ distribution Explanatory Variable(s): *self-assessment* – before and after test assessments of the quintile of performance on the tests; *personality* – a self-reported measure of harm avoidance, social closeness, social potency, and stress reaction	Data: Collected by authors, administrative data from a human resources department; 1,063 trainee truckers from a U.S. trucking company Methods: probit, ordered probit, linear spline	Controls: (1) actual test performance (2) controls in (1), harm avoidance, and education levels (3) controls in (2), ethnicity, sex, age, age squared, and household income; (4) controls in (3), before test belief, and post-test belief on other test Timing of Measurements: People are asked about their expected performance on the IQ and numeracy tests before and after they take the test. Later, they are also given the option to receive the results of their tests. Theory: (a) People misjudge their own ability due to systematically biased noisy signals (b) People value their self-assessed ability and avoid updating when the assessment is positive. (c) People sub-consciously misrepresent their own ability for strategic advantage.	Demand for information: A one quintile increase in a person's post-test belief about their test performance is positively associated with a 3.1 (1), 2.9 (2), 3.0 (3), 3.0 (4) percentage point higher probability of demanding information about the IQ test ($p<0.01$) and a 6.0 (1), 5.7(2), 5.8(3), 3.9(4) percentage point higher probability of demanding information about the numeracy test ($p<0.01$). Overconfidence: Harm avoidance and stress reaction are negatively correlated with overconfidence on the IQ test ($p<0.01$,$p<0.05$). Social potency is positively linked to overconfidence on the IQ test ($p<0.01$). Stress reaction is negatively associated with overconfidence on the numeracy test ($p<0.01$). Social potency is positively associated with overconfidence on the numeracy test ($p<0.05$).

Dohmen, Falk, Huffman et al. [2009a]	Outcome(s): *risk preference* – survey responses on an 11 point scale, relating to general risk preference and risk preference relating to car driving, financial matters, leisure and sports, career and health. Explanatory Variable(s): (see controls)	Data: Collected by the authors/ German Socio-Economic Panel (SOEP); 450 adults from Germany/22,019 people living in Germany Methods: interval regression, probit	Controls: (1) sex, age, and height (2) controls in (1) and parental education (3) controls in (2) and 2002 household wealth, (4) controls in (2), 2003 household income, (5) controls in (2), household income 2004, (6) controls in (2), household wealth, 2004 household income, additional demographic variables. Timing of Measurements: The measures are contemporaneous. Theory: People might have a stable, underlying preference for risk cross contexts.	Determinants of risk attitude: being female and age are negatively associated with willingness to take risks ($p<0.01$;(1)-(6)). Height is positively associated with a general willingness to take risks ($p<0.01$;(1)-(6)). Mother and father's education is positively associated with willingness to take risks ($p<0.01$;(1)-(5)). Stability of risk: The 6 measures of contextualized risk aversion are correlated with each, other ranging from 0.456 – 0.609.
Einav, Finkelstein, Pascu et al. [2010]	Outcome(s): *risk preference* – order rankings of observed decisions to purchase different purchase insurance for health, prescription drugs, dental, short-term and long-term disability and 401(k) plans Explanatory Variable(s): *predictable risk* – predictions from modeling risk based on observables; *idiosyncratic risk* – realization of risk in the net period	Data: Administrative data; 12,752 employees of Alcoa, Inc. (2004) Methods: Spearman correlations, OLS	Controls: (1) none (2) the menu of benefits the employee faced (3) controls in (2) and predictable and idiosyncratic risk Timing of Measurements: Most of the financial decisions were made in the same year. Theory: There is an underlying preference for risk that applies across many contexts.	The average correlations between the various domains are 0.192, 0.188, and 0.164. $p<0.01$ for all of the pairwise correlations, except between 401(l) and short-term disability ((1)-(3)), long-term disability ((1)-(3)), and dental (3). The main results withstand several robustness checks.

Kosfeld, Heinrichs, Zak et al. [2005]	Outcome(s): *trust* – willingness to "invest" in a real-stakes two-player trust game; *risk preference* – real stakes trust game played against a computer that randomly gave payoffs; *altruism* – the amount transferred back by the investee in the trust game (there are no monetary incentives to do so) Explanatory Variable(s): *biological determinant of trust* – nasal spray of oxytocin	Data: Experiment conducted by the authors; 194 male university students in Germany Methods: Mann-Whitney U-test, RCT	Controls: n/a (RCT) Timing of Measurements: The measure were contemporaneous. Theory: There is a notion of "trust" distinct from altruism and risk preference.	People who receive the oxytocin nasal spray, invest on average 17% more than those who do not ($p<0.05$). Risk behavior does not differ between the two groups. Trustees do not show more altruistic behavior when given oxytocin.
Le, Miller, Slutske et al. [2010]	Outcome(s): *risk preference* – response to a 10-point survey question about willingness to take risks in general, response to a 10-point survey questions about how conservative the subject is in making decisions to spend money Explanatory Variable(s): *genetic makeup* – differences in outcomes between monozygotic and dyzygotic twins	Data: Australian Twin Study of Gambling; 1,875 complete twin pairs Methods: OLS	Controls: (1) none (2) gender, age, education, and marital status (3) gender and gender - risk interaction (4) gender, gender - risk interaction, age, education, and marital status Timing of Measurements: The measures are contemporaneous. Theory: Risk preferences could be genetically heritable.	Heritability of the risk measure is 0.234 (1), 0.192 (2), 0.224 (3), and 0.196 (4) (t-stat=7.00,5.79, 4.09, 3.64) . Heritability of the conservative measure is 0.154 (1), 0.134 (2), 0.140 (3), and 0.135 (4) ($p<0.01$)

Sutter, Feri, Kocher et al. [2010]	Outcome(s): *social preferences* (*selfish* – the agent maximizes their own payoff; regardless of the other person's, *efficient* – maximizing the sum of the allow payoffs*; maximin* – maximizes the minimum of the two payoffs; *FS inequality* – values own payoff plus a weighted average of the difference between own payoff to the others payoffs; *ERC inequality* – people get disutility if their payoff deviates from the group average) – choices of allocating resources between peers in a real-stakes experiment Explanatory Variable(s): n/a	Data: Collected by the authors; 883 students aged 8 to 17 living in Australia (2008) Methods: maximum likelihood error-rate analysis	Controls: n/a Timing of Measurements: The measures are contemporaneous. Theory: Social preferences might change with age.	20% of children behave selfishly, a finding that is stable across age and gender. An increase in one year of age leads to a 0.044 increase in the probability of have efficiency preferences for males ($p<0.01$), but has not effect for females.

9. Summary and Conclusions

We summarize this chapter by providing provisional answers to the eight questions posed in Section 1.

(1) *How can we fit the psychological concept of personality into an economic framework? Can conventional models of preferences in economics explain the main theories in personality psychology?*

We have defined personality as a response function of agents that depends on situations (including incentives), endowments of traits, information, and resources within a conventional economic model. Psychologists analyze a richer class of actions than economists normally consider. We show how to integrate these actions into economic theory. The leading models of personality psychology are special cases of our model.[196]

(2) *What are the main measurement systems used in psychology for personality, and how are they validated? How are the different systems related to each other? What is the relationship between measures of personality and measures of psychopathology?*

In Section 5, we exposit the main systems for measuring personality, focusing primarily, but not exclusively, on the Big Five model. We consider the strengths and limitations of the systems and the relationships among competing systems. We show how measures of psychopathology are extreme manifestations of personality traits. We link specific diagnoses of pathology with conventional measures of psychological traits.

[196] Freudian models of the unconscious would make the traits that govern behavior, and especially η, unknown to agents but nonetheless governing choices. A pure model of behaviorism would feature the effects of constraints on choices. Borghans, Duckworth, Heckman et al. [2008] develop such a model.

(3) *What is the relationship between economic preference parameters and psychological measurements?*

We review an emerging body of research that relates economic preference parameters (risk aversion, time preference, ambiguity aversion, social preferences) to Big Five traits and to measures of self-esteem and personal control that are linked to the Big Five traits. Time preference is negatively correlated with IQ and the ability to control attention. Risk preference is negatively correlated with IQ and other measures of cognition. Higher IQ people are more consistent in their choices under uncertainty. While risk aversion is related to personality traits, the available evidence suggests that marginal ambiguity aversion is not. Social preferences are predicted by measured personality traits, but the body of evidence on this question is not large.

(4) *How stable across situations and over the life cycle are preference parameters and personality traits?*

We review the history of the person-situation debate between the social psychologists who maintain the primacy of the situation in determining behavior and the traits theorists who maintain the primacy of traits in explaining behavior. Behavioral economists, as a group, have adopted the situationist point of view. Extreme advocates of the situationist point of view claim that there is no personality construct. The issue hinges on the nonlinearity of action, effort, and productivity functions. In the presence of such nonlinearities, measured traits (e.g., actions) depend on situations and tasks.

A large body of evidence suggests that nonlinearity is an empirically important phenomenon. Nonetheless, a large body of evidence suggests that there are stable personality traits that predict a variety of behaviors in different situations. Personality is

not an ephemeral creation of situations, nor is its manifestation invariant across situations. Personality traits are not set in stone. They change over the life cycle. The evidence on the stability of preference parameters across situations and over life cycle is less ample. There is some evidence that simple separable models of preferences are inadequate descriptions of choice behavior. There is little evidence on the stability of traits over the life cycle.

(5) *What is the evidence on the predictive power of cognitive and personality traits?*

We present a large body of evidence that shows strong associations between personality traits and educational, labor market, health, and criminal outcomes.

(6) *What is the evidence on the causal power of personality on behavioral outcomes?*

Few of the correlational studies relating personality to outcomes have a firm causal basis. Personality psychologists do not yet attempt to establish the causal status of personality. Recent studies in economics establish causal status of certain personality traits on outcomes under their maintained assumptions, which are inevitably subject to debate. This is an area likely to flourish in the coming years.

(7) *Can personality be altered across the life cycle? Are interventions that boost personality traits likely fruitful avenues for policy?*

There is a small but growing body of intervention studies that establish that personality traits can be altered over long periods of time in response to interventions. Some of the major effects of early childhood intervention programs appear to operate through their lasting effects on personality. Family investment decisions also change personality. The evidence to date suggests that interventions that boost personality traits can be effective in promoting adult success.

(8) *Do the findings from psychology suggest that conventional economic theory should be enriched?*

The evidence from psychology enriches economics by providing a more nuanced interpretation of human choice and actions. It promises to provide a deeper understanding of conventional economic preference parameters and how they arise. Unfortunately, at the time of this writing, this promise remains unfulfilled. Given the current state of evidence against conventional economic preference specifications (see, e.g., Starmer [2000] and the evidence in Section 6) this line of research is promising.

References

Abbring, Jaap H. and Heckman, James J. (2007). "Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation." in *Handbook of Econometrics*. J. Heckman and E. Leamer, eds. Amsterdam, Elsevier. **6B:** 5145-5303.

Achenbach, TM, McConaughy, SH and Howell, CT (1987). "Child/Adolescent Behavioral and Emotional Problems: Implications of Cross-Informant Correlations for Situational Specificity." *Psychological bulletin* **101**(2): 213-232.

Ackerman, Phillip L. and Heggestad, Eric D. (1997). "Intelligence, Personality, and Interests: Evidence for Overlapping Traits." *Psychological Bulletin* **121**(2): 219-245.

ACT, Inc. (2007). *The ACT Technical Manual*. Iowa City, IA, ACT, Inc.

Agnew, Robert, Brezina, Timothy, Wright, John Paul and Cullen, Francis T. (2002). "Strain, Personality Traits, and Delinquency: Extending General Strain Theory." *Criminology* **40**(1): 43-72.

Aigner, Dennis J., Hsiao, Cheng, Kapteyn, Arie and Wansbeek, Tom (1984). "Latent Variable Models in Econometrics." in *Handbook of Econometrics*. Z. Griliches and M. D. Intriligator, eds. Amsterdam, Elsevier. **2:** 1321-1393.

Akerlof, GA and Dickens, WT (1982). "The Economic Consequences of Cognitive Dissonance." *The American Economic Review* **72**(3): 307-319.

Alexander, Karl L., Entwisle, Doris R. and Dauber, Susan L. (1993). "First-Grade Classroom Behavior: Its Short-and Long-Term Consequences for School Performance." *Child Development* **64**(3): 801-814.

Allport, Gordon W. and Odbert, Henry S. (1936). "Traitnames. A Psycho-Lexical Study." *Psychological Monographs.47*: 171.

Altmann, Steffen, Dohmen, Thomas and Wibral, Matthias (2008). "Do the Reciprocal Trust Less?" *Economics Letters* **99**(3): 454-457.

American Council On Education (2007). "ACT Technical Manual." ACT Inc.

American Council on Education (2009). "Technical Manual: 2002 Series GED Tests." American Council on Education

American Psychological Association (2007). *APA Dictionary of Psychology*. Washington, DC, American Psychological Association.

Andreoni, James (1995). "Warm-Glow Versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments." *The Quarterly Journal of Economics* **110**(1): 1.

Ardila, Alfredo, Pineda, David and Rosselli, Monica (2000). "Correlation between Intelligence Test Scores and Executive Function Measures." *Archives of Clinical Neuropsychology* **15**(1): 31.

Aristotle (1956). *The Nicomachean Ethics*. Cambridge, MA, Harvard University Press.

Ashton, Michael C., Lee, Kibeom, Perugini, Marco, Szarota, Piotr, de Vries, Reinout E., Di Blas, Lisa, Boies, Kathleen and De Raad, Boele (2004). "A Six-Factor Structure of Personality-Descriptive Adjectives: Solutions from Psycholexical Studies in Seven Languages." *Journal of Personality and Social Psychology* **86**(2): 356-366.

Atkinson, Anthony B. (1970). "On the Measurement of Inequality." *Journal of Economic Theory* **2**(3): 244-266.

Ayllon, Teodoro and Kelly, Kathy (1972). "Effects of Reinforcement on Standardized Test Performance." *Journal of Applied Behavior Analysis* **5**(4): 477-484.

Bagby, R. Michael, Bindseil, Kirstin D., Schuller, Deborah R., Rector, Neil A., Young, L. Trevor, Cooke, Robert G., Seeman, Mary V., McCay, Elizabeth A. and Joffe, Russell T. (1997). "Relationship between the Five-Factor Model of Personality and Unipolar, Bipolar, and Schizophrenic Patients." *Psychiatry Research* **70**: 83-94.

Bandiera, Oriana, Barankay, I and Rasul, I (2005). "Social Preferences and the Response to Incentives: Evidence from Personnel Data*." *Quarterly Journal of Economics*.

Bandura, Albert (1999). "Social Cognitive Theory of Personality." in *Handbook of Personality*. New York, NY, US, Guilford Press**:** 154-196.

Bardsley, Nicholas (2008). "Dictator Game Giving: Altruism or Artefact?" *Experimental Economics* **11**: 122-133.

Barenbaum, Nicole B. and Winter, David G. (2008). "History of Modern Personality Theory and Research." in *Handbook of Personality: Theory and Research*. O. P. John, R. W. Robins and L. A. Pervin, eds. New York, NY, Guilford Press**:** 3-26.

Barnett, W. Steven, Jung, Kwanghee, Yarosz, Donald J., Thomas, Jessica, Hornbeck, Amy, Stechuk, Robert and Burns, Susan (2008). "Educational Effects of the Tools of the Mind Curriculum: A Randomized Trial." *Early Childhood Research Quarterly* **23**(3): 299-313.

Barnett, W. Steven, Yarosz, Donald J., Thomas, Jessica and Hornbeck, Amy (2006). "Educational Effectiveness of a Vygotskian Approach to Preschool Education: A Randomized Trial." *Rutgers, NJ: National Institute for Early Education Research*.

Báron, Juan D. and Cobb-Clark, Deborah A. (2010). "Are Young People's Educational Outcomes Linked to Their Sense of Control?" IZA Discussion Paper No. 4907

Barrick, Murray R. and Mount, Michael K. (1991). "The Big Five Personality Dimensions and Job Performance: A Meta-Analysis." *Personnel Psychology* **44**(1): 1-26.

Barsky, Robert B., Juster, F. Thomas, Kimball, Miles S. and Shapiro, Matthew D. (1997). "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics* **112**(2): 537.

Baumeister, Roy F., Bratslavsky, Ellen, Muraven, Mark and Tice, Dianne M. (1998). "Ego Depletion: Is the Active Self a Limited Resource?" *Journal of Personality and Social Psychology* **74**(5): 1252.

Bechara, Antoine (2005). "Decision Making, Impulse Control and Loss of Willpower to Resist Drugs: A Neurocognitive Perspective." *Nature Neuroscience* **8**(11): 1458-1463.

Bechara, Antoine and Damasio, Antonio R. (2005). "The Somatic Marker Hypothesis: A Neural Theory of Economic Decision." *Games and Economic Behavior* **52**(2): 336-372.

Beckett, Celia, Maughan, Barbara, Rutter, Michael, Castle, Jenny, Colvert, Emma, Groothues, Christine, Kreppner, Jana, Stevens, Suzanne, O'Connor, Thomas G. and Sonuga-Barke, Edmund J. S. (2006). "Do the Effects of Early Severe Deprivation on Cognition Persist into Early Adolescence? Findings from the English and Romanian Adoptees Study." *Child development* **77**(3): 696.

Behncke, Stefanie (2009). "How Do Shocks to Non-Cognitive Skills Affect Test Scores?" *IZA Discussion Paper No. 4222*.

Benabou, Roland and Tirole, Jean (2003). "Intrinsic and Extrinsic Motivation." *Review of Economic Studies* **70**: 489-520.

Benjamin, Daniel J., Brown, Sebastian A. and Shapiro, Jesse M. (2006). "Who Is `Behavioral'? Cognitive Ability and Anomalous Preferences."

Benz, Matthias and Meier, Stephan (2008). "Do People Behave in Experiments as in the Field?—Evidence from Donations." *Experimental Economics* **11**: 268-281.

Bergen, Sarah E., Gardner, Charles O. and Kendler, Kenneth S. (2007). "Age-Related Changes in Heritability of Behavioral Phenotypes over Adolescence and Young Adulthood: A Meta-Analysis." *Twin Research and Human Genetics* **10**(3): 423.

Bickel, Warren K., Odum, Amy L. and Madden, Gregory J. (1999). "Impulsivity and Cigarette Smoking: Delay Discounting in Current, Never, and Ex-Smokers." *Psychopharmacology* **146**(4): 447.

Bierman, Karen L., Coie, John D., Dodge, Kenneth A., Greenberg, Mark T., Lochman, John E., McMahon, Robert J. and Pinderhughes, Ellen (2010). "The Effects of a Multiyear Universal Social-Emotional Learning Program: The Role of Student and School Characteristics." *Journal of Consulting and Clinical Psychology* **78**(2): 156-168.

Binet, Alfred, Simon, Theodore and Kite, Elizabeth S. (1916). *The Development of Intelligence in Children (the Binet-Simon Scale)*. Baltimore, MD, Williams & Wilkins Co.

Blair, Clancy (2006). "How Similar Are Fluid Cognition and General Intelligence? A Developmental Neuroscience Perspective on Fluid Cognition as an Aspect of Human Cognitive Ability." *Behavioral and Brain Sciences* **29**(2): 109.

Blair, Clancy, Gamson, David, Thorne, Steven and Baker, David (2005). "Rising Mean IQ: Cognitive Demand of Mathematics Education for Young Children, Population Exposure to Formal Schooling, and the Neurobiology of the Prefrontal Cortex." *Intelligence* **33**(1): 93-106.

Blair, Clancy and Razza, Rachel Peters (2007). "Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten." *Child Development* **78**(2): 647-663.

Blais, AR and Weber, EU (2006). "A Domain-Specific Risk-Taking (Dospert) Scale for Adult Populations." *Judgment and Decision Making* **1**(1): 33-47.

Block, Jack (1995). "A Contrarian View of the Five-Factor Approach to Personality Description." *Psychological Bulletin* **117**(2): 187-215.

Blonigen, Daniel M., Hicks, Brian M., Krueger, Robert F., Patrick, Christopher J. and Iacono, William G. (2006). "Continuity and Change in Psychopathic Traits as Measured Via Normal-Range Personality: A Longitudinal-Biometric Study." *Journal of Abnormal Psychology* **115**(1): 85.

Bloom, Dan, Gardenhire-Crooks, Alissa and Mandsager, Conrad L. (2009). "Reengaging High School Dropouts: Early Results of the National Guard Youth Challenge Program Evaluation." *MDRC Report, New York*.

Boake, Corwin (2002). "From the Binet-Simon to the Wechsler-Bellevue: Tracing the History of Intelligence Testing." *Journal of Clinical and Experimental Neuropsychology* **24**(3): 383.

Bodrova, Elena and Leong, Deborah J. (2001). *The Tools of the Mind Project: A Case Study of Implementing the Vygotskian Approach in American Early Childhood and Primary Classrooms*. Geneva, International Bureau of Education, UNESCO.

Bodrova, Elena and Leong, Deborah J. (2007). *Tools of the Mind: The Vygotskian Approach to Early Childhood Education*. Upper Saddle River, Pearson Education, Inc.

Bollen, Kenneth A. (1989). *Structural Equations with Latent Variables*. New York, Wiley.

Bonato, Daniela P. and Boland, Fred J. (1983). "Delay of Gratification in Obese Children." *Addictive Behaviors* **8**(1): 71.

Booth, Alison L. and Nolen, Patrick J. (2009). "Gender Differences in Risk Behaviour: Does Nurture Matter?" *IzA Discussion Paper No. 4026*.

Borghans, Lex , Golsteyn, Bart H. H., Heckman, James J. and Meijers, Huub (2009). "Gender Differences in Risk Aversion and Ambiguity Aversion." *Journal of the European Economic Association* **7**(2-3): 649-658.

Borghans, Lex, Duckworth, Angela L., Heckman, James J. and ter Weel, Bas (2008). "The Economics and Psychology of Personality Traits." *Journal of Human Resources* **43**(4): 972-1059.

Borghans, Lex, Golsteyn, Bart H. H., Heckman, James J. and Humphries, John Eric (2010). "IQ, Achievement, and Personality." Unpublished manuscript. University of Chicago, Department of Economics.

Borghans, Lex, Meijers, Huub and ter Weel, Bas (2008). "The Role of Noncognitive Skills in Explaining Cognitive Test Scores." *Economic Inquiry* **46**(1): 2-12.

Borkenau, P, Mauer, N, Riemann, R, Spinath, FM and Angleitner, A (2004). "Thin Slices of Behavior as Cues of Personality and Intelligence." *Journal of Personality and Social Psychology* **86**(4): 599-614.

Borkenau, P, Riemann, R, Angleitner, A and Spinath, FM (2001). "Genetic and Environmental Influences on Observed Personality: Evidence from the German Observational Study of Adult Twins." *Journal of Personality and Social Psychology* **80**(4): 655-668.

Borsboom, D., Mellenbergh, G. J. and Van Heerden, J. (2003). "The Theoretical Status of Latent Variables." *Psychological Review* **110**: 203-219.

Bouchard, Thomas J. and Loehlin, John C. (2001). "Genes, Evolution and Personality." *Behavior Genetics* **31**(3): 243-273.

Bowen, William G., Chingos, Matthew M. and McPherson, Michael S. (2009a). "Educational Attainment: Overall Trends, Disparities, and the Public Universities We Study." in *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton, NJ, Princeton University Press**:** 1-19.

Bowen, William G., Chingos, Matthew M. and McPherson, Michael S. (2009b). "Test Scores and High School Grades as Predictors." in *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton, NJ, Princeton University Press**:** 112-133.

Bowles, Samuel and Gintis, Herbert (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York, Basic Books.

Bowles, Samuel, Gintis, Herbert and Osborne, Melissa (2001a). "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature* **39**(4): 1137-1176.

Bowles, Samuel, Gintis, Herbert and Osborne, Melissa (2001b). "Incentive-Enhancing Preferences: Personality, Behavior, and Earnings." *American Economic Review* **91**(2): 155-158.

Bowman, Philip C. and Matthews, Charles V. (1960). "Motivations of Youth for Leaving School." University of Chicago Development Project

Boyle, Stephen H., Williams, Redford B., Mark, Daniel B., Brummett, Beverly H., Siegler, Ilene C. and Barefoot, John C. (2005). "Hostility, Age, and Mortality in a Sample of Cardiac Patients." *American Journal of Cardiology* **96**: 64-66.

Breuning, Stephen E. and Zella, William F. (1978). "Effects of Individualized Incentives on Norm-Referenced IQ Test Performance of High School Students in Special Education Classes." *Journal of School Psychology* **16**(3): 220.

Bridgman, Percy W. (1959). *The Way Things Are*. Cambridge, Harvard University Press.

Browning, Martin, Hansen, Lars Peter and Heckman, James J. (1999). "Micro Data and General Equilibrium Models." in *Handbook of Macroeconomics*. J. B. Taylor and M. Woodford, eds., Elsevier. **1A:** 543-633.

Buck, R. Creighton (2003). *Advanced Calculus*. Long Grove, IL, Waveland Press, Inc.

Burks, Stephen V., Carpenter, Jeffrey, Goette, Lorenz and Rustichini, Aldo (2010). "Overconfidence Is a Social Signaling Bias." *IzA Discussion Paper No. 4840*.

Burks, Stephen V., Carpenter, Jeffrey P., Goette, Lorenz and Rustichini, Aldo (2009). "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." *Proceedings of the National Academy of Sciences* **6**(109): 7745-7750.

Buss, David M. (2000). *The Dangerous Passion: Why Jealousy Is as Necessary as Love and Sex*. New York, Free Press.

Cahan, Sorel and Cohen, Nora (1989). "Age Versus Schooling Effects on Intelligence Development." *Child development* **60**(5): 1239.

Caliendo, Marco, Cobb-Clark, Deborah and Uhlendorff, Arne (2010). "Locus of Control and Job Search Strategies." IZA Discussion Paper No. 4750.

Caliendo, Marco, Fossen, Frank and Kritikos, Alexander (2008). "The Impact of Risk Attitudes on Entrepreneurial Survival." IZA Discussion Paper No. 3525.

Camerer, C F and Fehr, Ernst (2004). "Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists." in *Foundations of Human Sociality: Experimental and Ethnographic Evidence from 15 Small-Scale Societies*. J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr and H. Gintis, eds. New York, NY, Oxford University Press**:** 55-95.

Canli, Turhan (2004). "Functional Brain Mapping of Extraversion and Neuroticism: Learning from Individual Differences in Emotion Processing." *Journal of Personality* **72**(6): 1105-1132.

Canli, Turhan (2006). *Biology of Personality and Individual Differences*. New York, Guilford Press.

Canli, Turhan (2008). "Toward A" Molecular Psychology" Of Personality." in *Handbook of Personality: Theory and Research*. O. P. a. R. John, Richard W. and Pervin, Lawrence A., ed. New York, NY, The Guilford Press**:** 311-327.

Caplin, A and Leahy, J (2001). "Psychological Expected Utility Theory and Anticipatory Feelings*." *Quarterly Journal of economics* **116**(1): 55-79.

Carneiro, Pedro, Heckman, James J. and Vytlacil, Edward J. (2005). "Understanding What Instrumental Variables Estimate: Estimating Marginal and Average Returns to Education.

Carpenter, J, Verhoogen, E and Burks, S (2005). "The Effect of Stakes in Distribution Experiments." *Economics Letters* **86**: 393-398.

Carpenter, Patricia A., Just, Marcel A. and Shell, Peter (1990). "What One Intelligence Test Measures: A Theoretical Account of the Processing in the Raven Progressive Matrices Test." *Psychological Review* **97**(3): 404-431.

Carroll, John B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. New York, Cambridge University Press.

Casey, B.J., Tottenham, N. , Liston, C. and Durston, S. (2005). "Imaging the Developing Brain: What Have We Learned About Cognitive Development? ." *Trends in Cognitive Science* **9**(104-110).

Caspi, Avshalom (2000). "The Child Is Father of the Man: Personality Continuities from Childhood to Adulthood." *Journal of Personality and Social Psychology* **78**(1): 158.

Caspi, Avshalom, McClay, Joseph, Moffitt, Terrie E., Mill, Jonathan, Martin, Judy, Craig, Ian W., Taylor, Alan and Poulton, Richie (2002). "Role of Genotype in the Cycle of Violence in Maltreated Children." *Science* **297**(5582): 851-854.

Caspi, Avshalom, Moffit, Terrie E., Silva, Phil A., Stouthamer-Loeber, Magda, Krueger, Robert F. and Schmutte, Pamela S. (1994). "Are Some People Crime-Prone? Replications of the Personality-Crime Relationship across Countries, Genders, Races, and Methods." *Criminology* **32**(2): 163-195.

Caspi, Avshalom and Shiner, Rebecca L. (2006). "Personality Development." in *Handbook of Child Psychology*. W. Damon and R. M. Lerner, eds. Hoboken, NJ, John Wiley & Sons, Inc. **3:** 300-365.

Caspi, Avshalom, Sugden, Karen, Moffitt, Terrie E., Taylor, Alan, Craig, Ian W., Harrington, HonaLee, McClay, Joseph, Mill, Jonathan, Martin, Judy, Braithwaite, Antony and Poulton, Richie (2003). "Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-Htt Gene." *Science* **301**(5631): 386-389.

Cattan, Sarah (2010). "Heterogeneity and Selection in the Labor Market." PhD Thesis, Economics Department, University of Chicago.

Cattell, James McKeen (1890). "Mental Tests and Measurements." *Mind* **15**(59): 373.

Cattell, Raymond B. (1971). *Abilities: Their Structure, Growth, and Action*. Boston, Houghton Mifflin.

Cattell, Raymond B. (1971/1987). *Intelligence: Its Structure, Growth and Action*. Amsterdam, Elsevier.

Cawley, John, Heckman, James J. and Vytlacil, Edward J. (1999). "On Policies to Reward the Value Added by Educators." *Review of Economics and Statistics* **81**(4): 720-727.

Cebi, Merve (2007). "Locus of Control and Human Capital Investment Revisited." *Journal of Human Resources* **42**(4): 919-932.

Cervone, D and Pervin, LA (2009). *Personality: Theory and Research* Hoboken, NJ, John Wiley and Sons, Inc.

Chamorro-Premuzic, Tomas and Furnham, Adrian (2005). *Personality and Intellectual Competence*. Mahwah, N.J, L. Erlbaum Associates.

Chao, Li-Wei, Szrek, Helena, Sousa Pereira, Nuno and Pauly, Mark V. (2007). "Time Preference and Its Relationship with Age, Health, and Longevity Expectations." Universidade do Porto Faculdade de Economia do Porto 0706.

Chapman, Gretchen B. and Coups, Elliot J. (1999). "Time Preferences and Preventive Health Behavior: Acceptance of the Influenza Vaccine." *Medical Decision Making* **19**(3): 307.

Chapman, Gretchen B., Nelson, Richard and Hier, Daniel B. (1999). "Familiarity and Time Preferences: Decision Making About Treatments for Migraine Headaches and Crohn's Disease." *Journal of Experimental Psychology: Applied* **5**(1): 17.

Chesson, Harrell and Viscusi, W. Kip (2000). "The Heterogeneity of Time-Risk Tradeoffs." *Journal of Behavioral Decision Making* **13**(2): 251.

Chetty, Raj, Friedman, John N., Hilger, Nathaniel, Diane, Schanzenbach Whitmore and Yagan, Danny (2010). "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star." *NBER Working Paper*.

Clausen, John A. and Gilens, Martin (1990). "Personality and Labor Force Participation across the Life Course: A Longitudinal Study of Women's Careers." *Sociological Forum* **5**(4): 595.

Clingman, Joy and Fowler, Robert L. (1976). "The Effects of Primary Reward on the I.Q. Performance of Grade-School Children as a Function of Initial I.Q. Level." *Journal of applied behavior analysis* **9**(1): 19.

Cloninger, C. Robert (1987). "A Systematic Method for Clinical Description and Classification of Personality Variants: A Proposal." *Archives of General Psychiatry* **44**(6): 573.

Cloninger, C. Robert, Svrakic, Dragan M., Bayon, Carmen and Przybeck, Thomas R. (1999). "Measurement of Psychopathology as Variants of Personality." in *Personality and Psychopathology*. C. R. Cloninger, ed. American Psychiatric Press.

Cobb-Clark, Deborah and Tan, Michelle (2009). "Noncognitive Skills, Occupational Attainment, and Relative Wages." IZA Discussion Paper No. 4289.

Coleman, Margo and DeLeire, Thomas (2003). "An Economic Model of Locus of Control and the Human Capital Investment Decision." *Journal of Human Resources* **38**(3): 701-721.

Coller, Maribeth and Williams, Melonie B. (1999). "Eliciting Individual Discount Rates." *Experimental Economics* **2**(2): 107.

Conard, Maureen A. (2005). "Aptitude Is Not Enough: How Personality and Behavior Predict Academic Performance." *Journal of Research in Personality* **40**(2006): 339-346.

Conti, Gabriella and Heckman, James J. (2010). "Understanding the Early Origins of the Education-Health Gradient: A Framework That Can Also Be Applied to Analyze Gene-Environment Interactions." *Perspectives on Psychological Science* **5**(5): 585-605.

Conti, Gabriella, Heckman, James J., Lopes, Hedibert F. and Piatek, Remi (2010). "Constructing Economically Justified Aggregates: An Application of the Early Origins of Health." Unpublished manuscript.

Conti, Gabriella, Heckman, James J. and Urzua, Sergio (2010a). "Early Endowments, Education, and Health."

Conti, Gabriella, Heckman, James J. and Urzua, Sergio (2010b). "The Education-Health Gradient." *American Economic Review: Papers and Proceedings* **100**(2): 234-238.

Conti, Gabriella, Heckman, James J., Yi, J. and Zhang, J. (2010). "Early Health Shocks, Parental Responses, and Child Outcomes." Unpublished manuscript, Department of Economics, University of Hong Kong.

Conti, Gabriella and Pudney, Stephen (2007). "The Dynamics of Cognitive Development."

Costa, Paul T. and McCrae, Robert R. (1992a). "Four Ways Five Factors Are Basic." *Personality and Individual Differences* **13**(6): 653.

Costa, Paul T. and McCrae, Robert R. (1992b). *Revised Neo Personality Inventory (Neo Pi-R) and Neo Five-Factor Inventory (Neo-Ffi)*. Odessa, FL, Psychological Assessment Resources.

Costa, Paul T. and McCrae, Robert R. (1994). "Set Like Plaster? Evidence for the Stability of Adult Personality." in *Can Personality Change?* T. F. Heatherton and J. L. Weinberger, eds. Washington, DC, American Psychological Association**:** 21–40.

Costa, Paul T., McCrae, Robert R. and Siegler, Ilene C. (1999). "Continuity and Change over the Adult Life Cycle: Personality and Personality Disorders." in *Personality and Psychopathology*. C. R. Cloninger, ed. Washington, D. C., American Psychiatric Press.

Cragg, John G. and Donald, Stephen G. (1997). "Inferring the Rank of a Matrix." *Journal of Econometrics* **76**(1-2): 223-250.

Credé, Marcus and Kuncel, Nathan R. (2008). "Study Habits, Skills, and Attitudes: The Third Pillar Supporting Collegiate Academic Performance." *Perspectives on Psychological Science* **3**(6): 425-453.

Csikszentmihalyi, Mihaly (1996). *Creativity: Flow and the Psychology of Discovery and Invention*. New York, Harper Collins Publishers.

Cudeck, Robert and MacCullum, Robert (2007). *Factor Analysis at 100: Historical Developments and Future Directions*. Mahwah, NJ, Lawrence Erlibaum Associates.

Cunha, Flavio and Heckman, James J. (2007). "The Technology of Skill Formation." *American Economic Review* **97**(2): 31-47.

Cunha, Flavio and Heckman, James J. (2008). "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Journal of Human Resources* **43**(4): 738-782.

Cunha, Flavio and Heckman, James J. (2009). "The Economics and Psychology of Inequality and Human Development." *Journal of the European Economic Association* **7**(2-3): 320-364.

Cunha, Flavio, Heckman, James J., Lochner, Lance J. and Masterov, Dimitriy V. (2006). "Interpreting the Evidence on Life Cycle Skill Formation." in *Handbook of the Economics of Education*. E. A. Hanushek and F. Welch, eds. Amsterdam, North-Holland**:** 697-812.

Cunha, Flavio, Heckman, James J. and Navarro, Salvador (2005). "Separating Uncertainty from Heterogeneity in Life Cycle Earnings, the 2004 Hicks Lecture." *Oxford Economic Papers* **57**(2): 191-261.

Cunha, Flavio, Heckman, James J. and Schennach, Susanne M. (2010). "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* **78**(3): 883-931.

Daly, Michael, Delaney, Liam and Harmon, Colm P. (2009). "Psychological and Biological Foundations of Time Preferences." *Journal of the European Economic Association* **7**(2-3): 659-669.

Damasio, Antonio R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, Putnam.

Damasio, Hanna, Grabowski, Thomas, Frank, Randall, Galaburda, Albert M. and Damasio, Antonio R. (2005). "The Return of Phineas Gage: Clues About the Brain from the Skull of a Famous Patient." in *Social Neuroscience: Key Readings. Key Reading in Social Psychology.* J. T. Cacioppo and G. G. Berntson, eds. New York, NY, US, Psychology Press**:** 21-28.

de Wit, Harriet, Flory, Janine D., Acheson, Ashley, McCloskey, Michael and Manuck, Stephen B. (2007). "IQ and Nonplanning Impulsivity Are Independently Associated with Delay Discounting in Middle-Aged Adults." *Personality and Individual Differences* **42**(1): 111.

Deci, Edward L. and Ryan, Richard M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York, Plenum Press.

Dee, Thomas and West, Martin (2008). "The Non-Cognitive Returns to Class Size." *NBER Working Paper*.

DeLong, J. Bradford and Magin, Konstantin (2009). "The U.S. Equity Return Premium: Past, Present and Future." *Journal of Economic Perspectives* **23**(1): 193–208.

Devlin, Bernie, Daniels, Michael and Roeder, Kathryn (1997). "The Heritability of IQ." *Nature* **388**(6641): 468.

DeYoung, Colin G. (2007). "Between Facets and Domains: Ten Aspects of the Big Five." *Journal of Personality and Social Psychology*.

DeYoung, Colin G., Hirsh, Jacob B., Shane, Matthew S., Papademetris, Xenophon, Rajeevan, Nallakkandi and Gray, Jeremy R. (2010). "Testing Predictions from Personality Neuroscience." *Psychological Science* **21**(6): 820-828.

Diamond, Adele, Barnett, Steven, Thomas, Jessica and Munro, Sarah (2007). "Preschool Program Improves Cognitive Control." *Science* **318**: 1387-1388.

Dickens, William T. and Flynn, James R. (2001). "Heritability Estimates Versus Large Environmental Effects: The IQ Paradox Resolved." *Psychological Review* **108**(2): 346.

Digman, John M. (1997). "Higher-Order Factors of the Big Five." *Journal of Personality and Social Psychology* **73**(6): 1246-1256.

Ding, Xiaohao, Hartog, Joop and Sun, Yuze (2010). "Can We Measure Individual Risk Attitudes in a Survey?" IZA Discussion Paper, 4807.

Dohmen, Thomas, Falk, Armin, Huffman, David, Sunde, U., Schupp, J. and Wagner, G.G. (2009a). "Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences." Forthcoming, *Journal of the European Economic Association*.

Dohmen, Thomas, Falk, Armin, Huffman, David and Sunde, Uwe (2007). "Are Risk Aversion and Impatience Related to Cognitive Ability?" IZA Discussion Papers 2735. Bonn, Germany: Institute for the Study of Labor. Discussion Paper, 2735.

Dohmen, Thomas, Falk, Armin, Huffman, David and Sunde, Uwe (2008). "Representative Trust and Reciprocity: Prevalence and Determinants." *Economic Inquiry* **46**(1): 84-90.

Dohmen, Thomas, Falk, Armin, Huffman, David and Sunde, Uwe (2009b). "Homo Reciprocans: Survey Evidence on Behavioural Outcomes." *The Economic Journal* **119**(536): 592-612.

Drago, Francesco (2008). "Self-Esteem and Earnings." IZA Discussion Paper, 3577.

Du, Wanjiang, Green, Leonard and Myerson, Joel (2002). "Cross-Cultural Comparisons of Discounting Delayed and Probabilistic Rewards." *Psychological Record* **52**(4): 479.

Duckworth, Angela L. (2007). "Unpublished Dataset." University of Pennsylvania, Department of Psychology.

Duckworth, Angela L. (2009). "(over and) Beyond High-Stakes Testing." *American Psychologist* **64**(4): 279-280.

Duckworth, Angela L., Quinn, Patrick D. and Tsukayama, Eli (2010). "What No Child Left Behind Leaves Behind: The Roles of IQ and Self-Control in Predicting Standardized Achievement Test Scores and Report Card Grades." Unpublished manuscript. University of Pennsylvania, Department of Psychology.

Duckworth, Angela L. and Schulze, R. (2009). "Jingle Jangle: A Meta-Analytic Investigation of the Convergent Validity of Self-Control." Manuscript. University of Pennsylvania, Department of Psychology.

Duckworth, Angela L. and Seligman, Martin E. P. (2006). "Self-Discipline Gives Girls the Edge: Gender in Self-Discipline, Grades, and Achievement Scores." *Journal of Educational Psychology* **98**(1): 198-208.

Duckworth, Angela Lee and Seligman, Martin E. P. (2005). "Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents." *Psychological Science* **16**(12): 939-944.

Duckworth, Angela Lee, Tsukayama, Eli and May, Henry (in press). "Establishing Causality Using Longitudinal Hierarchical Linear Modeling: An Illustration Predicting Achievement from Self-Control." *Social Psychology and Personality Science*.

Duncan, Greg .J. and Magnuson, Katherine (2010). "The Nature and Impact of Early Achievement Skills, Attention Skills, and Behavior Problems." Unpublished Manuscript. University of California Irvine, Department of Education.

Duncan, Greg J., Dowsett, Chantelle J., Claessens, Amy, Mugnuson, Katherine, Huston, Aletha C., Klebanov, Pamela, Pagani, Linda S., Feinstein, Leon, Engel, Mimi, Brooks-Gunn, Jeanne, Sexton, Holly, Duckworth, Kathryn and Japel, Crista (2007). "School Readiness and Later Achievement." *Developmental Psychology* **43**(6): 1428-1446.

Duncan, Greg J. and Dunifon, Rachel (1998). ""Soft-Skills" And Long-Run Labor Market Success." in *Research in Labor Economics*. S. W. Polachek, ed. Stamford, CT, JAI Press. **17:** 123-149.

Dur, Robert, Non, Arjan and Roelfsema, Hein (2010). "Reciprocity and Incentive Pay in the Workplace." *Journal of Economic Psychology* **31**(4): 676-686.

Durlak, Joseph A. and Weissberg, Roger P. (in press). "The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions." *Child Development*.

Duyme, Michel, Dumaret, Annick-Camille and Tomkiewicz, Stanislaw (1999). "How Can We Boost IQs Of "Dull Children"? A Late Adoption Study." *Proceedings of the National Academy of Sciences* **96**(15): 8790.

Edlund, Calvin V. (1972). "The Effect on the Behavior of Children, as Reflected in the IQ Scores, When Reinforced after Each Correct Response." *Journal of Applied Behavior Analysis* **5**(3): 317-319.

Einav, Liran , Finkelstein, Amy , Pascu, Iuliana and Cullen, Mark R. (2010). "How General Are Risk Preferences? Choices under Uncertainty in Different Domains " *NBER Working Paper No. 15686*.

Epstein, Larry G. and Schneider, Martin (2003). "Recursive Multiple-Priors." *Journal of Economic Theory* **113**(1): 1-31.

Epstein, Larry G. and Zin, Stanley E. (1989). "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica: Journal of the Econometric Society* **57**(4): 937-969.

Epstein, Seymore (1979). "The Stability of Behavior: I. On Predicting Most of the People Much of the Time." *Journal of Personality and Social Psychology* **37**: 1097-1126.

Ericsson, K. Anders, Krampe, Ralf T. and Tesch-Romer, Clemens (1993). "The Role of Deliberate Practice in the Acquisition of Expert Performance." *Psychological Review* **100**(3): 363.

Ericsson, K. Anders and Ward, Paul (2007). "Capturing the Naturally Occurring Superior Performance of Experts in the Laboratory: Toward a Science of Expert and Exceptional Performance." *Current Directions in Psychological Science* **16**(6): 346.

Estle, Sara J., Green, Leonard, Myerson, Joel and Holt, Daniel D. (2007). "Discounting of Monetary and Directly Consumable Rewards." *Psychological Science* **18**(1): 58.

Eysenck, Hans Jurgen (1991). "Dimensions of Personality: 16, 5 or 3? Criteria for a Taxonomic Paradigm." *Personality and Individual Differences* **12**(8): 773.

Falk, Armin and Fischbacher, Urs (2006). "A Theory of Reciprocity." *Games and Economic Behavior* **54**(2): 293-315.

Falk, Armin and Heckman, James J (2009). "Lab Experiments Are a Major Source of Knowledge in the Social Sciences." *Science* **326**(5952): 535-538.

Fehr, Ernest (2009). "On the Economics and Biology of Trust." *Journal of the European Economic Association* **7**(2-3): 235-266.

Fehr, Ernst and Gächter, Simon (2000). "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* **14**(3): 159–181.

Fehr, Ernst and Schmidt, Klaus M. (1999). "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* **114**(3): 817-868.

Fehr, Ernst and Schmidt, Klaus M. (2006). "The Economics of Fairness, Reciprocity and Altruism–Experimental Evidence and New Theories." in *Handbook of the Economics of Giving, Altruism and Reciprocity*. S.-C. Kolm and J. M. Ythier, eds. Amsterdam, Elsevier. **1:** 615-691.

Feingold, A (1982). "The Validity of the Information and Vocabulary Subtests of the WAIS." *Journal of Clinical Psychology* **38**(1): 169-174.

Fergusson, David M. and Horwood, L. John (1998). "Early Conduct Problems and Later Life Opportunities." *Journal of Child Psychology and Psychiatry* **39**(8): 1097-1108.

Figner, Bernd, Knoch, Daria, Johnson, Eric J., Krosch, Amy R., Lisanby, Sarah H., Fehr, Ernst and Weber, Elke U. (2010). "Lateral Prefrontal Cortex and Self-Control in Intertemporal Choice." *Nat Neurosci* **13**(5): 538-539.

Finn, Jeremy D., Pannozzo, Gina M. and Voelkl, Kristin E. (1995). "Disruptive and Inattentive-Withdrawn Behavior and Achievement among Fourth Graders." *The Elementary School Journal* **95**(5): 421-434.

Fleeson, William (2001). "Toward a Structure- and Process-Integrated View of Personality: Traits as Density Distributions of States." *Journal of Personality and Social Psychology* **80**(6): 1011-1027.

Fleeson, William and Noftle, Erik (2008). "The End of the Person–Situation Debate: An Emerging Synthesis in the Answer to the Consistency Question." *Social and Personality Psychology Compass* **2**(4): 1667-1684.

Flinn, Christopher and Heckman, James J. (1982). "New Methods for Analyzing Structural Models of Labor Force Dynamics." *Journal of Econometrics* **18**(1): 115-168.

Flynn, James Robert (2007). *What Is Intelligence? Beyond the Flynn Effect*. New York, Cambridge University Press.

Foster, James E. and Sen, Amartya K. (1997). *On Economic Inequality*. New York, Oxford University Press.

Frederick, S, Loewenstein, G and O'Donoghue, T (2002). "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature* **40**(2): 351-401.

Frey, Meredith C. and Detterman, Douglas K. (2004). "Scholastic Assessment or g? The Relationship between the Scholastic Assessment Test and General Cognitive Ability." *Psychological Science* **15**(6): 373-378.

Friedman, Howard S. (2007). "Personality, Disease, and Self-Healing." in *Foundations of Health Psychology*. H. S. Friedman and R. C. Silvers, eds. New York, Oxford University Press**:** 172-199.

Friedman, Howard S., Kern, Margaret L. and Reynolds, Chandra A. (2010). "Personality and Health, Subjective Well-Being, and Longevity." *Journal of Personality* **78**: 179-215.

Friedman, Naomi P., Miyake, Akira, Corley, Robin P., Young, Susan E., DeFries, John C. and Hewitt, John K. (2006). "Not All Executive Functions Are Related to Intelligence." *Psychological Science* **17**(2): 172-179.

Funder, David C. (2008). "Persons, Situations, and Person-Situation Interactions." in *Handbook of Personality: Theory and Research*. O. P. John, R. W. Robins and L. A. Pervin, eds. New York, Guilford Press**:** 568-580.

Funder, David C. (2009). "Persons, Behaviors and Situations: An Agenda for Personality Psychology in the Postwar Era." *Journal of Research in Personality* **43**(2): 120-126.

Funder, David C. and Block, Jack (1989). "The Role of Ego-Control, Ego-Resiliency, and IQ in Delay of Gratification in Adolescence." *Journal of Personality and Social Psychology* **57**(6): 1041.

Gale, Catharine R., Batty, G. David and Deary, Ian J. (2008). "Locus of Control at Age 10 Years and Health Outcomes and Behaviors at Age 30 Years: The 1970 British Cohort Study." *Psychosomatic Medicine* **70**(4): 397-403.

Galton, Francis (1883). *Inquiries into Human Faculty and Its Development*. London, UK, Macmillan and Company.

Gardner, Howard (2004). *Frames of Mind: The Theory of Multiple Intelligences*.

Garon, Nancy, Bryson, Susan E. and Smith, Isabel M. (2008). "Executive Function in Preschoolers: A Review Using an Integrative Framework." *Psychological Bulletin* **134**(1): 31-60.

Geiser, Saul and Santelices, Maria Veronica (2007). "Validity of High School Grades in Predicting Student Success Beyond the Freshman Year: High-School Record Vs. Standardized Tests as Indicators of Four-Year College Outcomes." Center for Studies in Higher Education at the University of California, Berkeley CSHE.6.07.

Gilboa, I. and Schmeidler, D. (1989). "Maxmin Expected Utility with Non-Unique Prior." *Journal of Mathematical Economics* **18**: 141-153.

Gillham, Jane E. and Reivich, Karen J. (1999). "Prevention of Depressive Symptoms in School Children: A Research Update." *Psychological Science* **10**(5): 461.

Gneezy, Uri and List, John A (2006). "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments." *Econometrica* **74**: 1365-1384.

Goldberg, Lewis R., Sweeney, Dennis, Merenda, Peter F. and Hughes, John Edward, Jr. (1998). "Demographic Variables and Personality: The Effects of Gender, Age, Education, and Ethnic/Racial Status on Self-Descriptions of Personality Attributes." *Personality and Individual Differences* **24**(3): 393-403.

Goldberger, Arthur S. (1968). *Topics in Regression Analysis*. New York, Macmillan.

Goldberger, Arthur S. (1979). "Heritability." *Economica* **46**(184): 327-347.

Goldsmith, H. Hill, Buss, Arnold H., Plomin, Robert, Rothbart, Mary Klevjord, Thomas, Alexander, Chess, Stella, Hinde, Robert A. and McCall, Robert B. (1987). "Roundtable: What Is Temperament? Four Approaches." *Child development* **58**(2): 505.

Gorman, W. M. (1968). "The Structure of Utility Functions." *Review of Economic Studies* **35**(4): 367.

Gottfredson, L. S. (1998). "The General Intelligence Factor " *Scientific American Presents* **9**(4): 24-29.

Gottfredson, Linda S. (2002). "g: Highly General and Highly Practical." in *The General Factor of Intelligence: How General Is It?* R. J. Sternberg and E. L. Grigorenko, eds. Mahwah, NJ, Lawrence Erlbaum Associates Publishers**:** 331-380.

Gottfredson, Linda. S. (2008). "Of What Value Is Intelligence?" in *Wisc-Iv Applications for Clinical Assessment and Intervention*. A. Prifitera, D. Saklofske and L. G. Weiss, eds. Amsterdan, Elsevier.

Gottfredson, MR and Hirschi, T (1990). *A General Theory of Crime*. Palo Alto, CA, Stanford University Press.

Gottschalk, Peter (2005). "Can Work Alter Welfare Recipients' Beliefs?" *Journal of Policy Analysis and Management* **24**(3): 485-498.

Gough, H. G. and Heilbrun Jr., A. B. (1983). *The Adjective Check List Manual*. Palo Alto, CA, Consulting Psychologists Press.

Gough, Harrison G. (1964). "Graduation from High School as Predicted from the California Psychological Inventory." *Psychology in the Schools* **3**(3): 208-216.

Gray, Jeremy R. (2004). "Integration of Emotion and Cognitive Control." *Current Directions in Psychological Science* **13**(2): 46-48.

Green, D. R. (1974). *The Aptitude Achievement Distinction*. Monterey, McGraw-Hill.

Green, L, Fry, AF and Myerson, J (1994a). "Discounting of Delayed Rewards: A Life-Span Comparison." *Psychological Science-Cambridge* **5**: 33-33.

Green, Leonard, Fry, Astrid F. and Myerson, Joel (1994b). "Discounting of Delayed Rewards: A Life-Span Comparison." *Psychological Science* **5**(1): 33.

Green, Leonard and Myerson, Joel (2004). "A Discounting Framework for Choice with Delayed and Probabilistic Rewards." *Psychological Bulletin* **130**(5): 769.

Greene, AC, Sapp, GL and Chissom, B (1990). "Validation of the Stanford-Binet Intelligence Scale: With Exceptional Black Male Students." *Psychology in the Schools* **27**(1): 35-41.

Gul, Faruk and Pesendorfer, Wolfgang (2004). "Self-Control and the Theory of Consumption." *Econometrica* **72**(1): 119-158.

Gunderson, John G., Triebwasser, Joseph, Phillips, Katharine A. and Sullivan, Chantal N. (1999). "Personality and Vulnerability to Affective Disorders." in *Personality and Psychopathology*. C. R. Cloninger, ed. Washington, D.C., American Psychiatric Press.

Gustafsson, Jan-Eric (1988). "Hierarchical Models of Individual Differences in Cognitive Abilities." in *Advances in the Psychology of Human Intelligence*. R. J. Sternberg, ed. Hillsdale, NJ, Erlbaum. **4**.

Hampson, Sarah E. and Friedman, Howard S. (2008). "Personality and Health: A Lifespan Perspective." in *The Handbook of Personality: Theory and Research*. O. P. John, R. Robins and L. Pervin, eds. New York, Guilford**:** 770-794.

Hampson, Sarah E., Goldberg, Lewis R., Vogt, Thomas M. and Dubanoski, Joan P. (2007). "Mechanisms by Which Childhood Personality Traits Influence Adult Health Status: Educational Attainment and Healthy Behaviors." *Health Psychology* **26**(1): 121-125.

Hampson, Sarah E., Tildesley, Elizabeth, Andrews, Judy A., Luyckx, Koen and Mroczek, Daniel K. (2010). "The Relation of Change in Hostility and Sociability During Childhood to Substance Use in Mid Adolescence." *Journal of Research in Personality* **44**(1): 103-114.

Hanoch, Yaniv, Johnson, Joseph G. and Wilke, Andreas (2006). "Domain Specificity in Experimental Measures and Participant Recruitment: An Application to Risk-Taking Behavior." *Psychological Science* **17**(4): 300.

Hansen, Karsten T., Heckman, James J. and Mullen, Kathleen J. (2004). "The Effect of Schooling and Ability on Achievement Test Scores." *Journal of Econometrics* **121**(1-2): 39-98.

Hansen, Lars Peter (2005). "Comment On "Exotic Preferences for Macroeconomics" By D. K. Backus, B. R. Routledge and S. E. Zin." in *Nber Macroeconomics Annual*. M. Gertler and K. Rogoff, eds. Cambridge, MA, MIT Press**:** 391.

Hansen, Lars Peter and Sargent, Thomas J. (2008). *Robustness*. Princeton, NJ, Princeton University Press.

Harrison, Glenn W., Lau, Morten Igel and Williams, Melonie B. (2002). "Estimating Individual Discount Rates in Denmark: A Field Experiment." *American Economic Review* **92**(5): 1606.

Hartigan, John A. and Wigdor, Alexandra K. (1989). "Fairness in Employment Testing: Validity Generalization, Minority Issues, and the General Aptitude Test Battery." *Washington, DC, US: National Academy Press* **354**.

Hartlage, LC and Steele, CT (1977). "Wisc and Wisc-R Correlates of Academic Achievement." *Psychology in the Schools* **14**(1): 15-18.

Hartog, Joop (1980). "Earnings and Capability Requirements." *Review of Economics and Statistics* **62**(2): 230-240.

Hartog, Joop (2001). "On Human Capital and Individual Capabilities." *Review of Income and Wealth* **47**(4): 515-540.

Hathaway, Starke R., Reynolds, Phyllis C. and Monachesi, Elio D. (1969). "Follow-up of the Later Careers and Lives of 1,000 Boys Who Dropped out of High School." *Journal of Consulting and Clinical Psychology* **33**(3): 370-380.

Heckman, James J. (1976). "A Life-Cycle Model of Earnings, Learning, and Consumption." *Journal of Political Economy* **84**(4, Part 2): S11-S44.

Heckman, James J. (2007). "The Economics, Technology and Neuroscience of Human Capability Formation." *Proceedings of the National Academy of Sciences* **104**(3): 13250-13255.

Heckman, James J. (2008). "Schools, Skills and Synapses." *Economic Inquiry* **46**(3): 289-324.

Heckman, James J. and Honoré, Bo E. (1990). "The Empirical Content of the Roy Model." *Econometrica* **58**(5): 1121-1149.

Heckman, James J. and Humphries, John Eric (2010). "The Predictive Validity of IQ and Personality in Later Life." Department of Economics, University of Chicago.

Heckman, James J., Humphries, John Eric, Urzua, Sergio and Veramendi, Gregory (2010). "The Effects of Schooling on Labor Market and Health Outcomes." Unpublished Manuscript. University of Chicago, Department of Economics.

Heckman, James J. and LaFontaine, Paul A. (2010). "The American High School Graduation Rate: Trends and Levels." *Review of Economics and Statistics* **92**(2): 244-262.

Heckman, James J., Malofeeva, Lena, Pinto, Rodrigo and Savelyev, Peter A. (2010). "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." Unpublished manuscript. Department of Economics, University of Chicago.

Heckman, James J. and Masterov, Dimitriy V. (2007). "The Productivity Argument for Investing in Young Children." *Review of Agricultural Economics* **29**(3): 446-493.

Heckman, James J., Moon, Seong Hyeok, Pinto, Rodrigo, Savelyev, Peter A. and Yavitz, Adam Q. (2010a). "Analyzing Social Experiments as Implemented: A Reexamination of the

Evidence from the Highscope Perry Preschool Program." *Quantitative Economics* **1**: 1-46.

Heckman, James J., Moon, Seong Hyeok, Pinto, Rodrigo, Savelyev, Peter A. and Yavitz, Adam Q. (2010b). "The Rate of Return to the Highscope Perry Preschool Program." *Journal of Public Economics* **94**(1-2): 114-128.

Heckman, James J. and Navarro, Salvador (2007). "Dynamic Discrete Choice and Dynamic Treatment Effects." *Journal of Econometrics* **136**(2): 341-396.

Heckman, James J. and Rubinstein, Yona (2001). "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review* **91**(2): 145-149.

Heckman, James J. and Sedlacek, Guilherme L. (1985). "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market." *Journal of Political Economy* **93**(6): 1077-1125.

Heckman, James J., Stixrud, Jora and Urzua, Sergio (2006). "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* **24**(3): 411-482.

Heitz, Richard P., Unsworth, Nash and Engle, Randall W. (2005). "Working Memory Capacity, Attention Control, and Fluid Intelligence." in *Handbook of Understanding and Measuring Intelligence*. O. Wilhelm and R. W. Engle, eds. Thousand Oaks, CA, Sage Publications Inc.**:** 61.

Helson, Ravenna, Kwan, Virginia S. Y., John, Oliver P. and Jones, Constance (2002). "The Growing Evidence for Personality Change in Adulthood: Findings from Research with Personality Inventories." *Journal of Research in Personality* **36**(4): 287.

Helson, Ravenna and Picano, James (1990). "Is the Traditional Role Bad for Women?" *Journal of Personality and Social Psychology* **59**(2): 311.

Hembree, Ray (1988). "Correlates, Causes, Effects, and Treatment of Test Anxiety." *Review of Educational Research* **58**(1): 47-77.

Hennig-Schmidt, H, Rockenbach, Bettina and A (2010). "In Search of Workers' Real Effort Reciprocity-a Field and a Laboratory Experiment." *Journal of the European Economic Association* **49**.

Herrnstein, Richard J. and Murray, Charles A. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York, Free Press.

Hershey, J. C., Kunreuther, H. C. and Schoemaker, Paul J. H. (1982). "Sources of Bias in Assessment Procedures for Utility Functions." *Management Science* **28** (8): 936-954.

Hicks, John Richard (1946). *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory*. Oxford, Clarendon Press.

Hoffman, Moshe, Fessler, Dan, Gneezy, Uri and List, John (2010). "Height and Competitiveness." Unpublished manuscript, University of Chicago.

Hofstee, Willem K., de Raad, Boele and Goldberg, Lewis R. (1992). "Integration of the Big Five and Circumplex Approaches to Trait Structure." *Journal of Personality and Social Psychology* **63**(1): 146.

Hogan, Joyce and Holland, Brent (2003). "Using Theory to Evaluate Personality and Job-Performance Relations: A Socioanalytic Perspective." *Journal of Applied Psychology* **88**(1): 100-112.

Hogan, Robert and Hogan, Joyce (2007). *Hogan Personality Inventory Manual*. Tulsa, OK, Hogan Assessment Systems.

Hogan, Robert, Hogan, Joyce and Roberts, Brent W. (1996). "Personality Measurement and Employment Decisions: Questions and Answers." *American Psychologist* **51**(5): 469.

Hogan, Robert and Roberts, Brent W. (2001). *Personality Psychology in the Workplace*, American Psychological Association.

Holmlund, Helena and Silva, Olmo (2009). "Targeting Non-Cognitive Skills to Improve Cognitive Outcomes: Evidence from a Remedial Education Intervention." *IZA Discussion Papers*.

Holt, Michael M. and Hobbs, Tom R. (1979). "The Effects of Token Reinforcement, Feedback and Response Cost on Standardized Test Performance." *Behaviour research and therapy* **17**(1): 81.

Hopkins, Kenneth D. and Bracht, Glenn H. (1975). "Ten-Year Stability of Verbal and Nonverbal IQ Scores." *American Educational Research Journal* **12**(4): 469-477.

Horn, John L. (1970). "Organization of Data on Life-Span Development of Human Abilities." in *Life-Span Developmental Psychology: Research and Theory*. L. R. Goulet and P. B. Baltes, eds. New York, Academic Press**:** 423.

Horn, John L. and McArdle, John J. (2007). "Understanding Human Intelligence in Factor Analysis since Spearman." in *Factor Analysis at 100: Historical Developments and Future Directions*. R. Cudek and R. C. MacCallum, eds. Mahwah, NJ, Lawrence Erlbaum Associates.

Horvath, Paula and Zuckerman, Marvin (1993). "Sensation Seeking, Risk Appraisal, and Risky Behavior." *Personality and Individual Differences* **14**: 41-52.

Hough, Leaetta M. (1992). "The "Big Five" Personality Variables--Construct Confusion: Description Versus Prediction." *Human Performance* **5**(1-2): 139-155.

Hough, Leaetta M., Eaton, Newell K., Dunnette, Marvin D., Kamp, John D. and McCloy, Rodney A. (1990). "Criterion-Related Validities of Personality Constructs and the Effect of Response Distortion on Those Validities." *Journal of Applied Psychology* **75**(5): 581.

Hough, Leaetta M. and Ones, Deniz S. (2002). "The Structure, Measurement, Validity, and Use of Personality Variables in Industrial, Work, and Organizational Psychology." in *Handbook of Industrial, Work and Organizational Psychology, Volume 1: Personnel Psychology*. N. Anderson, D. S. Ones, H. K. Sinangil and C. Viswesvaran, eds. Thousand Oaks, CA, Sage Publications Inc.**:** 233.

Hough, Leaetta M. and Oswald, Frederick L. (2000). "Personnel Selection: Looking toward the Future--Remembering the Past." *Annual Review of Psychology* **51**: 631-664.

Hunter, JE (1986). "Cognitive Ability, Cognitive Aptitudes, Job Knowledge, and Job Performance." *Journal of vocational behavior* **29**(3): 340-362.

Jackson, Douglas Northrop (1974). *Personality Research Form Manual*. Goshen, N.Y, Research Psychologists Press.

Jackson, Joshua J., Bogg, Tim, Walton, Kate E., Wood, Dustin, Harms, Peter D., Lodi-Smith, Jennifer, Edmonds, Grant W. and Roberts, Brent W. (2009). "Not All Conscientiousness Scales Change Alike: A Multimethod, Multisample Study of Age Differences in the Facets of Conscientiousness." *Journal of Personality and Social Psychology* **96**(2): 446-459.

Jackson, Joshua J., Hill, Patrick L., Payne, Brennan R. , Roberts, Brent W. and Stine-Morrow, Elizabeth A. L. (2010). "Can an Old Dog Learn (and Want to Experience) New Tricks? Cognitive Training Increases Openness in Older Adults." Unpublished Manuscript. University of Illinois, Department of Psychology.

James, William (1890). *The Principles of Psychology*. New York, NY, H. Holt and Company.

James, William (1899). *Talks to Teachers on Psychology: And to Students on Some of Life's Ideals*. Boston, Press of Geo. H. Ellis Co. (Inc.)

Janosz, Michel, LeBlanc, Marc, Boulerice, Bernard and Tremblay, Richard E. (1997). "Disentangling the Weight of School Dropout Predictors: A Test on Two Longitudinal Samples." *Journal of Youth and Adolescence* **26**(6): 733-762.

Jencks, Christopher, Smith, Marshall, Acland, Henry, Bane, Mary Jo, Cohen, David K., Gintis, Herbert, Heyns, Barbara and Michelson, Stephen (1972). *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York, Basic Books.

Jensen, Arthur R. (1980). *Bias in Mental Testing*. New York, Free Press.

Jensen, Arthur R. (1998). *The g Factor: The Science of Mental Ability*. Westport, CT, Praeger.

John, Oliver P. (1990). "The "Big Five" Factor Taxonomy: Dimensions of Personality in the Natural Language and in Questionnaires." in *Handbook of Personality: Theory and Research*. L. A. Pervin and O. P. John, eds. New York, NY, Guilford Press**:** 66.

John, Oliver P., Caspi, Avshalom, Robins, Richard W. and Moffitt, Terrie E. (1994). "The "Little Five": Exploring the Nomological Network of the Five-Factor Model of Personality in Adolescent Boys." *Child Development* **65**(1): 160-178.

John, Oliver P. and Srivastava, Sanjay (1999a). "The Big Five Trait Taxonomy: History, Measurement and Theoretical Perspectives." in *Handbook of Personality: Theory and Research*. L. A. Pervin and O. P. John, eds. New York, The Guilford Press**:** 102-138.

John, Oliver P. and Srivastava, Sanjay (1999b). "The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives." in *Handbook of Personality: Theory and Research (2nd Ed.).* L. A. Pervin and O. P. John, eds. New York, NY, US, Guilford Press**:** 102-138.

Johnson, J, Wilke, A and Weber, EU (2004). "Beyond a Trait View of Risk Taking: A Domain-Specific Scale Measuring Risk Perceptions, Expected Benefits, and Perceived-Risk Attitudes in German-Speaking Populations." *Polish Psychological Bulletin* **35**(3): 153-172.

Johnson, Wendy, Bouchard, Thomas J., Jr., Krueger, Robert F., McGue, Matt and Gottesman, Irving I. (2004). "Just One g: Consistent Results from Three Test Batteries." *Intelligence* **32**(1): 95-107.

Judge, TA and Hurst, C (2007a). "Capitalizing on One's Advantages: Role of Core Self-Evaluations." *Journal of Applied Psychology* **92**(5): 1212-1227.

Judge, Timothy A., Bono, Joyce E., Erez, Amir and Locke, Edwin A. (2005). "Core Self-Evaluations and Job and Life Satisfaction: The Role of Self-Concordance and Goal Attainment." *Journal of Applied Psychology* **90**(2): 257-268.

Judge, Timothy A., Erez, Amir, Bono, Joyce E. and Thoresen, Carl J. (2002). "Are Measures of Self-Esteem, Neuroticism, Locus of Control, and Generalized Self-Efficacy Indicators of a Common Core Construct?" *Journal of Personality and Social Psychology* **83**(3): 693-710.

Judge, Timothy A. and Hurst, Charlice (2007b). "Capitalizing on One's Advantages: Role of Core Self-Evaluations." *Journal of Applied Psychology* **92**(5): 1212-1227.

Kahneman, Daniel and Tversky, Amos (1979). "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* **47**(2): 263-291.

Kelly, Francis J. and Veldman, Donald J. (1964). "Delinquency and School Dropout Behavior as a Function of Impulsivity and Nondominant Values." *The Journal of Abnormal and Social Psychology* **69**(2): 190-194.

Kern, Margaret L. and Friedman, Howard S. (2008). "Do Conscientious Individuals Live Longer?" *Health Psychology* **27**: 505-512.

Kern, Margaret L. and Friedman, Howard S. (2010a). "Personality and Pathways of Influence on Physical Health." In press, *Social and Personality Psychology Compass*.

Kern, Margaret L. and Friedman, Howard S. (2010b). "Why Do Some People Thrive While Other Succumb to Disease and Stagnation? Personality, Social Relations, and Resilience." in *New Frontiers in Resilient Aging: Life-Strengths and Well-Being in Late Life*. P. S. Fry and C. L. M. Keyes, eds. Cambridge, UK, Cambridge University Press**:** 162-184.

Kilburn, M. Rebecca, Hanser, Lawrence M. and Klerman, Jacob A. (1998). "Information in the NELS and Approaches to Imputing Test Scores." in *Estimating the Afqt Scores for National Educational Longitudinal Study (NELS) Respondents*. Santa Monica, CA, National Defense Research Institute, Arroya Center: Rand Corp.

Kirby, Kris N and Herrnstein, Richard J (1995). "Preference Reversals Due to Myopic Discounting of Delayed Reward." *Psychological Science* **6**(2): 83-89.

Kirby, Kris N. (1997). "Bidding on the Future: Evidence against Normative Discounting of Delayed Rewards." *Journal of Experimental Psychology: General* **126**(1): 54.

Knudsen, Eric I., Heckman, James J., Cameron, Judy and Shonkoff, Jack P. (2006). "Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce." *Proceedings of the National Academy of Sciences* **103**(27): 10155-10162.

Knutson, Brian, Wolkowitz, Owen M., Cole, Steve W., Chan, Theresa, Moore, Elizabeth A., Johnson, Ronald C., Terpstra, Jan, Turner, Rebecca A. and Reus, Victor I. (1998). "Selective Alteration of Personality and Social Behavior by Serotonergic Intervention." *Am J Psychiatry* **155**(3): 373-379.

Kobrin, JL, Patterson, BF, Shaw, EJ, Mattern, KD and Barbuti, SM (2008). "Validity of the SAT for Predicting First-Year College Grade Point Average." *The College Board* **5**: 1-10.

Kogan, N. and Wallach, M. A. (1967). "Risk Taking as a Function of the Situation, the Person and the Group." in *New Directions in Psychology III*. New York, Holt**:** 111-278.

Kohn, Alfie (1999). *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*. Boston MA, Houghton Mifflin Company.

Kohn, Melvin L. and Schooler, Carmi (1978). "The Reciprocal Effects of the Substantive Complexity of Work and Intellectual Flexibility: A Longitudinal Assessment." *American Journal of Sociology* **84**(1): 24.

Kosfeld, M, Heinrichs, M, Zak, PJ, Fischbacher, U and Fehr, E (2005). "Oxytocin Increases Trust in Humans." *Nature* **435**(7042): 673-676.

Köszegi, B (2006). "Ego Utility, Overconfidence, and Task Choice." *Journal of the European Economic Association* **4**(4): 673-707.

Kristensen, Petter and Bjerkedal, Tor (2007). "Explaining the Relation between Birth Order and Intelligence." *Science* **316**(5832): 1717.

Krueger, Robert F. and Johnson, Wendy (2008). "Behavioral Genetics and Personality: A New Look at the Integration of Nature and Nurture." in *Handbook of Personality: Theory and Research*. O. P. John, R. W. Robins and L. A. Pervin, eds. New York, NY, Guilford Press**:** 287-310.

Krueger, Robert F., South, Susan, Johnson, Wendy and Iacono, William G. (2008). "The Heritability of Personality Is Not Always 50%: Gene-Environment Interactions and Correlations between Personality and Parenting." *Journal of Personality*.

Kube, Sebastian, Maréchal, Michel André and Puppe, Clemens (2006). "Putting Reciprocity to Work--Positive Versus Negative Responses in the Field." University of St. Gallen Working Paper, 2006-27.

Kubzansky, Laura D., Sparrow, David, Vokonas, Pantel and Kawachi, Ichiro (2001). "Is the Glass Half Empty or Half Full? A Prospective Study of Optimism and Coronary Heart Disease in the Normative Aging Study." *Psychosomatic Medicine* **63**(6): 910.

Kuncel, Nathan R. and Hezlett, Sarah A. (2007). "Standardized Tests Predict Graduate Students' Success." *Science* **315**(5815): 1080-1081.

Kuncel, Nathan R., Ones, Deniz S. and Sackett, Paul R. (2010). "Individual Differences as Predictors of Work, Educational, and Broad Life Outcomes." *Personality and Individual Differences* **49**(4): 331-336.

Kvist, Ann Valentin and Gustafsson, Jan-Eric (2007). "The Relation between Fluid Intelligence and the General Factor as a Function of Cultural Background: A Test of Cattell's Investment Theory." *Institute for Labour Market Policy Evaluation* **23**.

Kyllonen, Patrick C., Walters, Alyssa M. and Kaufman, James C. (2005). "Noncognitive Constructs and Their Assessment in Graduate Education: A Review." *Educational Assessment* **10**(3): 153.

Larson, Gerald E., Saccuzzo, Dennis P. and Brown, James (1994). "Motivation: Cause or Confound in Information Processing/Intelligence Correlations?" *Acta Psychologica* **85**(1): 25.

Layard, Richard (2005). *Happiness: Lessons from a New Science*. London, Penguin.

Le, Anh T., Miller, Paul W., Slutske, Wendy S. and Martin, Nicholas G. (2010). "Are Attitudes Towards Economic Risk Heritable? Analysis Using the Australian Twin Study of Gambling." IZA Discussion Paper No. 4859.

Lear, Gabriel Richardson (2004). *Happy Lives and the Highest Good: Aristotle's Nicomachean Ethics*. Princeton, Princeton University Press.

Lejuez, C. W., Aklin, Will M., Zvolensky, Michael J. and Pedulla, Christina M. (2003). "Evaluation of the Balloon Analogue Risk Task (Bart) as a Predictor of Adolescent Real-World Risk-Taking Behaviours." *Journal of adolescence* **26**(4): 475.

Lejuez, C. W., Read, Jennifer P., Kahler, Christopher W., Richards, Jerry B., Ramsey, Susan E., Stuart, Gregory L., Strong, David R. and Brown, Richard A. (2002). "Evaluation of a Behavioral Measure of Risk Taking: The Balloon Analogue Risk Task (Bart)." *Journal of Experimental Psychology: Applied* **8**(2): 75.

Lesch, Klaus-Peter, Bengel, Dietmar, Heils, Armin, Sabol, Sue Z., Greenberg, Benjamin D., Petri, Susanne, Benjamin, Jonathan, Muller, Clemens R., Hamer, Dean H. and Murphy, Dennis L. (1996). "Association of Anxiety-Related Traits with a Polymorphism in the Serotonin Transporter Gene Regulatory Region." *Science* **274**: 1527-1531.

Levitt, Steven D. and List, John A. (2007). "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?" *Journal of Economic Perspectives* **21**(2): 153.

Lezak, Muriel D. (1987). "Relationships between Personality Disorders, Social Disturbances, and Physical Disability Following Traumatic Brain Injury." *The Journal of Head Trauma Rehabilitation* **2**(1): 57-69.

Lievens, Filip, Dilchert, Stephan and Ones, Deniz S. (2009). "Personality Scale Validities Increase Throughout Medical School." *Journal of Applied Psychology* **94**(6): 1514-1535.

Lillard, Angeline and Else-Quest, Nicole (2006). "The Early Years: Evaluating Montessori." *Science* **313**: 1893-1894.

List, John A (2006). "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions." *Journal of Political Economy* **114**: 1-37.

List, John A. (2003). "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics* **118**(1): 41-71.

List, John A. (2007). "On the Interpretation of Giving in Dictator Games." *Journal of Political Economy* **115**(3): 482.

List, John A. (2009). "Social Preferences: Some Thoughts from the Field." *Annual Review of Economics* **1**: 563-583.

Lizzeri, Alessandro and Siniscalchi, Marciano (2008). "Parental Guidance and Supervised Learning." *Quarterly Journal of Economics* **123**(3): 1161-1195.

Lleras, Christy (2008). "Do Skills and Behaviors in High School Matter? The Contribution of Noncognitive Factors in Explaining Differences in Educational Attainment and Earnings." *Social Science Research* **37**(3): 888-902.

Lochner, Lance J. and Moretti, Enrico (2004). "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review* **94**(1): 155-189.

Lounsbury, John W., Steel, Robert P., Loveland, James M. and Gibson, Lucy W. (2004). "An Investigation of Personality Traits in Relation to Adolescent School Absenteeism." *Journal of Youth and Adolescence* **33**(5): 457-466.

Lubbers, Miranda J., Van Der Werf, Margaretha P. C., Kuyper, Hans and Hendriks, A. A. Jolijn (2010). "Does Homework Behavior Mediate the Relation between Personality and Academic Performance?" *Learning and Individual Differences* **20**(3): 203-208.

Lubinski, David (2004). "Introduction to the Special Section on Cognitive Abilities: 100 Years after Spearman's (1904) '"General Intelligence," Objectively Determined and Measured'." *Journal of Personality and Social Psychology* **86**(1): 96.

Lucas, Richard E. and Donnellan, M. Brent (2009). "If the Person-Situation Debate Is Really over, Why Does It Still Generate So Much Negative Affect?" *Journal of Research in Personality* **43**: 146-149.

Lykken, David T. (2007). "A More Accurate Estimate of Heritability." *Twin Research and Human Genetics* **10**(1): 168.

Machin, Stephen, Marie, Olivier and Vujić, Sunčica (2010). "The Crime Reducing Effect of Education." CEP Discussion Paper No. 979.

Mandelbrot, Benoit B. (1961). "Stable Paretian Random Functions and the Multiplicative Variation of Income." *Econometrica* **29**(4): 517-543.

Mandelbrot, Benoit B. (1962). "Paretian Distributions and Income Maximization." *Quarterly Journal of Economics* **76**: 57-85.

Mandelbrot, Benoit B. (1963). "The Stable Paretian Income Distribution, When the Apparent Exponent Is near Two." *International Economic Review* **4**: 111-115.

Markon, Kristian E., Krueger, Robert F. and Watson, David (2005). "Delineating the Structure of Normal and Abnormal Personality: An Integrative Hierarchical Approach." *Jounral of Personality and Social Psychology* **88**(1): 139-157.

Martin, Leslie R., Friedman, Howard S. and Schwartz, Joseph E. (2007). "Personality and Mortality Risk across the Life Span: The Importance of Conscientiousness as a Biopsychosocial Attribute." *Health Psychology* **26**(4): 428-436.

Martin, Roy P. (1989). "Activity Level, Distractibility, and Persistence: Critical Characteristics in Early Schooling." in *Temperament in Childhood*. G. A. Kohnstamm, J. E. Bates and M. K. Rothbart, eds. Chichester, England, John Wiley and Sons Ltd.**:** 451-461.

Martins, Pedro S. (2010). "Can Targeted, Non-Cognitive Skills Programs Improve Achievement?" *IZA Discussion Paper No. 5266*.

Mataro, Maria, Jurado, Ángeles, Garcia-Sanchez, Carmen, Barraquer, Lluis, Costa-Jussa, Frederic. R. and Junque, Carmen (2001). "Long-Term Effects of Bilateral Frontal Brain Lesion: 60 Years after Injury with an Iron Bar." *Archives of neurology* **58**(7): 1139.

Matsumoto, K and Tanaka, K (2004). "Conflict and Cognitive Control." *Science* **303**(5660): 969-970.

Mayer, John D. and Salovey, Peter (1997). "What Is Emotional Intelligence?" in *Emotional Development and Emotional Intelligence: Educational Implications*. P. Salovey and D. J. Sluyter, eds. New York, NY, Basic Books**:** 3.

McAdams, Dan P. (1992). "The Five-Factor Model in Personality: A Critical Appraisal." *Journal of Personality. Special Issue: The five-factor model: Issues and applications* **60**(2): 329-361.

McAdams, Dan P. (2006). *The Person: A New Introduction to Personality Psychology*. Hoboken, N.J., J. Wiley and Sons.

McArdle, John J., Hamagami, Fumiaki, Meredith, William and Bradway, Katherine P. (2000). "Modeling the Dynamic Hypotheses of Gf-Gc Theory Using Longitudinal Life-Span Data." *Learning and Individual Differences* **12**(1): 53-79.

McCabe, Lisa A., Cunnington, Marisol and Brooks-Gunn, Jeanne (2004). "The Development of Self-Regulation in Young Children: Individual Characteristics and Environmental Contexts." in *Handbook of Self-Regulation: Research, Theory, and Applications*. R. F. Baumeister and K. D. Vohs, eds. New York, NY, Guilford Press**:** 340.

McClelland, David C., Koestner, Richard and Weinberger, Joel (1989). "How Do Self-Attributed and Implicit Motives Differ?" *Psychological Review* **96**(4): 690.

McClelland, Megan M., Cameron, Claire E., Connor, Carol McDonald, Farris, Carrie L., Jewkes, Abigail M. and Morrison, Frederick J. (2007). "Links between Behavioral Regulation and Preschoolers' Literacy, Vocabulary, and Math Skills." *Developmental Psychology* **43**(4): 947-959.

McClure, Samuel M., Laibson, David I., Loewenstein, George F. and Cohen, Jonathan D. (2004). "Separate Neural Systems Value Immediate and Delayed Monetary Rewards." *Science* **306**(5695): 503-507.

McCrae, Robert R. and Costa, Paul T. (1990). *Personality in Adulthood*. New York, Guilford Press.

McCrae, Robert R. and Costa, Paul T. (1994). "The Stability of Personality: Observation and Evaluations." *Current Directions in Psychological Science* **3**(6): 173-175.

McCrae, Robert R. and Costa, Paul T. (1996). "Toward a New Generation of Personality Theories: Theoretical Contexts for the Five-Factor Model." in *The Five-Factor Model of Personality: Theoretical Perspectives*. J. S. Wiggins, ed. New York, NY, Guilford Press**:** 51.

McCrae, Robert R. and Costa, Paul T. (1997a). "Conceptions and Correlates of Openness to Experience." in *Handbook of Personality Psychology.* R. Hogan, J. A. Johnson and S. R. Briggs, eds. San Diego, CA, US, Academic Press**:** 825-847.

McCrae, Robert R. and Costa, Paul T. (1997b). "Personality Trait Structure as a Human Universal." *American Psychologist* **52**(5): 509-516.

McCrae, Robert R. and Costa, Paul T. (2003). *Personality in Adulthood: A Five-Factor Theory Perspective*. New York, Guilford Press.

McCrae, Robert R. and Costa, Paul T. Jr. (2008). "The Five-Factor Theory of Personality." in *Handbook of Personality: Theory and Research*. O. P. John, R. W. Robins and L. A. Pervin, eds. New York, NY, The Guilford Press**:** 159-181.

McGue, Matt, Bacon, Steven and Lykken, David T. (1993). "Personality Stability and Change in Early Adulthood: A Behavioral Genetic Analysis." *Developmental psychology* **29**(1): 96.

McGue, Matt, Bouchard, Thomas J., Iacono, William G. and Lykken, David T. (1993). "Behavioral Genetics of Cognitive Ability: A Life-Span Perspective." in *Nature, Nurture and Psychology*. R. Plomin and G. E. McClearn, eds. Washington, DC, American Psychological Association**:** 59.

McHenry, JJ, Hough, LM, Toquam, JL, Hanson, M and S, Ashworth (1990). "Project a Validity Results: The Relationship between Predictor and Criterion Domains." *Personnel Psychology* **43**(2): 335-354.

Mealey, Linda (1995). "The Sociobiology of Sociopathy: An Integrated Evolutionary Model." *Behavioral and Brain Sciences* **18**(03): 523-541.

Meier, Stephan (2007). "A Survey of Economic Theories and Field Evidence on Pro-Social Behavior." in *Economics and Psychology: A Promising New Cross-Disciplinary Field*. B. S. Frey and A. Stutzer, eds. Cambridge, MA, MIT Press**:** 51-88.

Mershon, Bryan and Gorsuch, Richard L. (1988). "Number of Factors in the Personality Sphere: Does Increase in Factors Increase Predictability of Real-Life Criteria?" *Journal of Personality and Social Psychology* **55**(4): 675-680.

Michalko, KT and Saklofske, DH (1996). "A Psychometric Investigation of the Wechsler Individual Achievement Test with a Sample of Saskatchewan Schoolchildren." *Canadian Journal of School Psychology* **12**(1): 44-54.

Miller, Earl K. (2000). "The Prefrontal Cortex and Cognitive Control." *Nature Reviews Neuroscience* **1**(October): 59-65.

Miller, Earl K. and Cohen, Jonathan D. (2001). "An Integrative Theory of Prefrontal Cortex Function." *Annual Review of Neuroscience* **24**(1): 167-202.

Mischel, Walter (1968). *Personality and Assessment*. New York, Wiley.

Mischel, Walter (2004). "Toward an Integrative Science of the Person." *Annual Review of Psychology* **55**: 1-22.

Mischel, Walter (2009). "From *Personality and Assessment* (1968) to Personality Science, 2009." *Journal of Research in Personality* **43**: 282-290.

Mischel, Walter, Ayduk, O., Berman, M. G., Casey, B. J., Gotlib, I. H., Jonides, J., Kross, E., Teslovich, T., Wilson, N. L., Zayas, V. and Shoda, Y. (2010). "'Willpower' over the Life Span: Decomposing Self-Regulation." *Social Cognitive and Affective Neuroscience*.

Mischel, Walter and Metzner, Ralph (1962). "Preference for Delayed Reward as a Function of Age, Intelligence, and Length of Delay Interval." *The Journal of Abnormal and Social Psychology* **64**(6): 425.

Mischel, Walter and Shoda, Yuichi (1995). "A Cognitive-Affective System Theory of Personality: Reconceptualizing Situations, Dispositions, Dynamics, and Invariance in Personality Structure." *Psychological Review* **102**(2): 246-268.

Mischel, Walter and Shoda, Yuichi (2008). "Toward a Unifying Theory of Personality: Integrating Dispositions and Processing Dynamics within the Cognitive-Affective Processing System." in *Handbook of Personality Psychology*. O. P. John, R. W. Robins and L. A. Pervin, eds. New York, Guilford Press.

Mischel, Walter, Shoda, Yuichi and Rodriguez, Monica L. (1989). "Delay of Gratification in Children." *Science* **244**(4907): 933-938.

Miyake, Akira, Friedman, Naomi P., Emerson, Michael J., Witzki, Alexander H. and Howerter, Amy (2000). "The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis." *Cognitive Psychology* **41**(1): 49-100.

Moffitt, Terrie E., Caspi, Avshalom and Rutter, Michael (2005). "Strategy for Investigating Interactions between Measured Genes and Measured Environments." *Archives of General Psychiatry* **62**(5): 473-481.

Monterosso, John R. and Luo, Shan (2010). "An Argument against Dual Valuation System Competition: Cognitive Capacities Supporting Future Orientation Mediate Rather Than Compete with Visceral Motivations." *Journal of Neuroscience, Psychology, and Economics* **3**(1): 1-14.

Morgeson, Frederick P., Campion, Michael A., Dipboye, Robert L., Hollenbeck, John R., Murphy, Kevin and Schmitt, Neal (2007). "Reconsidering the Use of Personality Tests in Personnel Selection Contexts." *Personnel Psychology* **60**(3): 683.

Moskowitz, D. S. (1982). "Coherence and Cross-Situational Generality in Personality: A New Analysis of Old Problems." *Journal of Personality and Social Psychology* **43**(4): 754-768.

Moutafi, Joanna, Furnham, Adrian and Tsaousis, Ioannis (2006). "Is the Relationship between Intelligence and Trait Neuroticism Mediated by Test Anxiety?" *Personality and Individual Differences* **40**(3): 587.

Mroczek, Daniel K. and Spiro, Avron (2007). "Personality Change Influences Mortality in Older Men." *Psychological Science* **18**(5): 371-376.

Mueller, Gerrit and Plug, Erik (2006). "Estimating the Effect of Personality on Male and Female Earnings." *Industrial and Labor Relations Review* **60**(1): 3-22.

Mueser, Peter R. (1979). "The Effects of Noncognitive Traits." in *Who Gets Ahead? The Determinants of Economic Success in America*. C. Jencks, S. Bartlett, M. Corcoran, J. Crouse, D. Eaglesfield, G. Jackson, K. McClelland, P. Mueser, M. Olneck, J. Schwartz, S. Ward and J. Williams, eds. New York, NY, Basic Books**:** 122-158.

Murray, Henry Alexander (1938). *Explorations in Personality: A Clinical and Experimental Study of Fifty Men of College Age*. New York, London, Oxford University Press.

Neal, Derek A. and Johnson, William R. (1996). "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* **104**(5): 869-895.

Neisser, Ulric, Boodoo, Gwyneth, Bouchard, Thomas J., Jr., Boykin, A. Wade, Brody, Nathan, Ceci, Stephen J., Halpern, Diane F., Loehlin, John C., Perloff, Robert, Sternberg, Robert J. and Urbina, Susana (1996). "Intelligence: Knowns and Unknowns." *American Psychologist* **51**(2): 77-101.

Nettle, Daniel (2006). "The Evolution of Personality Variation in Humans and Other Animals." *American Psychologist* **61**(6): 622-631.

Newman, Denise L., Caspi, Avshalom, Moffitt, Terrie E. and Silva, Phil A. (1997). "Antecedents of Adult Interpersonal Functioning: Effects of Individual Differences in Age 3 Temperament." *Developmental psychology* **33**(2): 206.

Niederle, Muriel, Möbius, Markus, Niehaus, Paul and Rosenblat, Tanya (2010). "Managing Self-Confidence: Theory and Experimental Evidence." Unpublished manuscript. Iowa State University, Department of Economics.

Noftle, Erik E. and Robins, Richard W. (2007). "Personality Predictors of Academic Outcomes: Big Five Correlates of GPA and SAT Scores." *Journal of Personality and Social Psychology* **93**(1): 116-130.

Noyes, C. Reinold (1945). "Director's Comment." in *Income from Independent Professional Practice*. M. Friedman and S. S. Kuznets, eds. New York, National Bureau of Economic Research**:** 405-410.

Nyhus, Ellen K. and Pons, Empar (2005). "The Effects of Personality on Earnings." *Journal of Economic Psychology* **26**(3): 363-384.

O'Leary, UM, Rusch, KM and Guastello, SJ (2006). "Estimating Age-Stratified WAIS-R IQs from Scores on the Raven's Standard Progressive Matrices." *Journal of Clinical Psychology* **47**: 277-284.

O'Gorman, J. G. and Baxter, E. (2002). "Self-Control as a Personality Measure." *Personality and Individual Differences* **32**: 533-539.

Oliver, Pamella H., Guerin, Diana W. and Gottfried, Allen W. (2007). "Temperamental Task Orientation: Relation to High School and College Educational Accomplishments." *Learning and Individual Differences* **17**(3): 220-230.

Omizo, MM (1980). "The Differential Aptitude Tests as Predictors of Success in a High School for Engineering Program." *Educational and Psychological Measurement* **40**(1): 197-203.

Ones, Deniz S. and Viswesvaran, Chockalingam (1998). "The Effects of Social Desirability and Faking on Personality and Integrity Assessment for Personnel Selection." *Human Performance* **11**(2-3): 245.

Overton, Willis F. (2007). *Monographs of the Society for Research in Child Development* **72**(3).

Pailing, Patricia E. and Segalowitz, Sidney J. (2004). "The Error-Related Negativity as a State and Trait Measure: Motivation, Personality, and Erps in Response to Errors." *Psychophysiology* **41**(1): 84.

Parco, J. E., Rapoport, a. and Stein, W. E. (2002). "Effects of Financial Incentives on the Breakdown of Mutual Trust." *Psychological Science* **13**: 292-297.

Paulhus, Delroy L. (1984). "Two-Component Models of Socially Desirable Responding." *Journal of Personality and Social Psychology* **46**(3): 598.

Paunonen, Sampo V. and Ashton, Michael C. (2001). "Big Five Factors and Facets and the Prediction of Behavior." *Journal of Personality and Social Psychology* **81**(3): 524-539.

Perkins, David N. and Tishman, Shari (2001). "Dispositional Aspects of Intelligence." in *Intelligence and Personality: Bridging the Gap in Theory and Measurement*. J. M. Collis and S. Messick, eds. Mahwah, NJ, Lawrence Erlbaum Associates, Publishers**:** 233-257.

Pesonen, Anu-Katriina, Räikkönen, Katri, Heinonen, Kati, Andersson, Sture, Hovi, Petteri, Järvenpää, Anna-Liisa, Eriksson, Johan G. and Kajantie, Eero (2008). "Personality of Young Adults Born Prematurely: Helsinki Study of Very Low Birth Weight Adults." *Journal of Child Psychology and Psychiatry* **49**(6): 609-617.

Peters, Ellen, Vastfjall, Daniel, Slovic, Paul, Mertz, C. K., Mazzocco, Ketti and Dickert, Stephan (2006). "Numeracy and Decision Making." *Psychological Science* **17**(5): 407.

Phelps, Elizabeth A. (2006). "Emotion and Cognition: Insights from Studies of the Human Amygdala." *Annual Review of Psychology* **57**: 27-53.

Piatek, Rémi and Pinger, Pia (2010). "Maintaining (Locus of) Control? Assessing the Impact of Locus of Control on Education Decisions and Wages." Institute for the Study of Labor (IZA), Discussion Paper No. 5289.

Pinker, Steven (2002). *The Blank Slate: The Modern Denial of Human Nature*. New York, Penguin Books.

Plomin, Robert, DeFries, John C., Craig, Ian W. and McGuffin, Peter (2002). *Behavioral Genetics in the Postgenomic Era*. Washington, D.C., American Psychological Association.

Plomin, Robert and Nesselroade, John R. (1990). "Behavioral Genetics and Personality Change." *Journal of Personality* **58**(1): 191.

Ponitz, Claire Cameron, McClelland, Megan M., Jewkes, Abigail M., Connor, Carol McDonald, Farris, Carrie L. and Morrison, Frederick J. (2008). "Touch Your Toes! Developing a Direct Measure of Behavioral Regulation in Early Childhood." *Early Childhood Research Quarterly* **23**(2): 141-158.

Ponitz, Claire Cameron, McClelland, Megan M., Matthews, J. S. and Morrison, Frederick J. (2009). "A Structured Observation of Behavioral Self-Regulation and Its Contribution to Kindergarten Outcomes." *Developmental Psychology* **45**(3): 605-619.

Poropat, Arthur E. (2009). "A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance." *Psychological Bulletin* **135**(2): 322-338.

Posner, Michael I. and Rothbart, Mary K. (2000). "Developing Mechanisms of Self-Regulation." *Development and psychopathology* **12**(3): 427.

Project Head Start (1969). *The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development*. Bladensburg, MD, Westinghouse Learning Corporation.

Putnam, Samuel P., Ellis, Lesa K. and Rothbart, Mary K. (2001). "The Structure of Temperament from Infancy through Adolescence." in *Advances in Research on Temperament*. A. Eliasz and A. Angleitner, eds. Lengerich, Germany, Pabst Science Publishers**:** 165.

Rabin, M (2000). "Diminishing Marginal Utility of Wealth Cannot Explain Risk Aversion." in *Choices, Values, and Frames*. D. Kahneman and A. Tversky, eds. New York, NY, Cambridge University Press**:** 202-208.

Raven, J., Raven, J. C. and Court, J.H. (1988). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX, Harcourt Assessment.

Raven, John C. (1962). *Advanced Progressive Matrices: Sets I and II*. London, H.K. Lewis.

Raver, C. Cybele and Zigler, Edward F. (1997). "Social Competence: An Untapped Dimension in Evaluating Head Start's Success." *Early Childhood Research Quarterly* **12**(4): 363-385.

Read, Daniel and Read, Nicoleta Liliana (2004). "Time Discounting over the Lifespan." *Organizational behavior and human decision processes* **94**(1): 22.

Reuter, M, Montag, C, Altmann, S, Bendlow, F, Elger, C, Kirsch, P, Weber, B and Falk, A (2009). "Genetically Determined Differences in Human Trust Behavior: The Role of the

Oxytocin Receptor Gene." Unpublished manuscript. Department of Economics, University of Bonn.

Revelle, W., Wilt, J. and Condon, D. (2010). "Individual Differences and Differential Psychology: A Brief History and Prospect." in *Handbook of Individual Differences*. T. Chamorro-Premuzic, A. Furnham and S. von Stumm, eds. In Press.

Revelle, William (1997). "Extraversion and Impulsivity: The Lost Dimension?" in *The Scientific Study of Human Nature: Tribute to Hans J. Eysenck at Eighty*. H. Nyborg, ed. Amsterdam, Pergamon/Elsevier Science**:** 189.

Rindermann, Heiner (2007). "The g-Factor of International Cognitive Ability Comparisons: The Homogeneity of Results in Pisa, Timss, Pirls and IQ-Tests across Nations." *European Journal of Personality* **21**(5): 667.

Roberts, Brent W. (1997). "Plaster or Plasticity: Are Adult Work Experiences Associated with Personality Change in Women?" *Journal of Personality* **65**(2): 205.

Roberts, Brent W. (2006). "Personality Development and Organizational Behavior." in *Research on Organizational Behavior*. B. Staw, ed. Oxford, UK, Elsevier Science/JAI Press. **27:** 1-41.

Roberts, Brent W. (2007). "Contextualizing Personality Psychology." *Journal of Personality* **75**(6): 1071-1082.

Roberts, Brent W. (2009). "Back to the Future: Personality and Assessment and Personality Development." *Journal of Research in Personality* **43**(2): 137-145.

Roberts, Brent W. and Chapman, Christopher N. (2000). "Change in Dispositional Well-Being and Its Relation to Role Quality: A 30-Year Longitudinal Study." *Journal of Research in Personality* **34**(1): 26.

Roberts, Brent W., Chernyshenko, Oleksandr S., Stark, Stephen and Goldberg, Lewis R. (2005). "The Structure of Conscientiousness: An Empirical Investigation Based on Seven Major Personality Questionnaires." *Personnel Psychology* **58**(1): 103.

Roberts, Brent W. and DelVecchio, Wendy F. (2000). "The Rank-Order Consistency of Personality Traits from Childhood to Old Age: A Quantitative Review of Longitudinal Studies." *Psychological Bulletin* **126**(1): 3-25.

Roberts, Brent W., Harms, Peter, Smith, Jennifer L., Wood, Dustin and Webb, Michelle (2006). "Using Multiple Methods in Personality Psychology." in *Handbook of Multimethod Measurement in Psychology*. M. Eid and E. Diener, eds. Washington, D.C., American Psychological Association**:** 321-335.

Roberts, Brent W., Helson, Ravenna and Klohnen, Eva C. (2002). "Personality Development and Growth in Women across 30 Years: Three Perspectives." *Journal of Personality* **70**(1): 79.

Roberts, Brent W., Kuncel, Nathan R., Shiner, Rebecca, Caspi, Avshalom and Goldberg, Lewis R. (2007). "The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes." *Perspectives on Psychological Science* **2**(4): 313-345.

Roberts, Brent W. and Mroczek, Daniel (2008). "Personality Trait Change in Adulthood." *Current Directions in Psychological Science* **17**(1): 31-35.

Roberts, Brent W., Walton, Kate E. and Viechtbauer, Wolfgang (2006). "Patterns of Mean-Level Change in Personality Traits across the Life Course: A Meta-Analysis of Longitudinal Studies." *Psychological Bulletin* **132**(1): 1-25.

Roberts, Brent W., Wood, Dustin and Caspi, Avshalom (2008). "Personality Development." in *Handbook of Personality: Theory and Research*. O. P. John, R. W. Robins and L. A. Pervin, eds. New York, NY, Guilford Press.

Roberts, Richard D., Goff, Ginger Nelson, Anjoul, Fadi, Kyllonen, P. C., Pallier, Gerry and Stankov, Lazar (2000). "The Armed Services Vocational Aptitude Battery (Asvab): Little More Than Acculturated Learning (Gc)!?" *Learning and Individual Differences* **12**(1): 81-103.

Roberts, Richard D., Markham, Pippa M., Matthews, Gerald and Zeidner, Moshe (2005). "Assessing Intelligence: Past, Present, and Future." in *Handbook of Understanding and Measuring Intelligence*. O. Wilhelm and R. W. Engle, eds. Thousand Oaks, CA, Sage Publications Inc.**:** 333.

Robson, Arthur J. (1996). "A Biological Basis for Expected and Non-Expected Utility." *Journal of Economic Theory* **68**: 397-424.

Robson, Arthur J. (2001). "Why Would Nature Give Individuals Utility Functions?" *The Journal of Political Economy* **109**(4): 900-914.

Robson, Arthur J. and Samuelson, Larry (2007). "The Evolution of Intertemporal Preferences." *AEA Paper and Proceedings*.

Robson, Arthur J. and Samuelson, Larry (2009). "The Evolution of Time Preference with Aggregate Uncertainty." *American Economic Review* **9**(5): 1925-1953.

Rodríguez-Planas, Nuria (2010). "Mentoring, Educational Services, and Economic Incentives: Longer-Term Evidence on Risky Behaviors from a Randomized Trial." *IZA Discussion Paper*.

Rosenberg, Morris (1965). *Society and the Adolescent Self-Image*, Princeton University Press.

Rosenberg, Morris (1989). "Society and the Adolescent Self-Image (Rev. Ed.)." *Middletown, CT, England: Wesleyan University Press* **347**.

Ross, L. and Nisbett, R. E. (1991). *The Person and the Situation: Perspectives of Social Psychology*. New York, McGraw-Hill.

Rothbart, Mary K., Ahadi, Stephan A. and Evans, David E. (2000). "Temperament and Personality: Origins and Outcomes." *Journal of Personality and Social Psychology* **78**(1): 122.

Rothlisburg, B (1987). "Comparing the Stanford-Binet, Fourth Edition to the Wisc-R: A Concurrent Validity Study." *Journal of School Psychology* **25**: 193-196.

Rotter, Julian B. (1966). "Generalized Expectancies for Internal Versus External Control of Reinforcement." *Psychological Monographs* **80**: 1-28.

Rowe, D. C. (1981). "Environmental and Genetic Influences on Dimensions of Perceived Parenting: A Twin Study." *Developmental Psychology* **17**: 203-208.

Rowe, D. C. (1983). "A Biometrical Analysis of Perceptions of the Family Environment: A Study of Twin and Singleton Sibling Kinships." *Child Development* **54**: 416-423.

Roy, A.D. (1951). "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* **3**(2): 135-146.

Rueda, M. Rosario, Rothbart, Mary K., McCandliss, Bruce D., Saccomanno, Lisa and Posner, Michael I. (2005). "Training, Maturation, and Genetic Influences on the Development of Executive Attention." *Proceedings of the National Academy of Sciences* **102**(41): 14931-14936.

Rust, J (2008). "Dynamic Programming." in *The New Palgrave Dictionary of Economics*. S. N. Durlauf and L. E. Blume, eds. London, UK, Palgrave Macmillan, Ltd.

Rutter, Michael (2006a). *Genes and Behavior: Nature--Nurture Interplay Explained*. Oxford, UK, Blackwell Publishers.

Rutter, Michael (2006b). "Implications of Resilience Concepts for Scientific Understanding." in *Resilience in Children*. B. M. Lester, A. Masten and B. McEwen, eds. Malden, MA, Blackwell Publishing**:** 1.

Ryden, A., Sullivan, M., Torgerson, J. S., Karlsson, J., Lindroos, A. K. and Taft, C. (2003). "Severe Obesity and Personality: A Comparative Controlled Study of Personality Traits." *International Journal of Obesity* **27**(12): 1534-1540.

Sackett, Paul R., Borneman, Matthew J. and Connelly, Brian S. (2008). "High Stakes Testing in Higher Education and Employment: Appraising the Evidence for Validity and Fairness." *American Psychologist* **63**(4): 215-227.

Sackett, Paul R., Hardison, Chaitra M. and Cullen, Michael J. (2004). "On Interpreting Stereotype Threat as Accounting for African American-White Differences on Cognitive Tests." *The American Psychologist* **59**(1): 7-13.

Safire, William (1989). "On Language; the Elision Fields." *New York Times*, 13 August 1989.

Sahm, Claudia R. (2007). "Does Risk Tolerance Change? University of Michigan.

Salgado, Jésus F. (1997). "The Five Factor Model of Personality and Job Performance in the European Community." *Journal of Applied Psychology* **82**(1): 30-43.

Samuelson, Paul A. (1956). "Social Indifference Curves." *The Quarterly Journal of Economics* **70**(1): 1-22.

Saucier, Gerard (2003). "An Alternative Multi-Language Structure for Personality Attributes." *European Journal of Personality* **17**(3): 179-205.

Savage, Leonard Jimmie (1954). *The Foundations of Statistics*. New York, Wiley.

Savelyev, Peter A. (2010). "Conscientiousness, Education, and Longevity of High-Ability Individuals." Unpublished manuscript. University of Chicago, Department of Economics.

Scarr, Sandra, Weinberg, Richard A. and Waldman, Irwin D. (1993). "IQ Correlations in Transracial Adoptive Families." *Intelligence* **17**(4): 541-556.

Schaie, Klaus Warner (1994). "The Course of Adult Intellectual Development." *American Psychologist* **49**(4): 304.

Schmeichel, Brandon J., Vohs, Kathleen D. and Baumeister, Roy F. (2003). "Intellectual Performance and Ego Depletion: Role of the Self in Logical Reasoning and Other Information Processing." *Journal of Personality and Social Psychology* **85**(1): 33.

Schmidt, FL, Hunter, JE and Larson, M (1988). "General Cognitive Ability Versus General and Specific Aptitudes in the Prediction of Training Performance: Some Preliminary Findings." Navy Personnel Research and Development Center

Schmidt, Frank L. and Hunter, John (2004). "General Mental Ability in the World of Work: Occupational Attainment and Job Performance." *Journal of Personality and Social Psychology* **86**(1): 162-173.

Schmidt, Frank L. and Hunter, John E. (1983). "Individual Differences in Productivity: An Empirical Test of Estimates Derived from Studies of Selection Procedure Utility." *Journal of Applied Psychology* **68**(3): 407-414.

Schmidt, Frank L. and Hunter, John E. (1998). "The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings." *Psychological Bulletin* **124**(2): 262.

Schoemaker, Paul J. H. (1982). "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations." *Journal of Economic Literature* **20**(2): 529-563.

Schooler, Carmi, Mulatu, Mesfin Samuel and Oates, Gary (1999). "The Continuing Effects of Substantively Complex Work on the Intellectual Functioning of Older Workers." *Psychology and aging* **14**(3): 483.

Schuerger, James M. and Witt, Anita C. (1989). "The Temporal Stability of Individually Tested Intelligence." *Journal of Clinical Psychology* **45**(2): 294-302.

Schulz, Richard, Bookwala, Jamila, Knapp, Judith E., Scheier, Michael and Williamson, Gail M. (1996). "Pessimism, Age, and Cancer Mortality." *Psychology and Aging* **11**: 304-309.

Schumpeter, J. A. (1954). *History of Economic Analysis*. New York, Oxford University Press.

Schwartz, Shalom H. (1992). "The Universal Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries." *Advances in Experimental Social Psychology* **25**: 1.

Schweinhart, Lawrence J., Montie, Jeanne, Xiang, Zongping, Barnett, W. Steven, Belfield, Clive R. and Nores, Milagros (2005). *Lifetime Effects: The High/Scope Perry Preschool Study through Age 40*. Ypsilanti, MI, High/Scope Press.

Segal, Carmit (2008). "Motivation, Test Scores, and Economic Success." Department of Economics and Business, Universitat Pompeu Fabra, Working Paper No. 1124.

Sell, Aaron, Tooby, John and Cosmides, Leda (2009). "Formidability and the Logic of Human Anger." *Proceedings of the National Academy of Science* **106**(35): 15073-15078.

Shamosh, Noah and Gray, Jeremy R. (2007). "Delay Discounting and Intelligence: A Meta-Analysis." Department of Psychology, Yale University, Unpublished manuscript.

Shifman, S., Johannesson, M., Bronstein, M., Chen, S. X. , Collier, D. A. , Craddock, N. J. , Kendler, K. S., Li, T., O'Donovan, M., O'Neill, F. A., Owen, M. J., Walsh, D., Weinberger, D. R., Sun, C., Flint, J. and Darvasi, A. (2008). "Genome-Wide Association Identifies a Common Variant in the Reelin Gene That Increases the Risk of Schizophrenia Only in Women." *PLoS Genetics* **4**(2).

Shiner, Rebecca and Caspi, Avshalom (2003). "Personality Differences in Childhood and Adolescence: Measurement, Development, and Consequences." *Journal of Child Psychology and Psychiatry* **44**(1): 2-32.

Shiner, Rebecca L. (1998). "How Shall We Speak of Children's Personalities in Middle Childhood? A Preliminary Taxonomy." *Psychological Bulletin* **124**(3): 308-332.

Siegler, Robert S. (1992). "The Other Alfred Binet." *Developmental Psychology* **28**(2): 179.

Sijtsma, Klaas (2009). "On the Use, the Misuse, and the Very Limited Usefulness Of cronbach's Alpha." *Psychometrika* **74**(1): 107-120.

Skiadas, C (1998). "Recursive Utility and Preferences for Information." *Economic Theory* **12**(2): 293-312.

Slonim, R and Roth, AE (1998). "Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic." *Econometrica* **66**: 569-596.

Slovic, Paul (1962). "Convergent Validation of Risk Taking Measures " *Journal of Abnormal and Social Psychology* **65**(1): 68-71.

Slovic, Paul (1972). "Psychological Study of Human Judgment: Implications for Investment Decision Making." *The Journal of Finance* **27**(4).

Smith, Mary Lee, Glass, Gene V. and Miller, Thomas I. (1980). *The Benefits of Psychotherapy*. Baltimore, Johns Hopkins University Press.

Social and Character Development Research Consortium (2010). "Efficacy of Schoolwide Programs to Promote Social and Character Development and Reduce Problem Behavior

in Elementary School Children (Ncer 2011–2001)." National Center for Education Research, Institute of Education Sciences, U.S. Department of Education

Spear, Linda Patia (2000a). "The Adolescent Brain and Age-Related Behavioral Manifestations." *Neuroscience & Biobehavioral Reviews* **24**(4): 417.

Spear, Linda Patia (2000b). "Neurobehavioral Changes in Adolescence." *Current Directions in Psychological Science* **9**(4): 111.

Spearman, Charles (1904). ""General Intelligence," Objectively Determined and Measured." *American Journal of Psychology* **15**: 201-293.

Spearman, Charles (1927). *The Abilities of Man; Their Nature and Measurement*. New York, The Macmillan Company.

Stankov, Lazar (2005). "g Factor: Issues of Design and Interpretation." in *Handbook of Understanding and Measuring Intelligence*. O. Wilhelm and R. W. Engle, eds. Thousand Oaks, CA, Sage Publications Inc.**:** 279.

Stanovich, Keith E. (1999). *Who Is Rational? : Studies of Individual Differences in Reasoning*. Mahwah, N.J, Lawrence Erlbaum Associates.

Starmer, Chris (2000). "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk." *Journal of Economic Literature* **38**(2): 332-382.

Steele, Claude M. and Aronson, Joshua (1998). "Stereotype Threat and the Test Performance of Academically Successful African Americans." in *The Black-White Test Score Gap*. C. Jencks and M. Phillips, eds. Washington, D.C., Brookings Institution Press**:** 401-427.

Steinberg, Laurence (2008). "A Social Neuroscience Perspective on Adolescent Risk-Taking." *Developmental Review* **28**(1): 78-106.

Steinberg, Laurence, Graham, S., O'Brien, L., Woolard, J., Cauffman, E. and Banich, M. (2009). "Age Differences in Future Orientation and Delay Discounting." *Child Development* **80**(1): 28-44.

Sternberg, Robert J. (2000). *Practical Intelligence in Everyday Life*. New York, NY, Cambridge University Press.

Sternberg, Robert J. (2001). "Successful Intelligence: Understanding What Spearman Had Rather Than What He Studied." in *Intelligence and Personality: Bridging the Gap in Theory and Measurement*. J. M. Collis and S. Messick, eds. Mahwah, NJ, Lawrence Erlbaum Associates**:** 347.

Stevens, Courtney, Fanning, Jessica, Coch, Donna, Sanders, Lisa and Neville, Helen (2008). "Neural Mechanisms of Selective Auditory Attention Are Enhanced by Computerized Training: Electrophysiological Evidence from Language-Impaired and Typically Developing Children." *Brain Research* **1205**: 55-69.

Sticht, Thomas G. (1982). "Literacy, Oracy, and Vocational Aptitude as Predictors of Attrition and Promotion in the Armed Services." Unpublished manuscript. University of Pennsylvania, Department of Psychology.

Sticht, Thomas G. (1995). "The Military Experience and Workplace Literacy: A Review and Synthesis for Policy and Practice." National Center on Adult Literacy, University of Pennsylvania Working paper, TR94-01.

Störmer, Susi and Fahr, René (2010). "Individual Determinants of Work Attendance: Evidence on the Role of Personality " IZA Discussion Paper No. 4927.

Sutter, Matthias, Feri, Francesco, Kocher, Martin G., Martinsson, Peter, Nordblom, Katarina and Rützler, Daniela (2010). "Social Preferences in Childhood and Adolescence: A Large-Scale Experiment." *IZA Discussion Paper No. 5016*.

Sylva, Kathy (1997). "The Quest for Quality in Curriculum." in *Lasting Differences: The High/Scope Preschool Curriculum Comparison Study through Age 23*. L. J. Schweinhart and D. P. Weikart, eds. Ypsilanti, High/Scope Press**:** 89-93.

Tang, Tony. Z., DeRubeis, Robert J., Hollon, Steven. D., Amsterdam, Jay., Shelton, Richard. and Schalet, Benjamin. (2009). "Personality Change During Depression Treatment: A Placebo-Controlled Trial." *Archives of general psychiatry* **66**(12): 1322.

Taubman, Paul J. and Wales, Terrence J. (1973). "Higher Education, Mental Ability and Screening." *Journal of Political Economy* **81**(1): 28-55.

Tausch, Franziska, Potters, Jan and Riedl, Arno (2010). "Preferences for Redistribution and Pensions: What Can We Learn from Experiments?" *IZA Discussion paper*.

Tellegen, Auke (1985). "Structures of Mood and Personality and Their Relevance to Assessing Anxiety, with an Emphasis on Self-Report." in *Anxiety and the Anxiety Disorders*. A. H. Tuma and J. D. Maser, eds. Hillsdale, NJ, Lawrence Erlbaum Associates**:** 681.

Terman, Lewis M. (1916). *The Measurement of Intelligence: An Explanation of and a Complete Guide for the Use of the Stanford Revision and Extension of the Binet-Simon Intelligence Scale*. Boston, Houghton Mifflin Co.

Terman, Lewis M. and Merrill, Maud A. (1960). *Stanford-Binet Intelligence Scale: Manual for the Third Revision Form L-M*. Boston, Houghton Mifflin.

Terman, Lewis M. and Oden, Melita H. (1947). *Genetic Studies of Genius Vol. Iv, the Gifted Child Grows Up*, Stanford University Press.

Thaler, Richard H (1980). "Toward a Positive Theory of Consumer Choice." *Journal of Economic Behavior & Organization* **1**(1): 39-60.

Thaler, Richard H and Ziemba, WT (1988). "Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries." *The Journal of Economic Perspectives* **2**(2): 161-174.

Thaler, Richard H. (2008). "A Short Course in Behavioral Economics."

Turkheimer, Eric, Haley, Andreana, Waldron, Mary, D'Onofrio, Brian and Gottesman, Irving I. (2003). "Socioeconomic Status Modifies Heritability of IQ in Young Children." *Psychological Science* **14**(6): 623-628.

Urzua, Sergio (2008). "Racial Labor Market Gaps: The Role of Abilities and Schooling Choices." *Journal of Human Resources* **43**(4): 919-971.

Uylings, Harry B. M. (2006). "Development of the Human Cortex and the Concept Of "Critical" Or "Sensitive" Periods." *Language Learning* **56**(Suppl 1): 59.

Valiente, Carlos, Lemery-Chalfant, Kathryn and Castro, Kimberly S. (2007). "Children's Effortful Control and Academic Competence." *Merrill-Palmer Quarterly* **53**(1): 1-25.

Valiente, Carlos, Lemery-Chalfant, Kathryn and Swanson, Jodi (2010). "Prediction of Kindergartners' Academic Achievement from Their Effortful Control and Emotionality: Evidence for Direct and Moderated Relations." *Journal of Educational Psychology* **102**(3): 550-560.

Valiente, Carlos, Lemery-Chalfant, Kathryn, Swanson, Jodi and Reiser, Mark (2008). "Prediction of Children's Academic Competence from Their Effortul Control, Relationships, and Classroom Participation." *Journal of Educational Psychology* **100**(1): 67-77.

van Eijck, Koen and de Graaf, Paul M. (2004). "The Big Five at School: The Impact of Personality on Educational Attainment." *Netherlands' Journal of Social Sciences* **41**(1): 24-40.

van Praag, Bernard M. S. (1985). "Linking Economics with Psychology: An Economist's View." *Journal of Economic Psychology* **6**(3): 289-311.

van Praag, Bernard M. S. and van Weeren, J. (1988). "Memory and Anticipation Processes and Their Significance for Social Security and Income Inequality." in *Applied Behavioral Economics Volume 2*. S. Maital, ed. New York, NY, New York University Press**:** 731-751.

Vaughn, Brian E., Kopp, Claire B. and Krakow, Joanne B. (1984). "The Emergence and Consolidation of Self-Control from Eighteen to Thirty Months of Age: Normative Trends and Individual Differences." *Child Development* **55**(3): 990-1004.

Vazsonyi, Alexander T. , Pickering, Lloyd E., Junger, Marianne and Hessing, Dick (2001). "An Empirical Test of a General Theory of Crime: A Four-Nation Comparative Study of Self-Control and the Prediction of Deviance." *Journal of Research in Crime and Delinquency* **38**: 91-131.

Viswesvaran, Chockalingam and Ones, Deniz S. (1999). "Meta-Analyses of Fakability Estimates: Implications for Personality Measurement." *Educational and Psychological Measurement* **59**(2): 197.

Vitaro, Frank, Brendgen, Mara, Larose, Simon and Tremblay, Richard E. (2005). "Kindergarten Disruptive Behaviors, Protective Factors, and Educational Achievement by Early Adulthood." *Journal of Educational Psychology* **97**(4): 617-629.

Wang, Lin (1993). *The Differential Aptitude Test: A Review and Critique*. Annual Meeting of the Southwest Educational Research Association, Austin, TX.

Watson, D., Clark, L. A. and Chmielewski, M. (2008). "Structures of Personality and Their Relevance to Psychopathology: II. Further Articulation of a Comprehensive Unified Trait Structure, John Wiley & Sons. **76:** 1545-1586.

Webb, Edward (1915). "Character and Intelligence." *British Journal of Psychology* **1**(3): 1.

Weber, Elke U. (2001). "Personality and Risk Taking." in *International Encyclopedia of the Social and Behavioral Sciences*. N. J. Smelser and P. B. Baltes, eds. Oxford, Pergamon**:** 11274.

Weber, Elke U., Blais, AnnRenee and Betz, Nancy E. (2002). "A Domain-Specific Risk-Attitude Scale: Measuring Risk Perceptions and Risk Behaviors." *Journal of Behavioral Decision Making* **15**(4): 263.

Weber, Elke U. and Hsee, Christopher (1998). "Cross-Cultural Differences in Risk Perception, but Cross-Cultural Similarities in Attitudes Towards Perceived Risk." *Management Science* **44**(9): 1205.

Weber, Elke U. and Milliman, Richard A. (1997). "Perceived Risk Attitudes: Relating Risk Perception to Risky Choice." *Management Science* **43**(2): 123.

Wechsler, David (1939). *The Measurement of Adult Intelligence*. Baltimore, MD, Williams & Wilkins Co.

Wechsler, David (1943). "Non-Intellective Factors in General Intelligence." *Journal of Abnormal and Social Psychology* **38**(1): 101-103.

Weikart, David P., Epstein, A. S., Schweinhart, L. and Bond, J. T. (1978). *The Ypsilanti Preschool Curriculum Demonstration Project: Preschool Years and Longitudinal Results*. Ypsilanti, MI, High/Scope Press.

Whisenton, Joffre T. and Lorre, M. Ray (1970). "A Comparison of the Values, Needs, and Aspirations of School Leavers with Those of Non-School Leavers." *Journal of Negro Education. Vol.* **39**(4): 325-332.

Widiger, Thomas A. and Costa Jr., Paul T. (2002). "Five-Factor Model Personality Disorder Research." in *Personality Disorders and the Five-Factor Model of Personality*. P. T. Costa Jr. and T. A. Widiger, eds. Washington, D.C., American Psychological Association.

Widiger, Thomas A. and Mullins-Sweatt, Stephanie N. (2009). "Five-Factor Model of Personality Disorder: A Proposal for DSM-V." *Annual Review of Clinical Psychology* **5**: 197-220.

Widiger, Thomas A., Trull, Timothy J., Clarkin, John F., Sanderson, Cynthia and Costa Jr., Paul T. (2002). "A Description of the Dsm-Iv Personality Disorders with the Five-Factor Model of Personality." in *Personality Disorders and the Five-Factor Model of Personality*. P. T. Costa Jr. and T. A. Widiger, eds. Washington, D.C., American Psychological Association.

Wilhelm, Oliver and Engle, Randall W. (2005). "Intelligence: A Diva and a Workhorse." in *Handbook of Understanding and Measuring Intelligence*. Thousand Oaks, CA, Sage Publications.

Willingham, Warren W. (1985). *Success in College: The Role of Personal Qualities and Academic Ability*. New York, College Entrance Examination Board.

Willingham, Warren W., Pollack, Judith M. and Lewis, Charles (2002). "Grades and Test Scores: Accounting for Observed Differences." *Journal of Educational Measurement* **39**(1): 1-37.

Willis, Robert J. and Rosen, Sherwin (1979). "Education and Self-Selection." *Journal of Political Economy* **87**(5, Part 2): S7-S36.

Winship, Christopher and Korenman, Sanders (1997). "Does Staying in School Make You Smarter? The Effect of Education on IQ in the Bell Curve." in *Intelligence, Genes, and Success: Scientists Respond to the Bell Curve*. B. Devlin, S. Feinberg, D. Resnick and K. Roeder, eds. New York, Springer, Copernicus**:** 215-234.

Wolfe, Raymond N. and Johnson, Scott D. (1995). "Personality as a Predictor of College Performance." *Educational and Psychological Measurement* **55**(2): 177-185.

Wong, Maria Mei-ha and Csikszentmihalyi, Mihaly (1991). "Motivation and Academic Achievement: The Effects of Personality Traits and the Duality of Experience." *Journal of Personality* **59**(3): 539-574.

Wood, Dustin and Roberts, Brent W. (2006). "Cross-Sectional and Longitudinal Tests of the Personality and Role Identity Structural Model (PRISM)." *Journal of Personality* **74**(3): 779-809.

Young, JW and Kobrin, JL (2001). "Differential Validity, Differential Prediction, and College Admission Testing: A Comprehensive Review and Analysis." College Entrance Examination Board

Zentner, Marcel and Bates, John E. (2008). "Child Temperament: An Integrative Review of Concepts, Research Programs, and Measures." *European Journal of Developmental Science* **2**(1-2): 7-37.

Zhou, Qing, Main, Alexandra and Wang, Yun (2010). "The Relations of Temperamental Effortful Control and Anger/Frustration to Chinese Children's Academic Achievement

and Social Adjustment: A Longitudinal Study." *Journal of Educational Psychology* **102**(1): 180-196.

Zigler, Edward and Butterfield, Earl C. (1968). "Motivational Aspects of Changes in IQ Test Performance of Culturally Deprived Nursery School Children." *Child Development* **39**(1): 1-14.

Zuckerman, Marvin (1994). *Behavioral Expressions and Biosocial Bases of Sensation Seeking*. New York, Cambridge University Press.

Zuckerman, Marvin (2007). *Sensation Seeking and Risky Behavior*. Washington, DC, American Psychological Association.