CHAPTER 1

Personality Psychology and Economics[1]

Mathilde Almlund[*], Angela Lee Duckworth[**], James Heckman[*,*], and Tim Kautz[*]**

[*]Department of Economics, University of Chicago, Chicago, IL 60637
[**]University of Pennsylvania, Philadelphia, Pennsylvania, PA 19104
[***]University College Dublin, American Bar Foundation

Contents

[1] This research was supported by grants from NIH R01-HD054702, R01-HD065072, and K01-AG033182; the University of Chicago; the Institute for New Economic Thinking (INET); A New Science of Virtues: A Project of the University of Chicago; the American Bar Foundation; a conference series from the Spencer Foundation; the JB & MK Pritzker Family Foundation; the Buffett Early Childhood Fund; the Geary Institute, University College Dublin, Ireland; and an anonymous foundation. The opinions expressed in this report are those of the authors and do not necessarily reflect the views of any of the funders. 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 A7.2) and health (presented in Web Appendix A7.1), 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 Amanda Agan, Dan Benjamin, Pietro Biroli, Dan Black, Daniel Cervone, Deborah Cobb-Clark, Flavio Cunha, Kathleen Danna, Thomas Dohmen, Steven Durlauf, Joel Han, Moshe Hoffman, John Eric Humphries, Miriam Gensowski, Bob Krueger, Jongwook Lee, Xiliang Lin, Dan McAdams, Terrance Oey, Lawrence Pervin, Pia Pinger, Armin Rick, Brent Roberts, Molly Schnell, Bas ter Weel, and Willem van Vliet. We also benefited from a workshop at the University of Illinois, Department of Psychology, on an early draft of this paper and presentations of portions of this paper at the Spencer/INET workshop at the University of Chicago, December 10–11, 2010, and at the IZA workshop on Cognitive and Non-Cognitive Skills, January 25–27, 2011, Bonn, Germany. Additional material that supplements the text is presented in a Web Appendix (http://jenni.uchicago.edu/personality_economics/). Parts of this paper build on an earlier study by Borghans, Duckworth, Heckman, and ter Weel (2008).

Handbook of the Economics of Education, Volume 4
ISSN 0169-7218, DOI: 10.1016/B978-0-444-53444-6.00001-8

Abstract

This chapter explores the power of personality traits both as predictors and as causes of academic and economic success, health, and criminal activity. Measured 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 traits and preferences. An extreme version of the situationist view claims that there are no stable personality traits or preference parameters that persons carry across different situations. Those who hold this view claim that personality psychology has little relevance for economics.

The biological and evolutionary origins of personality traits 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.

Measured 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 traits on economic and social outcomes. Personality traits are more malleable over the life cycle compared with cognition, which becomes highly rank stable around age 10. Interventions that change personality are promising avenues for addressing poverty and disadvantage.

Keywords

Personality
Behavioral Economics
Cognitive Traits
Wages
Economic Success
Human Development
Person-situation Debate

1. INTRODUCTION

The power of cognitive ability in predicting social and economic success is well documented.[2] Economists, psychologists, and sociologists now actively examine determinants of social and economic success beyond those captured by cognitive ability.[3] However, a substantial imbalance remains in the scholarly and policy literatures in the emphasis placed on cognitive ability compared to other traits. This chapter aims to correct this imbalance. It considers how personality psychology informs economics and how economics can inform personality psychology.

A recent analysis of the Perry Preschool Program shows that traits other than those measured by IQ and achievement tests causally determine life outcomes.[4] This experimental intervention enriched the early social and emotional environments of disadvantaged children of ages 3 and 4 with subnormal IQs. It primarily focused on fostering the ability of participants to plan tasks, execute their plans, and review their work in social groups.[5] In addition, it taught reading and math skills, although this was not its main focus. Both treatment and control group members were followed into their 40s.[6]

Figure 1.1 shows that, by age 10, the mean IQs of the treatment group and the control group were the same. Many critics of early childhood programs seize on this and related evidence to dismiss the value of early intervention studies.[7] Yet on a variety of measures of socioeconomic achievement, the treatment group was far more successful than the control group.[8] The annual rate of return to the Perry Program was in the range 6–10% for boys and girls separately.[9] These rates of return are statistically significant and above the returns to the US stock market over the postwar period.[10] The intervention changed something other than IQ, which produced strong treatment effects. Heckman, Malofeeva, Pinto, and Savelyev (first draft 2008, revised 2011) show that the personality

[2] See, e.g., the studies cited in Becker (1964) and the discussion of ability bias in Griliches (1977).

[3] See Bowles, Gintis, and Osborne (2001a) and Borghans, Duckworth, Heckman, and ter Weel (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, and ter Weel (2008) link economics with psychology.

[4] We draw on the research of Heckman, Malofeeva, Pinto, and Savelyev (first draft 2008, revised 2011). See Weikart, Epstein, Schweinhart, and Bond (1978); Sylva (1997); Schweinhart et al. (2005); and Heckman, Moon, Pinto, Savelyev, and Yavitz (2010a) for descriptions of the Perry program.

[5] Sylva (1997) shows that the Perry Program has important features that are shared with programs designed to foster self-control in children, e.g., *Tools of the Mind* (Bodrova and Leong, 2001).

[6] Plans are underway to follow the Perry sample through age 50.

[7] See the Westinghouse study of Head Start (Project Head Start, 1969).

[8] See Heckman, Malofeeva, Pinto, and Savelyev (first draft 2008, revised 2011) and Heckman, Moon, Pinto, Savelyev, and Yavitz (2010a).

[9] See Heckman, Moon, Pinto, Savelyev, and Yavitz (2010b).

[10] See DeLong and Magin (2009) for estimates of the return on equity.

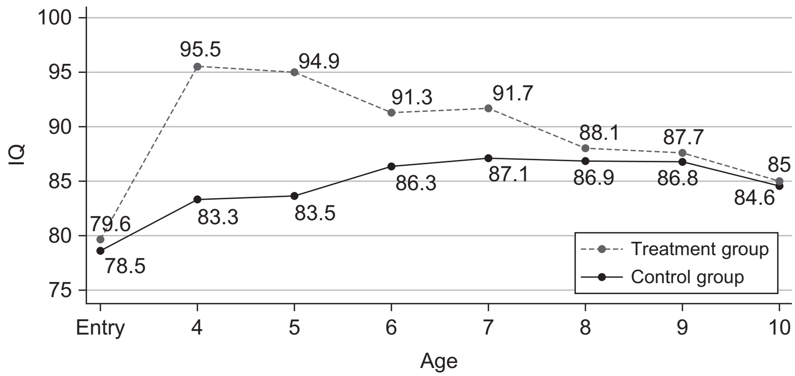

Figure 1.1 Perry Preschool Program: IQ, by Age and Treatment Group.

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

traits of the participants were beneficially improved in a lasting way.[11] This chapter is about those traits.

Personality psychologists mainly focus on empirical associations between their measures of personality traits and a variety of life outcomes. Yet for policy purposes, it is important to know mechanisms of *causation* to explore the viability of alternative policies.[12] We use economic theory to formalize the insights of personality psychology and to craft models that are useful for exploring the causal mechanisms that are needed for policy analysis.

We interpret personality as a *strategy function* for responding to life situations. Personality traits, along with other influences, produce measured personality as the output of personality strategy functions. We discuss how psychologists use measurements of the performance of persons on tasks or in taking actions to identify personality traits and cognitive traits. We discuss fundamental identification problems that arise in applying their procedures to infer traits.

Many economists, especially behavioral economists, are not convinced about the predictive validity, stability, or causal status of economic preference parameters or personality traits. They believe, instead, that the constraints and incentives in situations

[11] 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 for girls there was also improvement in Openness to Experience. See Heckman, Malofeeva, Pinto, and Savelyev (first draft 2008, revised 2011). 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.

[12] See Heckman (2008a).

almost entirely determine behavior.[13] This once popular, extreme situationist view is no longer generally accepted in psychology. Most psychologists now accept the notion of a stable personality as defined in this chapter.[14] Measured personality exhibits both stability and variation across situations.[15]

Although personality traits are not merely situation-driven ephemera, they are 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. 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, which we pose here and answer in the concluding section:

1. *How can we fit psychological constructs of personality into an economic framework? Can conventional models of preferences in economics characterize the main theories in personality psychology?*
2. *What are the main measurement systems used in psychology for representing personality and personality traits, and how are they validated? How are different systems related to each other? What is the relationship between standard measures of personality and measures of psychopathology and child temperament?*
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?*

This chapter is organized as follows. Section 2 presents a definition of personality that captures central ideas in the literature on personality psychology. It also presents a brief history of personality psychology and the person-situation debate that paralyzed the field for 20 years and that still influences behavioral economics. Section 3 defines measured personality as a response function using an economic model of preferences, expectations, and constraints. Our model distinguishes measured personality from personality traits. We interpret personality as a response function mapping variables

[13] See Thaler (2008) for an example of this point of view.

[14] See, e.g., Mischel and Shoda (1995, 2008).

[15] McAdams (2006, p. XVIII); Funder (2009); Mischel (2009); Roberts (2007, 2009); and Revelle, Wilt, and Condon (2011) discuss the stability question.

that characterize traits and situations to manifest (measured) personality. 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. We sketch a dynamic model of trait formation.

Section 4 discusses alternative criteria that psychologists use to define traits. It examines the strengths and limitations of each approach. We link our abstract definition to linear factor models that are commonly used to identify personality and cognitive traits.

Section 5 presents the main systems used to measure personality and cognition and discusses the relationship among the systems. We illustrate a nonidentification result developed in Section 3 by showing how scores on IQ tests are greatly affected by incentives and context. We present additional evidence showing that the scores on achievement tests depend on cognitive and personality measurements, with a substantial predictive role for personality measures. Measures of "IQ" commonly used in economics and social science conflate measures of cognition and personality.

Section 6 discusses economic preferences and examines the evidence relating economic preference parameters to psychological parameters. Section 7 surveys the evidence on the predictive validity of personality measures for education, crime, health, and labor market outcomes. The material presented in the main text summarizes a large and growing empirical literature. A Web Appendix presents additional detail on the literature relating cognition and personality in each of these areas of economic and social life.[16]

Section 8 presents evidence on the *causal* impact of personality on outcomes and evidence on the stability and malleability of personality traits and preferences. We extend the theoretical framework for trait formation introduced in Section 3 and discuss a corresponding measurement system. We discuss the evidence from intervention studies. Section 9 concludes with provisional answers to the eight questions.

2. PERSONALITY AND PERSONALITY TRAITS: DEFINITIONS AND A BRIEF HISTORY OF PERSONALITY PSYCHOLOGY

Personality psychology attempts to describe the whole person.[17] It considers both universal traits and individual differences. It examines the ways in which people are unique. As a sign of its breadth, personality psychology considers cognitive functioning as one aspect of personality.

[16] 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.

[17] Cervone and Pervin (2009) provide a clear introduction to personality psychology.

In considering the content of personality psychology, it is helpful to distinguish *personality traits, personality as a response function*, and *measured personality*. Personality is a response function that maps personality traits to measured (manifest) personality.

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, is used throughout the personality psychology literature.[18] We formalize these notions in Section 3.

Roberts' definition of personality traits refers to the stability of certain patterns of behavior—actions or responses to situations that people take, including patterns of thoughts or feelings. Perceptions, expectations of future events, and preferences may shape behavior, feelings, and thoughts. In this way, cognitive activities help to determine measured personality.

There are many different models of personality.[19] A prototypical model that captures many features of a wide class of models in personality psychology is one due to Roberts (2006). He presents the schematic displayed in Fig. 1.2 to relate personality traits to behavior.[20] He distinguishes mental abilities from personality traits (the items in the boxes will be discussed in later sections of this chapter). These, along with preferences (motives, interests, and values) and narratives (the stories people tell themselves in organizing their lives and making meanings of them), shape one's identity and reputation, including the views of the person by others and the person's perception of how others perceive him or her. Identity and reputation in turn shape the roles of individuals in the economy and the society and the larger culture to which they belong. *Personality* is the *system* of relationships that map traits and other determinants of behavior into measured actions.

In Roberts' vision of personality, feedback processes operate among all components of Fig. 1.2. Thus, his broad conception of personality includes the possibility that identity shapes traits and abilities, perhaps through a mechanism such as epigenetics, in which environment affects gene expression.[21] Measured personality results from interactions

[18] However, some personality psychologists use this or a very similar definition to define *personality* and not personality traits. 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). In this chapter, we define personality as a property of a system of equations, and measured personality is the output of those equations.

[19] See the models in John, Robins, and Pervin (2008).

[20] Graphical models like Fig 1.2 are the rule in personality psychology. Explicit formal models are rare. Section 3 presents a formal model.

[21] See, e.g., Rutter (2006a).

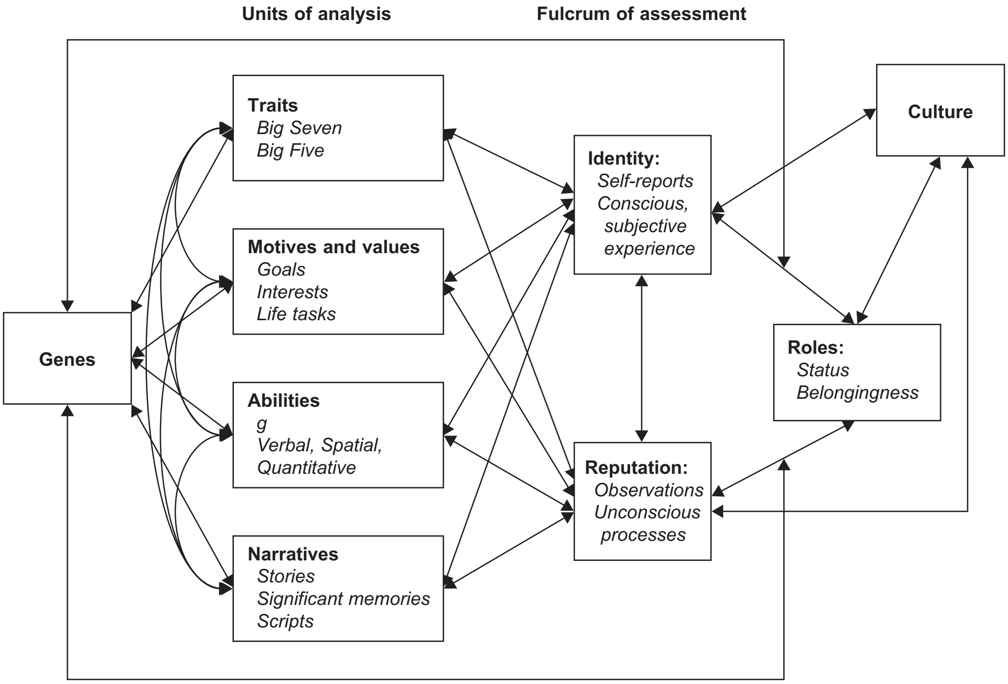

Figure 1.2 Roberts' Model of Personality as the Output of a System.

Source: Roberts (2006).

among components of the system. Personality traits are one determinant of personality and need to be carefully distinguished from the full expression of personality, which is generated by traits interacting with other factors. Personality is seen as a system of behaviors, thoughts, and feelings that emerge from the interacting components.

In Section 3, we formalize aspects of Roberts' framework for personality within an economic model of production, choice, and information. Figure 1.2 presages our discussion of a basic identification problem discussed in Section 3. Measurements and behaviors that arise from responses to incentives and interactions with culture are used to infer personality traits and abilities. Personality traits and cognitive abilities, along with the other "units of analysis" in Fig. 1.2, produce the observed behaviors that are used to infer the generating traits. To infer traits from behaviors requires "parsing out" or standardizing for all of the other contributing factors that produce the observed behavior—a challenging task. The inability to parse and localize behaviors that depend on a single trait or ability leads to a fundamental identification problem. Behavior depends on incentives created by situations. Accurately measuring personality traits requires standardizing for the situation.

2.1. A Brief History of Personality Psychology[22]

Interest in how individual behavior differs in common situations is as old as human history. The importance of personality traits for determining 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, is playing truant."*
>
> ***(Binet and Simon, 1916, p. 254)***

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."[23] 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).[24]

Arthur Jensen, an intellectual heir of Spearman, who is widely regarded as a proponent of *g* as an explanatory factor of success and failure in many domains of life, 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)***

2.1.1 The Pioneers of 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 behaviors and traits. Psychologists studied personality traits along with intelligence, interests, and motivation and measured differences and similarities within and across individuals.

[22] See Revelle, Wilt, and Condon (2011) for an informative history of personality psychology.
[23] Here and elsewhere through this essay, we capitalize personality traits.
[24] Many other psychologists who developed and promoted IQ tests expressed similar sentiments. See the Web Appendix Section A2.1.

A systematic approach to the study of personality was conceived by early psychologists 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), after eliminating synonyms and words not associated with traits. They designed and administered studies of trait inventories to large samples of individuals and applied the same factor analytic methods developed by Galton, Spearman, Binet, Pearson, Cattell, and Thorndike to these assessments in order to isolate *g* to identify the structure of cognitive abilities.

The fruits of several decades of research in this tradition beginning in the 1970s have produced a widely (but not universally) shared taxonomy of traits, known as the Big Five, that is arrived at through factor analysis of observer and self-reports of behaviors.[25] 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 levels.

Table 1.1 presents the Big Five traits. They are Openness to Experience (also called Intellect or Culture), Conscientiousness, Extraversion, Agreeableness, and Neuroticism (also called Emotional Stability).[26] The Big Five factors represent personality traits at the broadest level of abstraction. They summarize a large number of distinct, more specific, personality facets.

The Big Five traits are defined without reference to any context (i.e., situation). This practice leads to an identification problem that we discuss in Section 3. The behaviors

Table 1.1 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).

[25] Goldberg (1993), Barenbaum and Winter (2008), John and Srivastava (1999), and Krueger and Johnson (2008) discuss the Big Five.
[26] The acronym OCEAN is sometimes used to summarize these traits.

used to measure the traits are also determined by factors other than the Big Five traits. John (1990), Goldberg (1993), and Costa and McCrae (1992a) present evidence that most of the variables used to assess personality traits 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 traits, 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.

2.1.2 The Person-Situation Debate, Its Lingering Influence in Economics, and the Subsequent Flourishing of Personality Psychology

In 1968, Walter Mischel published a monograph entitled *Personality and Assessment*, challenging the most important theoretical assumptions and empirical findings of personality psychology. An acrimonious "person-situation" debate ensued, which pitted those who favored situational factors as explaining behavior against those who considered personality traits as more consequential. During this time, considered by many to be a dark age in the history of personality psychology, the general zeitgeist favored experimental social psychological approaches that focused on the importance of the situation compared to the individual traits featured in personality psychology.

Mischel noted that correlations between behavioral task measures of personality and questionnaire measures seldom, if ever, exceeded 0.3.[27,28] 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 wrote

> *"… 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 went on to write 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. His view was that global traits, in attempting to summarize behavioral dispositions without regard to situational contingencies, were "excessively crude, gross units

[27] There is great irony that none of the correlations of cognitive measures with outcomes that are reported in Table A1 in the Web Appendix are as high as 0.3, but no one questions the power of cognition in predicting outcomes in social life. Few studies in social psychology show correlations as high as 0.2 (Richard, Bond, and Stokes-Zoota, 2003).

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

to encompass adequately the extraordinary complexity and subtlety of the discriminations that people constantly make" (p. 301).

Mischel (2004) now suggests that behaviors can be consistent 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."[29] Indeed, in his view, the systematic variation of behavior across situations points to underlying motivations, beliefs, schemas, strategies, and other factors that collectively and interactively lead to coherence in any individual's measured personality. His revised view of personality is broadly consistent with Roberts' Fig. 1.2.

In Section 3, we formalize the "if-then" relationship using an economic model. We show 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 accepting evidence from personality psychology into economics. Like most heated debates in social science, this one occurred in the absence of much data. In Section 5, we discuss the body of evidence that has emerged over the past four decades on the existence of stable personality traits.

The debate over the relative importance of person and situation in the 1960s and 1970s reflected deeper currents in psychology and social science more generally, that still run strong. Behaviorism, associated with B. F. Skinner, was influential. It posited that experience explains all aspects of behavior. There was the widely held notion that situation and experience were all powerful—that people were born as blank slates.[30] This captured the interventionist spirit of the times. Interindividual heterogeneity in traits was ignored. 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. 14)*

Many behavioral economists hold a similar view, and they often appeal to Mischel as a guiding influence. For example, in a recent roundtable discussion, Richard Thaler 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)***

[29] That is, unobserved heterogeneity.
[30] Pinker (2002).

Many studies in behavioral economics attempt to establish inconsistency in behavior across situations, in violation of standard assumptions of stable preferences used in mainstream economics. For instance, several studies find very low correlations in risk-taking behavior across situations.[31]

Personality psychology survived the behaviorist assault and is a prospering field. A rich body of correlational evidence, which we summarize in Section 7, shows that for many outcomes, measured personality traits are predictive and, sometimes more predictive than standard measures of cognition, that traits are stable across situations, and situations also matter.

Mounting evidence that behavior has a biological basis suggests that personality is an important determinant of behavior. The evidence from behavioral genetics shows that measured personality traits are as heritable as cognitive traits. Studies in neuroscience show that alterations in brain structure and function through accidents, disease, and by experiments affect measured personality. They reinforce the evidence from heritability studies. This evidence and other evidence shows that measured personality is not situation-specific ephemera. We review this evidence in Section 8.

3. CONCEPTUALIZING PERSONALITY AND PERSONALITY TRAITS WITHIN ECONOMIC MODELS

Personality psychologists rarely use formal models to define or measure their constructs. In order to introduce their knowledge to economists, we formalize their frameworks. This makes the concepts of personality psychology more precise and provides a basis for measurement and policy analysis.

We introduce a series of progressively more comprehensive models to integrate concepts from personality psychology into economics.[32] Roberts' framework (Fig. 1.2) captures the main features of the influential models used in personality psychology. We use it as a point of departure. Psychology adds new and often more nuanced descriptions of human behavior to the standard descriptions used in economics.

In the nineteenth century, economics and psychology were closely aligned. Economists then spoke of the "hedonic calculus" used by people weighing choices.[33] One of the advances made in neoclassical economics in the first half of the twentieth century was to

[31] See, 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).

[32] Borghans, Duckworth, Heckman, and ter Weel (2008) develop a variety of economic models for integrating personality psychology into economic models. We build on their analysis. We review these frameworks in Section A3 of the Web Appendix.

[33] See, e.g., Schumpeter (1954).

focus on choices and the objective (easily measured) factors (such as prices and incomes) that determine choices. Revealed preference became a central tool of economics and was implemented using the marginal rate of substitution between choices—a key parameter that emerged from the neoclassical revolution.[34] This parameter did not require measurable utility or knowledge of the mental states of the agents making choices. Mental states and measureable utility, once the province of economists, were eliminated by Occam's Razor.

Measurable utility was used in utilitarian economics but fell out of favor (see Samuelson, 1956, and 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, 1954).

Most economists view mental states as unnecessary baggage except insofar as they affect choices. Thus, the traits, abilities, and narratives used by Roberts in Fig. 1.2 are of interest to most economists only if they affect choices through preferences, constraints, and effects on information processing capabilities. Motives and values are captured in part by economic preference parameters. Until recently, "happiness" and "aggregate utility," as well as other subjective mental states that do not affect behavior (choices), were considered uninteresting to most economists.[35]

Preferences, constraints, and expectations provide the most direct way to introduce psychological variables into economic models. We begin our analysis with a bare-bones approach that focuses on constraints. For example, cognitive and personality traits affect earnings capacity because they enhance productivity (see, e.g., Bowles, Gintis, and Osborne, 2001a), and at least up to a point, more of a trait can generate more resources that enlarge choice sets and hence directly affect behavior.

3.1. An Approach Based on Comparative Advantage

The Roy model (1951) of comparative advantage provides a useful starting point. Heckman, Stixrud, and Urzua (2006) use the Roy model to introduce psychological variables into the study of social and economic outcomes.[36] Personality traits are treated as endowments, and choices are determined by personality traits and other factors as they affect productivity in skills.

[34] See Hicks (1946).

[35] However, see 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.

[36] See Roy (1951), Heckman and Sedlacek (1985), and Heckman and Honoré (1990).

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 represented by θ and the "effort" they expend on the task, e_j:

$$P_j = \phi_j(\theta, e_j), \quad j \in \mathcal{J} = \{1,\ldots,J\}, \quad e_j \in \mathcal{E},\ \theta \in \Theta. \tag{1.1}$$

The traits are the endowments of agents that govern behavior. Examples of traits include height, personality characteristics, problem-solving ability, 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.

In much applied research, effort and traits are often assumed to be measured so that over the relevant range, assuming differentiability with respect to e_j and θ,

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

but neither condition is strictly required. Excess effort (overexertion; too much attention to detail) may be counterproductive so that function ϕ_j need 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. Different traits may have different productivities in different tasks, leading to comparative advantage in different tasks for people with different endowments.[37]

Efforts may complement traits $\left(\frac{\partial^2 \phi_j}{\partial e_j \partial \theta'} > 0\right)$ or may substitute for them $\left(\frac{\partial^2 \phi_j}{\partial e_j \partial \theta'} < 0\right)$. A variety of intermediate cases might exist where some effort-trait relationships are complementary and others are substitution relationships. Some people may solve complex math problems with no effort, whereas others may have to 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, and Tice (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, agents can productively engage in only one of the J tasks at any time. This restriction can be interpreted as a case in which effort can only be applied to a single task. A reward-maximizing

[37] Cattan (2011) shows that sociability has negative returns in some sectors but positive returns in other sectors.

agent with trait θ and endowment $\bar{e}$ faces the problem of picking the maximal task to perform, $\hat{j}$ where

$$\hat{j} = \underset{j \in \{1,\ldots,J\}}{\operatorname{argmax}} \{R_j \phi_j(\theta, \bar{e})\}. \tag{1.2}$$

In this case, θ and $\bar{e}$ play the same role. People with different effort and capability endowments will generally choose different tasks.[38,39] 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 for their shortfall in ability by exerting effort. For certain tasks (e.g., creating new branches of mathematics), there may be threshold levels of θ such that for $\theta < \overline{\theta}_j$, $\phi_j(\theta, e_j) = 0$ for all $e_j < \bar{e}$. (The person needs a given level of trait θ, no matter how hard they try.) The higher R_j, the more likely will the person choose to perform task j. The particular choice of which j to perform depends on the productivity of traits in different tasks.

3.2. Allowing for Multitasking

More generally, at a point in time, people may perform multiple tasks.[40] A less discrete version of the Roy model builds on the same foundations, allows people to perform multiple tasks at any time, and postulates that $\phi_j(\theta, e_j)$ is concave and increasing in e_j.[41] The agent chooses effort levels e_j across the J tasks to maximize total rewards:

$$\max_{\{e_j\}_{j=1}^{J}} \sum_{j=1}^{J} R_j \phi_j(\theta, e_j) \tag{1.3}$$

$$\text{subject to} \qquad \sum_{j=1}^{J} e_j = \bar{e}.^{42}$$

[38] A straightforward extension works with utilities and not rewards so we define utility $U(P_1,\ldots,P_J)$ 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 $U(d_P \odot P)$ 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.

[39] See Heckman, Stixrud, and Urzua (2006); Cattan (2011); and the evidence in Section 7.

[40] 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.

[41] Failure of concavity can take us back to case I.

[42] The first-order conditions for this problem are standard: $R_j \frac{\partial \phi_j}{\partial e_j} \geq \lambda$ and $e_j \geq 0$, $j = 1,\ldots,J$, where λ is the vector of multipliers associated with the effort constraint. Some people may allocate no effort to some tasks. P_j may be zero if $e_j = 0$, but this is not strictly required. Again, it is straightforward to generalize this reward function to a general utility function $U(P_1,\ldots,P_J)$.

As the reward for activity j, R_j, increases, everything else constant, the effort devoted to j will increase.[43,44] This model is consistent with effort that compensates for shortfalls in endowments, as well as effort that reinforces initial endowments. The choice of effort depends on the pattern of complementarity and substitutability. Different situations may be associated with different rewards for the same task. Such variation can produce differences in performance across tasks of the sort featured in the person–situation debate discussed in Section 2. One needs to standardize for the incentives to exert effort across tasks and differences in the endowments of effort in order to use measurements of performance on tasks to identify traits, θ.

3.3. Identifying Personality Traits

Before considering more general models, it is useful to discuss basic identification problems that arise in simple settings and in more general models. At the current level of generality, all traits can potentially affect productivity in all tasks. However, some tasks may require only a single trait or a subset of all of the traits. Following a traditional dichotomy in psychology that is explicit in Roberts' Fig. 1.2, divide θ into "mental," μ and "personality," π traits: θ_μ and θ_π, each of which may in turn be a vector.[45]

Psychological measurement systems sometimes use productivity measured in different tasks to identify θ_μ and θ_π.[46] This is the way Carroll (1993) defines mental ability where the task is performed on "mental" tests. To use performance on a task (or on multiple measures of the task) to identify a trait requires that performance on certain tasks (performance on a test, performance in an interpersonal situation, etc.) depends exclusively on one component of θ, say $\theta_{1,j}$. In that case,

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

Even if we can measure productivity P_j in task j, and only one component of θ affects P_j, to identify the level of a trait, one must control for the level of effort applied to j in order to use P_j to infer the level of $\theta_{1,j}$. That is, one must standardize for the effort at a benchmark level, say e^*, to use P_j to identify a measure of the trait that is uniform across different situations that elicit different levels of effort.[47]

The activity of picking a task (or a collection of tasks) to measure a particular trait ($\theta_{1,j}$ in our example) is called *operationalization* in psychology. *Construct validity* refers to

43 $\frac{\partial^2 \phi_j}{\partial\theta\partial e'_j} > 0$ 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.

44 In case I, agents will pick j.

45 Effort endowment might also be divided in the same fashion $(\bar{e}_\mu, \bar{e}_\pi)$ but we do not explicitly develop this possibility.

46 They also use observer reports and tests, which can be interpreted as observation on performance of tasks.

47 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.

whether or not a purported measure of the trait constructed in the stage of operationalization correlates with measures deemed to represent the trait. Considerable judgment is required to operationalize a trait and independently validate it. There is clear danger of circularity. Economists should carefully scrutinize how the measures they borrow from psychology are operationalized and validated in that literature. We should not necessarily assume that the measures created in that field have been rigorously established. We discuss how major constructs are validated in Section 5.

Assuming that construct validity has been established, if effort is involved in the performance of a task used to uniquely define a trait, the measurement of performance must be standardized in order to use measured productivity, P_j, to identify the trait. Otherwise, the endowment of effort and all of the factors that contribute to the exertion of effort, including the reward to the task, R_j, will contaminate the estimate of the trait. Failure to adjust for effort produces the kind of variability across situations with different rewards that was much discussed in the person–situation debate. We present examples of such contamination of measurement by the operation of incentives on effort in Section 5.

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

Consider the following case of two productivity measurements on tasks j and k:

$$\begin{aligned} P_j &= \phi_j(\theta_{1,\mu}, \theta_{1,\pi}, e_j) \\ P_k &= \phi_k(\theta_{1,\mu}, \theta_{1,\pi}, e_k), \quad j \neq k. \end{aligned}$$

One might have such measurements if data are available on the productivity of the same person performing two different tasks. Standardize measurements at a common level of effort, $e_j = e_k = e^*$.[48] If the functional forms of the $\phi_j(\cdot)$ and $\phi_k(\cdot)$ are known, and the system of equations satisfies a local rank condition, then one can solve for the pair ($\theta_{1,\mu}$, $\theta_{1,\pi}$) at e^*.[49]

48 Note that if the support of e_j and e_k is disjoint, no e^* exists, and hence, no such standardization is possible.

49 Let $\theta = (\theta_{1,\mu}, \theta_{1,\pi})$. Assume that the functional forms of $\phi_j(\cdot)$ and $\phi_k(\cdot)$ are known. Formally, a solution from P_j and P_k for $\theta_{1,\mu}$ and $\theta_{1,\pi}$ requires that the Jacobian of the system of equations for P_j and P_k,

$$\left[\frac{\partial \phi_j}{\partial \theta} \quad \frac{\partial \phi_k}{\partial \theta}\right]'_{e_j = e_k = e^*}$$

be nonvanishing in open neighborhoods around any point of solutions for θ (see, e.g., Buck, 2003).

The rank condition might not be satisfied, and the functional forms ϕ_j and ϕ_k might not be known. The productivity functions need not be monotone in $\theta_{1,\mu}$ or $\theta_{1,\pi}$. Interacting systems might produce multiple equilibria so that the same values of θ produce different values of (P_j, P_k).[50] Interacting systems might also have no solution.

Note that even if these problems do not arise, only the pair $(\theta_{1,\mu}, \theta_{1,\pi})$ 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 in which 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), an intrinsic identification problem arises in using measures of productivity in tasks to infer traits.[51] A dedicated measurement for at least one component is an essential requirement for identification. Other components can be defined relative to that measurement.[52,53]

3.4. Extensions of the Roy Model

Many empirical economists use the simple gross income maximizing framework of the Roy model to study the effects of personality on outcomes. The model is amended in many papers by including a cost $C_j(\theta, e_j)$ for obtaining rewards so that instead of criterion (1.2), the agent picks $\hat{j}$ that maximizes the net reward

$$\hat{j} = \underset{j\in\{1,\ldots,J\}}{\arg\max}\{R_j\phi_j(\theta, \bar{e}) - C_j(\theta, \bar{e})\}.$$

In the analogous extension for criterion (1.3), the agent maximizes

$$\sum_{j=1}^{J} R_j\phi_j(\theta, e_j) - C_j(\theta, e_j)$$

[50] Thus, there is a correspondence between (P_j, P_k) and θ, but no unique functional relationship.

[51] 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.

[52] This problem arises in linear factor models, but it is a more general problem. See, e.g., Anderson and Rubin (1956) for a definitive treatment of linear factor models. The scales in any factor model are arbitrary and are always defined with respect to a normalization (i.e., a dedicated measurement that defines the factor). The more general nonlinear model considered in the text faces the same problem.

[53] In general, without knowledge of the functional forms of the $\phi_j(\cdot)$, $j = 1,\ldots,J$, the problem of solving for two measurements P_j, P_k to infer $\theta_{1,\mu}$ and $\theta_{1,\pi}$ at a common level of $e_j = e_k$ is intractable. Many alternative solutions are possible. The traditional factor analysis literature reviewed in Section 5 below assumes linearity of the $\phi_j(\cdot)$, $j = 1,\ldots,J$. But even in that literature, attention focuses primarily on identifying the *distribution* of $(\theta_{1,\mu}, \theta_{1,\pi})$, not individual values $(\theta_{1,\mu}, \theta_{1,\pi})$ when P_j, $j = 1,\ldots,J$ is measured with error, although methods for solving for individual values of $(\theta_{1,\mu}, \theta_{1,\pi})$ and correcting for measurement error of the resulting estimates are available in the literature and are widely applied (see, e.g., Heckman, Malofeeva, Pinto, and Savelyev, first draft 2008, revised 2011; Savelyev, 2010; Heckman and Williams, 2011). Cunha, Heckman, and Schennach (2010) establish conditions under which it is possible to nonparametrically identify the functional form of $\phi_j(\cdot)$, $j = 1,\ldots,J$ and the distributions of $(\theta_{1,\mu}, \theta_{1,\pi})$ in the presence of measurement error on P_j, $j = 1,\ldots,J$.

with respect to the choice of e_j. 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 productivity traits from preference traits.[54]

The framework of the Roy model is widely used in recent analyses of the role of personality and cognition.[55] It has precedents in the work of Mandelbrot (1962), Heckman and Sedlacek (1985), and Heckman and Honoré (1990). In most applications, the $\phi_j(\theta, e_j)$ and $C_j(\theta, e_j)$ (or their logarithms) are assumed to be linear or log linear in θ and e_j. For example:

$$P_j = \alpha'_\theta \theta + \alpha'_e e_j$$
$$C_j = \beta'_\theta \theta + \beta'_e e_j.$$

The analyst models both the choice of the task and the output from the chosen task. A third (mixed) case can arise in which some clusters of tasks are mutually exclusive, so the agent can perform only one task within each cluster of tasks, but the agent can simultaneously engage in tasks across clusters.

3.5. Adding Preferences and Goals

Preferences and goals (see Fig. 1.2) may also shape effort.[56] 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 $e = (e_1, \ldots, e_J)$. Agents 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 final consumption goods X with associated prices W. A utility function can be specified over X, P, and e with preference parameter vector $\psi \in \Psi$.[57] Thus, we write

$$U(X, P, e; \psi), \tag{1.4}$$

[54] Heckman and Navarro (2007) and Abbring and Heckman (2007) present conditions for 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.

[55] See, e.g., Heckman, Stixrud, and Urzua (2006); Heckman, Humphries, Urzua, and Veramendi (2011); Báron and Cobb-Clark (2010); and Cattan (2011).

[56] In some versions of the preceding models with costs, preferences can be embodied in psychic costs.

[57] Robson (1996, 2001) and Robson and Samuelson (2007, 2009) discuss the evolutionary origin of preference parameters.

where the agent maximizes (1.4) subject to the constraints

$$Y + R'P = W'X, \tag{1.5}$$

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

$$\sum_{j=1}^{J} e_j = \bar{e}. \tag{1.6}$$

Preference specification (1.4) captures the notions that agents have preferences over goods, agents may value the output of tasks in their own right, and agents may value the effort devoted to tasks.[58]

The parameter ψ determines the trade-offs in preferences among X, P, and e. In one interpretation, subjective measures of well-being (Layard, 2005) attempt to directly measure (1.4).[59] Parameters that affect subjective well-being but not choices can be identified from the measures of well-being, but not from choices.[60]

3.6. Adding Learning and Uncertainty

All of the preceding models can be extended to account for learning and uncertainty. Let $\mathcal{I}$ be the information possessed by the agent, and let E denote mathematical expectation. An agent can be interpreted as making decisions based on

$$E[U(X, P, e; \psi); \mathcal{I}], \tag{1.7}$$

where ψ may be in the agent's information set (i.e., the agent knows his preferences).

Different theories specify different amounts of information available to agents. They might be uncertain about their preferences, ψ, traits, θ, the prices they face, W, the rewards to productivity, R, the outcomes of purchase decisions, X, and their endowments of effort, $\bar{e}$. The theory can be suitably modified to account for this uncertainty.

The use of the expectations operator begs the question of how agents construct the information set and how subjective expectations are formed. Psychological traits θ may affect information perception and processing. Several recent studies that apply personality traits to the economics of search suggest that agents with a higher perception of the

[58] Goods might also be direct arguments of the productivity functions, but, for simplicity, we do not analyze that case.

[59] However, the happiness literature is not strictly wedded to the notion that happiness is the same as our U, which is used only to characterize choice behavior.

[60] The model can readily be extended to cover more general cases. There is no need to impose the linear reward structure ($R'P$). The resources raised from productive tasks can be a nonlinear in P. Another simple extension of the model is the case in which there is no financial gain from engaging in tasks, but the agent receives a direct utility benefit from doing so. In this case, constraint (1.5) 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.

control they have over their lives have greater confidence in the arrival of job offers.[61] Overconfidence may be a trait that causes persons to inflate their perceived productivity.[62] A production function for information may depend on components of the trait vector, $\theta_{\mathcal{I}}$, 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. We discuss the evidence on how psychological traits affect information updating in Section 6.

One might object to the rationality and self-perception implicit in this formulation. As in Freud (1909, reprinted 1990), decision making might be made by a subconscious mind lacking self-perception. Decision making may be unconscious and agents may not recognize their desired goals. Nonetheless, constraints limit their revealed choice behavior. Borghans, Duckworth, Heckman, and ter Weel (2009) develop a model in which agents have random preferences and make choices at random within their feasible set. Variations in constraints drive the measured behavior of group averages but do not predict the behavior of any individual.

3.7. Definition of Personality within an Economic Model

Personality *traits* are the components of $\bar{e}$, θ, and ψ that affect behavior. One might define measured personality as the performance (the P_j) and effort (the e_j) that arise from solutions to any of the optimization problems previously discussed. Thus, the derived productivity and effort functions would constitute the systems generating measured personality as a response to constraints, information, and preferences, that is, as a system of functions that solve out for the P_j and e_j that agents choose in terms of their choice parameters.[63]

This approach to defining personality would not capture the full range of behaviors or *actions* considered by personality psychologists as constituting manifestations of personality. The actions considered by psychologists include a variety of activities that economists normally do not study, for example, cajoling, beguiling, bewitching, charming, etc. Thus, in selling a house, various actions might be taken, for example, smiling, persuading people by reason, threatening, scowling, showing affection, etc. Actions also include emotions, feelings, and thoughts and are not restricted to be activities that promote physical productivity. 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.

[61] McGee (2010) and Caliendo, Cobb-Clark, and Uhlendorff (2010).

[62] See, e.g., Akerlof and Dickens (1982); Caplin and Leahy (2001); Köszegi (2006); and Möbius, Niederle, Niehaus, and Rosenblat (2010).

[63] As previously noted in a simpler setting, no solutions may exist or multiple solutions may exist (so, there is a system of correspondences) between traits and personality outcomes.

To capture these more general notions, we introduce the concept of "actions" that are 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 i^{th} possible action to perform task j by $a_{i,j}$, $i \in \{1, \ldots, K_j\}$. Array the actions in a vector $a_j = (a_{1,j}, \ldots, a_{K_j,j}) \in \mathcal{A}$. 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:

$$P_j = \tau_j(a_{1,j}, a_{2,j}, \ldots, a_{K_j,j}). \tag{1.8}$$

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

$$a_{i,j} = \nu_{i,j}(\theta, e_{i,j}), \tag{1.9}$$

where

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

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 naturally gregarious person may find it easier to engage in social interactions than others. 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 arises 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 that gives him the highest productivity, and the other possible actions are not taken.

Agents may have utility over actions beyond the utility derived from consuming the outputs of tasks. For example, an agent may prefer accomplishing a task by working hard rather than by cheating. Different beliefs, thoughts, and feelings may have different effects on outcomes. Introducing actions in this fashion allows for the possibility that some actions are valued in their own right and do not directly contribute to productivity in any of the J tasks. Let M be the index set for the set of possible actions, including actions that do not directly contribute to productivity. In this more general formulation

$$a_{i,m} = \nu_{i,m}(\theta, e_{i,m}),\ m \in M,$$

where $\mathcal{A} \subseteq \mathcal{M}$.

We define utility over actions. Let a denote the choice of actions, some of which may not be associated with any particular task. Using the same information as used to characterize (1.7), the agent solves

$$\max E[U(a, X, P, e; \psi); \mathcal{I}]$$

with respect to X and e given the 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 ψ) and on the available information. Part of learning may consist of agents learning about the set of actions that are available to them, $\mathcal{A}(\mathcal{I})$.

One can extend the framework to introduce the effects of the situation in the person-situation debate, by considering specific situations represented 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 (1.9), the action function can be expanded to be dependent on situation h:

$$a_{i,j,h} = \nu_{i,j}(\theta, e_{i,j}, h), \tag{1.10}$$

and productivity on a task can be specified as a function of the action taken to perform the task in situation h:

$$P_{j,h} = \tau_j(a_{1,j,h}, \ldots, a_{K_j,j,h}) \tag{1.11}$$

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

$$P_{j,h} = \tau_j(\theta, a_{1,j,h}, \ldots, a_{K_j,j,h}, h).^{64} \tag{1.12}$$

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.[65] Persons taking an achievement test sometimes perform worse if they are told that their scores will influence social perceptions of their group as is found in the stereotype threat literature.[66]

The situation represents a key notion in the "person-situation" debate discussed in Section 2. Equations (1.10)–(1.12) capture the "if-then" notion of Mischel and Shoda (1995). Under specification (1.12), agents with the same actions, efforts, and 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.[67]

[64] A more general formulation would treat $h \in \mathcal{H}$ as mutually exclusive descriptions of situations and not claim to represent all situations by a base set of characteristics and would index all of the $\nu_{i,j}$ functions by h.

[65] Included in situation h might be the act of being observed by third parties and other possible sources of social interactions.

[66] Steele and Aronson (1998) and Sackett, Hardison, and Cullen (2004).

[67] At the cost of further notation, we could make the set of possible situations task specific.

Let $T \in \mathcal{T}$ be the vector of traits $(\theta, \psi, \bar{e})$. At any point in time, traits are endowments. In the general case, the solution to the constrained maximization problem involves choosing goods X, situation h, actions $a_{i,j}$, and efforts e_j, $j \in \{1, \ldots, J\}$ subject to the constraints. h is fixed if agents cannot choose the situation. For simplicity, we analyze this case. Relaxing this assumption is straightforward but is notationally more cumbersome.

In the case of fixed h, the solution to the maximization problem produces a set of response functions.[68] Preference parameters, ψ, characterize the trade-offs and goals that help shape manifest behavior. The agent's response functions (assumed to exist) are

$$X = X(R, W, T, h, Y, \mathcal{I}) \tag{1.13}$$

$$e = e(R, W, T, h, Y, \mathcal{I}) \tag{1.14}$$

$$a = a(R, W, T, h, Y, \mathcal{I}). \tag{1.15}$$

Productivity P across tasks is derived from the actions, efforts, and traits of the agents.[69]

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 system of functions. People may have different personalities depending on their trait endowments, constraints, and situations. Their actions—not the traits—constitute the data used to identify the traits.

Introducing actions widens the set of data from which one might 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.

As noted in the introduction to Section 2, many personality psychologists define personality as "enduring patterns of thoughts, feelings, and behaviors" that reflect tendencies of persons to respond in certain ways under certain circumstances. Our notion of action a is broad enough to encompass the wide array of behaviors considered by the personality psychologists. We previously defined personality traits T as generators of behavior.

One way to capture the notion of enduring actions is to average the a functions (1.15) for a person with a given trait vector $T = t$ over situations and efforts. Thus, for a given task j and trait vector t, the average action for information set $\mathcal{I}$ can be defined as

$$\bar{a}_{T,i,j,\mathcal{I}} = \int\limits_{S_{T,\mathcal{I}}(h,\, e_{i,j})} \nu_{i,j}(\theta, e_{i,j}, h)\; g\Big(h, e_{i,j} \,|\, T = (\theta, \psi, \bar{e}), \mathcal{I}\Big)\, \mathrm{d}h\, \mathrm{d}e_{i,j},$$

68 The same warnings as previously issued apply. No solutions may exist or they may be multiple valued.

69 In the case of h chosen, we get a system of derived demands for X, h, $a_{i,j}$, e_j.

where $\mathcal{S}_{T,\mathcal{I}}(h, e_{i,j})$ is the support of $(h, e_{i,j})$ given T and $\mathcal{I}$, and $g\left(h, e_{i,j} \mid T = (\theta, \psi, \bar{e}), \mathcal{I}\right)$ is the density of $(h, e_{i,j})$ given $T = (\theta, \psi, \bar{e})$ and information set $\mathcal{I}$. $\bar{a}_{T,i,j,\mathcal{I}}$ is the "enduring action" i of agents across situations in task j with information $\mathcal{I}$, that is, the average personality. Note that if $\nu_{i,j}$ is separable in T the marginal effect of personality trait vector θ is the same in all situations. One can define the "enduring traits" in a variety of ways, for example by averaging over tasks, j, situations, h, or both. Only under separability will one obtain the same marginal effect of θ. Epstein (1979) and a subsequent literature present evidence against nonseparability but in favor of an "enduring trait" that is common across situations.

3.8. 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, T, information, $\mathcal{I}$, situations, h, and actions, a, as state variables that 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 briefly discuss the dynamics of trait and state formation, leaving a more complete discussion to Section 8.

To capture the evidence from a large and growing literature, we consider the dynamic evolution of traits.[70] Let T^{v} be traits at age v, $v \in \{0, \ldots, V\} = \mathcal{V}$. Traits may change through family and self-investment (Cunha and Heckman, 2007, 2009), schooling, biology, or experience. Information $\mathcal{I}^{v}$ may be updated through various channels of learning. All task outputs, actions, and goods inputs may be time dated.

Investment in period v is an action or set of actions that an individual (or a person or group acting for the individual) may take in period v. 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^{v}) and through aspects of the situation in which the agent is found, h^{v} where h^{v} is the vector of attributes of the situation:

$$T^{v+1} = \eta^{v}(\underbrace{T^{v}}_{\text{self}-\text{productivity}}, \underbrace{IN^{v}}_{\text{investment}}, h^{v}), \quad v \in \mathcal{V}, \tag{1.16}$$

where the first set of arguments arises from self- and cross-productivity (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.[71]

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

[71] The actions taken by agents might also enter as arguments to this technology.

Note that if elements of T^v 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^v 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).[72]

As emphasized by Mischel and Shoda (1995) and Roberts and Jackson (2008), situations may change over time as a function of past actions, past situations, investment, information, and the like. We present this possibility by the following equation of motion:

$$h^{v+1} = \chi^v(h^v, IN^v, a^v). \tag{1.17}$$

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

Information $\mathcal{I}^v$ may change over the life cycle through experimentation and exogenous learning:

$$\mathcal{I}^{v+1} = \rho^v(\mathcal{I}^v, a^v, T^v, IN^v, h^v). \tag{1.18}$$

This learning mechanism incorporates the beliefs of agents about the available data. Thus, people may learn about their environments and themselves in part as a consequence of their own actions and in part as a consequence of the exogenous arrival of information. Equations of motion (1.16)–(1.18) are very general. We consider special cases of them used in the empirical literature in Section 8.

A rich and evolving literature investigates dynamic preferences when agents do not possess full knowledge of their future environments (see, e.g., Hansen, 2005; Hansen and Sargent, 2008; Rust, 2008; Epstein and Zin, 1989; Epstein and Schneider, 2003; Skiadas, 1998). That literature is too large to summarize in this chapter. Preferences need not be separable over time, and there may be time inconsistency of choices associated with hyperbolic discounting.[73] We discuss commonly used dynamic preference specifications in Section 6.

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

Personality psychologists generally do not present formal models. The formalization, to our knowledge, in this section is the first mathematically precise definition of personality traits and measured personality. The models we have sketched in this section capture central features of leading models in personality psychology.

[72] See Lear (2004). A habit can be defined as an effortless performance of a task, that is, an action that requires no effort. High levels of traits might allow people to perform actions effortlessly.

[73] See Kirby and Herrnstein (1995) and Gul and Pesendorfer (2004).

By its authors' own admission, the McCrae and Costa (2008) Five Factor Theory is not a fully articulated model. Their model 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 assumed to learn about their own traits, but precise learning mechanisms are not discussed. 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 $\mathcal{I}^v$ plays a central role in Five Factor Theory. People learn about their traits through actions and experience, but the exact mechanisms are not precisely formulated. Equation (1.18) captures these notions. Situations may also evolve as a function of actions and experience, but no role is assigned to investment in Five Factor Theory.

Thus, a restricted version of (1.17) 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, McCrae and Costa 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 investment or experience do not affect the evolution of traits. Thus, the arguments of equation (1.16) are suppressed, but traits may still exogenously evolve as a function of age and the biology of the individual. Even though traits evolve as part of an exogeneous maturation process, persons may learn about themselves (their traits) by taking actions and by being acted on by the external environment.

"Social cognitive" theories are rivals to trait theories based on the Big Five.[74] Albert Bandura, Daniel Cervone, and Walter Mischel are central figures in this literature. Roberts' diagram (Fig. 1.2) captures key aspects of this theory, and Roberts himself can be viewed as a member of both camps. 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 reject the "cognitive–noncognitive" distinction that is often used in economics. They view manifest personality as an outcome of cognitive processes. A major role is assigned to agency—individual goals and motives that produce actions. Their goals and motives are captured by our ψ. The arrival of information is captured by $\mathcal{I}$. Although the literature in personality psychology often contrasts these two schools of thought, they are not distinct to us. Only in one extreme version of the social-cognitive theory are traits entirely absent. In that version, agent behavior is entirely

[74] See Cervone and Pervin (2009).

shaped by situations. More generally, Mischel and Shoda (2008) focus on the role of situation in shaping actions, efforts, and productivities and allow for traits to influence actions. The "sociogenomic" model of Roberts and Jackson (2008) also considers the dynamics of personality formation.

Thus, both schools of thought accept specification (1.9) or its extension (1.10), and both would be comfortable with response systems (1.13)–(1.15). 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, whereas trait theorists typically 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 in producing traits. 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 not consider this possibility. Instead they emphasize self-learning about traits that evolve by fixed biological principles unrelated to the experiences of individuals.[75]

4. MEASURING PERSONALITY

Unlike other personal traits, like height or weight, personality traits cannot be directly measured. Observed productivities, efforts, and actions are used to infer traits. This leads directly to the analysis of latent variables and to factor models that underlie much of the analysis of trait psychology. This is an area where psychology and the econometrics of measurement error, and latent variables more generally, fruitfully interact. Factor models underlie the concepts of validity of measurements that are used in psychology.

4.1. Linear Factor Models

Linear factor models are widely used in personality psychology and in psychometric models for mental test scores. We review the use of these models in psychology. Versions are already in widespread use in economics.[76] To capture essential points, we abstract from many of the issues discussed in Section 3. We consider measurements arising from productivity in tasks. We thus focus solely on the outputs of tasks, abstracting from actions, efforts, and situations. With suitable extensions of the notation used here, we can extend the factor model to the more general models discussed in Section 3.

[75] Cervone (2004) contrasts the two schools of thought.

[76] See, e.g., Heckman, Stixrud, and Urzua (2006); Heckman, Humphries, Urzua, and Veramendi (2011); Piatek and Pinger (2010); Cattan (2011); and Cunha and Heckman (2008).

We assume additive separability of the arguments of Eq. (1.1). The stripped-down model writes task performance of person n on task j, $P_{n,j}$, based on traits arrayed in a vector T_n in the following manner:

$$P_{n,j} = \mu_j + \lambda_j' T_n + \Delta_{n,j},\ n = 1, \ldots, N,\ j = 1, \ldots, J, \tag{1.19}$$

where μ_j is the mean productivity in 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 $(L < J)$ for the factor model 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 (1.19) captures the notions that (a) latent traits, T_n, generate a variety of outcomes, (b) task outputs are imperfect measures of the traits, T_n, because $\Delta_{n,j}$ also determines task output, and (c) tasks other than tests or observer reports may also proxy the underlying traits, that is, latent traits generate both test scores and behaviors. A correlation of outcomes across tasks can arise because tasks depend on the same vector of traits.[77] Outcomes across tasks may be correlated even if the components of T_n are not.[78]

4.2. Discriminant and Convergent Validity

In this simplified framework, most personality psychologists 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.[79] 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 or measurements of productivities) is commonly used to test for construct validity. This approach exploits the notion that a particular battery of measurements captures a component of T_n, for example, $T_{n,l}$, $l = 1,\ldots,L$, and not other components. Many measurements may be taken on $T_{n,l}$, and having multiple measurements helps to control for measurement error.

77 The strength of the correlation depends in part on the magnitudes of λ_j and λ_k across the two tasks, j and k.

78 Cunha, Heckman, and Schennach (2010) present a nonparametric identification analysis for a general nonseparable model allowing for measurement error in measures of performance. In the notation of Eq. (1.19), they nonparametrically identify the distribution of T_n and the distribution of $\Delta_{n,j}$, $j = 1,\ldots,J$. The latter is identified without assuming full independence among the measurement errors.

79 Borsboom, Mellenbergh, and Van Heerden (2003) compare the approach taken in Section 3 of defining traits *a priori* within a model with the operationalist approach (Bridgman, 1959) of defining a trait by whatever measurements are available on it. Operationalism begs the questions that arise in operationalization and construct validity.

All measurements are really just outcomes on a type of task, although the effort applied may vary greatly across tasks. The literature in psychology usually assigns a special status to tests, self-reports, and observer reports of latent traits and also uses direct measures of productivity, such as supervisor ratings.[80] Behaviors, tests, observer reports, and self-reports all can be used to proxy the underlying traits. These include repeated measurements on the same types of assessment mechanisms, as well as measurements on different behaviors and assessments that are assumed to be generated by common traits.[81]

A standard approach to defining constructs in personality psychology is based on factor analysis. This approach takes a set of measurements that are designed to capture a construct and measures within-cluster and across-cluster correlations of the measurements to isolate latent factors $T_{n,l}$, $l = 1,\ldots,L$, 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 (e.g., success in college, performance on the job, earnings). This process led 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, which we discuss 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.[82] This fuels controversy among psychologists advocating competing taxonomies.

To state these issues more formally, let $P^q_{n,l}$ be the qth measurement on trait l for person n. Using a linear factor representation, the qth measurement of factor l for person n can be represented as

$$\begin{gathered} P^q_{n,l} = \mu^q_l + \lambda^q_l T_{n,l} + \epsilon^q_{n,l}, \\ q = 1,\ldots,Q_l, \quad n = 1,\ldots,N, \quad l = 1,\ldots,L. \end{gathered} \tag{1.20}$$

The factor $T_{n,l}$ is assumed to be statistically independent of the "measurement errors," $\epsilon^q_{n,l}$, $q = 1,\ldots,Q_l$. Different factors are sometimes assumed to be independent ($T_{n,l}$ independent of $T_{n,l'}$ for $l \neq l'$). The measurement errors (or "uniquenesses") are usually assumed to be mutually independent within and across constructs.[83]

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

$$P^q_{n,l} = \mu^q_l + (\lambda^q)' T_n + \epsilon^q_{n,l}, \quad q = 1,\ldots,Q_l, \tag{1.21}$$

80 See Groth-Marnat (2009).

81 Different measurements may load onto different traits.

82 Indeed, as documented in Section 7, the factors associated with personality are also correlated with some measures of cognitive factors, but not all.

83 The literature in economics relaxes the independence assumptions. See Cunha and Heckman (2008) and Cunha, Heckman, and Schennach (2010) and the literature they cite. They present conditions under which independence can be eliminated and identification of factors is still possible.

where λ^q is a vector with possibly as many as L nonzero components. The $\epsilon^q_{n,l}$ are assumed to be independent of T_n and mutually independent within and across constructs (l and l' are two constructs). The task has *discriminant validity* for trait l if the only nonzero component of λ^q is λ^q_l. The μ^q_l and λ^q_l can depend on measured characteristics of the agent, Q_n.[84] The task has *convergent validity* if measures within the construct are highly correlated.

More precisely, conventional psychometric validity of a collection of items or test scores for different constructs has three aspects. (1) Factor T_l for construct l is statistically independent of factor $T_{l'}$ for construct $l' \neq l$, *discriminant validity*.[85] (2) A factor T_l is assumed to account for the intercorrelations among the items or tests within a construct l. (3) Item-specific and random-error variance are low (intercorrelations among items are high within a cluster).[86] Criteria (2) and (3) define *convergent validity*.[87]

Oblique factor analysis picks factors and factor loadings that allow the factors to be correlated across traits. Its criterion is to maximize the correlation of measurements on a trait and minimize the correlation of measurements across traits, but not imposing that cross-trait correlation be zero. See Harman (1976) and Gorsuch (1983) for a discussion of alternative criteria in oblique factor analysis.

4.3. Predictive Validity

An alternative criterion for validating measurement systems is based on 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, on 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 $(\epsilon^q_{n,l} \neq 0)$. Proxies for $T_{n,l}$ can appear to be

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

[85] This is sometimes weakened to a condition of zero correlation.

[86] 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.

[87] Nothing in these standard 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 would produce outcomes and factors that satisfy the criteria. The multiple traits would be captured into a single factor. Only if different traits differentially affect different outcomes can one identify different traits.

[88] See http://www.hoganassessments.com/products_services/hpi.aspx and also Hogan and Roberts (2001).

separate determinants (or "causes") instead of surrogates for an underlying one-dimensional construct or factor. Thus, suppose that measurement system (1.20) is the correct specification and that a set of measurements display both convergent and discriminant validity. As long as there are measurement errors in the measures 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.[89] For this reason, it is necessary to correct for measurement error in using predictive validity to identify and measure traits.

A second problem with this approach to validation is reverse causality. This is especially problematic when interpreting correlations between personality measurements and outcomes. Outcomes may influence 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 solve this problem. For example, the anticipation of a future pay raise may increase present self-esteem.

Psychologists sometimes address the problem of reverse causality by using early measures of traits determined well before the outcomes are measured to predict later outcomes.[90] This approach is problematic if the traits the analyst seeks to identify evolve over time and the contemporary values of traits drive behavior. This practice trades a reverse causality problem with a version of an errors in variables problem. Early measures of the traits may be poor proxies for the traits that drive measured current behavior. In our review of the literature in Section 7, we distinguish studies that attempt to control for reverse causality and those that do not.

Heckman, Stixrud, and Urzua (2006) 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.

Many psychologists focus on prediction, not causality.[91] Establishing predictive validity will often be enough to achieve the goal of making personnel assignment and student placement decisions.[92] However, for policy analysis, including analyses of new programs designed to augment the skills of the disadvantaged, causal models are required in order to generate policy counterfactuals.[93]

89 This is a standard result in the econometrics of measurement error. See, e.g., Aigner, Hsiao, Kapteyn, and Wansbeek (1984).

90 This approach is based on the *post hoc ergo propter hoc* fallacy.

91 There is a long tradition in psychology of conducting predictive analysis based on factor analysis (see, e.g., the essays in Cudeck and MacCullum, 2007), but, to our knowledge, there is no systematic treatment of the problem of reverse causality in that field.

92 See, e.g., Hogan and Roberts (2001); and Hogan and Hogan (2007).

93 See Heckman (2008a).

The papers of Heckman, Stixrud, and Urzua (2006) and Cunha and Heckman (2008), develop frameworks for circumventing the problems that arise in using predictive validity to define and measure personality constructs. 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 frameworks for testing *discriminant validity*. They allow the factors across different clusters of constructs to be correlated and show how to test for the presence of correlations across the factors.

They use an extension of factor analysis to represent proxies for low-dimensional factors. They test for the number of latent factors required to fit the data and rationalize the proxies.[94] Generalizing the analysis of Hansen, Heckman, and Mullen (2004), Heckman, Stixrud, and Urzua (2006) allow for lifetime experiences and investments to determine, in part, the coefficients of the factor model and to affect the factor itself. Cunha, Heckman, and Schennach (2010) and Cunha and Heckman (2008) 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 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.

4.4. Faking

"Faking" may corrupt measurements designed to proxy latent factors. 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.[95] Subconscious motives to see themselves as virtuous may produce the same faking behavior, even when responses are anonymous. 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 literature on "stereotype threat" shows that the framing of an achievement test can affect the performance of the test taker.[96] Some evidence suggests that faking has a surprisingly minimal effect on predicting job performance.[97] Correcting for faking using

[94] Conti, Heckman, Lopes, and Piatek (2010) discuss alternative approaches to selecting the number of latent factors. See also Cragg and Donald (1997).

[95] See Viswesvaran and Ones (1999), Sternberg (2001), and Sternberg et al. (2000).

[96] See Steele and Aronson (1998) and Sackett, Hardison, and Cullen (2004).

[97] See Hough, Eaton, Dunnette, Kamp, and McCloy (1990); Hough and Ones (2002); and Ones and Viswesvaran (1998).

scales designed to measure deliberate lying does not seem to improve predictive validity.[98] Nevertheless, as noted in Section 3, when measuring cognitive and personality traits, one should standardize for incentives and environment.

The linear factor model does not capture a variety of interesting interactions among traits. Cunha, Heckman, and Schennach (2010) and the papers they cite develop a non-linear nonnormal factor analysis that allows for measurement errors to be correlated across measures and over time. We report estimates based on their nonlinear factor analyses in Section 8.

4.5. The Causal Status of Latent Variables

Some psychologists question the causal status of latent variables extracted from factor analyses of measurements across individuals.[99] Such factor analytic studies summarize inter-individual variation but do not necessarily inform analysts about the effects of exogenously changing the factor in producing outcomes across individuals. In addition, variations of traits within persons may have very different effects than variations across persons.

The distinction between the effects of changing traits within and across persons is traditional in econometrics.[100] Econometric models that capture this distinction could be fruitfully applied to psychology, as can hierarchical linear models.[101]

These methods do not address the deeper problem that most of the estimates of "the effects" of psychological traits on outcomes (either from "within" or "across" studies) have no causal status. Structural equation methods have been used to estimate causal relationships using cross-person variation. They rely on the usual toolkit of simultaneous equations exclusion restrictions to secure identification.[102] Standard experimental and econometric techniques for inferring causality from within-person changes have only recently been applied to estimate causal effects of personality.[103] We review this literature in Section 8.

5. IMPLEMENTING THE MEASUREMENT SYSTEMS

How do psychologists measure individual differences? In this section, we analyze the major measurement systems for cognition and personality. We examine the relative performance of cognition and personality in predicting a variety of outcomes. For cognition, there is a fairly well-established set of terminologies and conventions. Aptitude tests are designed to measure differences in the rates at which individuals learn (i.e., fluid intelligence).

[98] See Morgeson et al. (2007).
[99] See Borsboom, Mellenbergh, and Van Heerden (2003) and Cervone (2005).
[100] See, e.g., Mundlak (1978) and Hsiao (2003).
[101] See Raudenbush and Bryk (2001).
[102] See, e.g., Aigner, Hsiao, Kapteyn, and Wansbeek (1984) for a review of this classical literature.
[103] See, e.g., Cunha, Heckman, and Schennach (2010) and Heckman, Malofeeva, Pinto, and Savelyev (first draft 2008, revised 2011).

Achievement tests are designed to measure acquired knowledge (i.e., crystallized intelligence). For personality, a variety of alternative measurement systems are proposed, and this is a source of confusion. We attempt to compare and equate these systems of measurement. We link them to measures of childhood temperament and psychopathology, which are also used to describe individual differences. We note that the problems of operationalization and construct validity are present in analyzing any measures of traits.

5.1. Cognition

Intelligence (also called cognitive ability and general mental ability) is defined by psychologists to include 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 et al., 1996, p. 77).[104] These are clearly distinct traits, and the literature distinguishes more finely among them. The term "IQ" is often used synonymously with intelligence but in fact refers specifically to scores on intelligence tests. 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."[105]

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 more specific mental abilities loading on other factors.[106] *g* is widely interpreted as general mental ability.[107] An extreme version of *g*-theory that is no longer widely accepted is that *g* accounts for all the correlation among different tests of cognition.[108]

Psychometricians have expanded this notion to create a hierarchy of "orders." The order of a factor indicates its generality in explaining a variety of tests of cognitive ability deemed to satisfy construct validity. Tests have different emphases (e.g., verbal ability, numeracy, coding speed, and other tasks). A first-order factor is predictive in all cognitive tasks, $j = 1, \ldots, J$ in Eq. (1.19). In modern parlance, this general correlation is called "*g*" but it is no longer viewed as the sole predictor of cognitive test scores. A lower order factor is predictive of performance 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 1.3

[104] Psychologists have attempted to broaden the concept of intelligence beyond this list. Most notably, Gardner (2004) suggests that the notion of intelligence should also include creativity and the ability to solve practical, real-world problems. He includes in his theory of multiple intelligences, musical intelligence, kinesthetic intelligence, and interpersonal and intrapersonal intelligence, among others.

[105] See Wilhelm and Engle (2005).

[106] See Johnson, Bouchard, Krueger, McGue, and Gottesman (2004); Jensen (1998); Lubinski (2004); and Spearman (1904, 1927).

[107] See Gottfredson (2002).

[108] See, e.g., Carroll (1993).

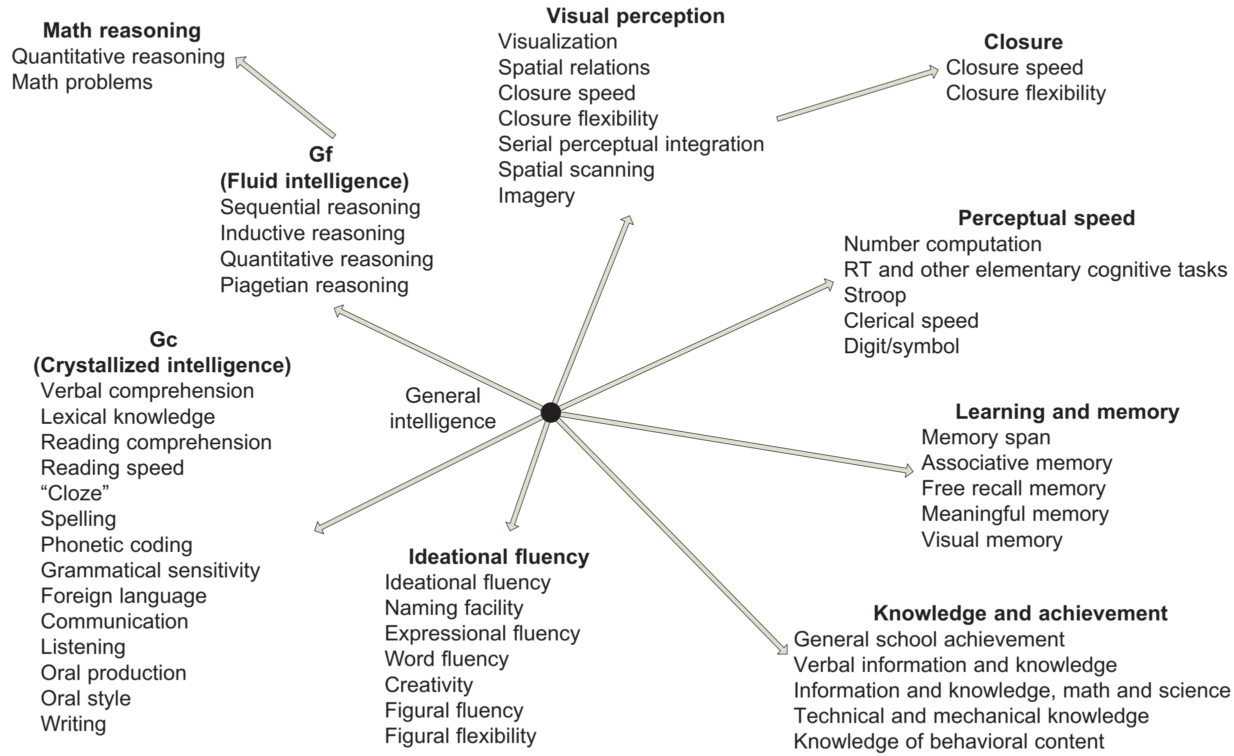

Figure 1.3 A Hierarchical Scheme of General Intelligence and Its Components.

Source: Recreated from Ackerman and Heggestad (1997).

reports one possible partition of general intelligence due to Ackerman and Heggestad (1997), who summarize the work of Carroll (1993) on the multiple facets of intelligence.[109]

5.1.1 Fluid versus Crystallized Intelligence

There is less agreement about the number and identity of lower order factors.[110] Carroll (1993) proposed a general intelligence factor *g* and several more specific second-order factors, including, but not limited to, what Cattell (1971, 1987) dubbed crystallized and fluid intelligence. Crystallized intelligence, Cattell proposed, comprises acquired skills and knowledge and, thus, is partly dependent on 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.[111] Two pieces of evidence are worth highlighting. First, crystallized intelligence tends to increase monotonically for most of the life cycle, whereas fluid intelligence tends to peak in very early adulthood then to decline.[112] 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.[113] 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 second-order

[109] Carroll's own organization of his evidence is somewhat different. See Carroll (1993, p. 626).

[110] Carroll (1993) 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 (e.g., Lubinski, 2004, and Horn, 1970).

[111] 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.

[112] See McArdle, Hamagami, Meredith, and Bradway (2000).

[113] See Dickens and Flynn (2001).

factors are not only correlated with the first-order factor g but also contribute additional explanatory power to predicting some clusters of test score outcomes. Achievement tests, such as the Armed Forces Qualifying Test used by economists and psychologists alike, are heavily weighted toward crystallized intelligence,[114] whereas tests like the Raven Progressive Matrices (1962) are heavily weighted toward fluid intelligence.[115] Several studies have shown that fluid intelligence is much more strongly related to g than are measures of crystallized intelligence.[116] Moreover, lay intuitions of intelligence (i.e., what most people mean by "being smart") correspond more closely with the ability to learn than with possession of already acquired knowledge.[117] 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. Some would argue that g has been usurped by fluid intelligence. A closer reading is that what is commonly meant by intelligence encompasses a number of distinct traits captured in the lower order factors of Fig. 1.3.

5.1.2 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.[118] Different tasks in life require different skills in different degrees.[119] Table 1.2 shows the domains of validation and the estimated validities of a number of widely used tests of cognition. Note 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.[120]

The SAT college entrance exam is moderately successful in predicting grades in college, which the SAT was designed to do.[121] However, high-school grades are better

[114] See Roberts et al. (2000).

[115] See 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% 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. Common developmental factors may affect both of these traits and that fluid intelligence promotes the acquisition of crystallized intelligence.

[116] See Cattell (1971, 1987), Gustafsson (1988), and Kvist and Gustafsson (2008).

[117] See Gottfredson (1998).

[118] See Mandelbrot (1962).

[119] See, e.g., Roy (1951); Mandelbrot (1962); Willis and Rosen (1979); Heckman and Sedlacek (1985); and Heckman, Stixrud, and Urzua (2006).

[120] See Kuncel, Ones, and Sackett (2010).

[121] See Young and Kobrin (2001).

Table 1.2 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	0.35–0.53	Kobrin, Patterson, Shaw et al. (2008)	
ACT	Grades in early years of college	0.42	ACT, Incorporated (2007)	
Stanford–Binet	Correlations with other intelligence tests	0.77–0.87 with WISC-R	Rothlisberg (1987) and Greene, Sapp, and Chissom (1990)	
WISC (Wechsler Intelligence Scale for Children)	Correlations with academic achievement	WISC: 0.443–0.751 with WRAT tests, 0.482–0.788 with first-grade grades, 0.462–0.794 with second-grade grades; WISC-R: 0.346–0.760 with WRAT tests, 0.358–0.537 with first-grade grades, 0.420–0.721 with second-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, 0.38–0.43 with college grades, 0.62 with high-school grades	Feingold (1982)	

Continued

Table 1.2 Predictive Validities of Various Tests of Fluid and Crystallized Intelligence—continued

Cognitive Achievement and IQ Tests

Test	Domain over Which It Is Validated	Estimated Validities	Source	Notes
Raven's Standard Progressive Matrices	Correlations with other intelligence tests	0.74–0.84 with WAIS-R	O'Leary, Rusch, and Guastello (1991)	
GATB (General Aptitude Test Battery)	Supervisor rating performance in training programs and in job performance	0.23–0.65	Hunter (1986)	Large range due to variety of jobs
ASVAB (Armed Services Vocational Aptitude Battery)	Performance in military training programs and military attrition rates	0.37–0.78 for training (mean = 0.56); −0.15 for attrition	Schmidt, Hunter, and Larson (1988) for performance in training programs; Sticht, Hooke, and Caylor (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	0.33–0.49 for HS Senior GPA	Technical Manual: 2002 Series GED Tests	
DAT (Differential Aptitude Tests)	Correlations with academic achievement	0.13–0.62 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 (1996)	CAT, California Achievement Test

predictors of college performance.[122] The rival American College Test (ACT) is validated in a similar fashion but uses broader measures of college performance, such as grades in higher years of college rather than freshman year grades.[123] The Graduate Record Exam is validated by performance in graduate school.[124] 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 such as fixing a rifle or repairing a radio.[125] One can interpret *The Bell Curve* by Herrnstein and Murray and the flood of papers it stimulated as conducting validity studies of the AFQT using real-world outcomes of the sort studied in Tables A1 and A2 in the Web Appendix. 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.[126]

5.2. Personality Traits

We have noted in Sections 2 and 3 that sharp contrasts between cognition and personality are not easy to make. Consider, for example, the 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 et al., 2000). Furthermore, the Big Five factor of Openness to Experience has as facets curiosity ("ideas") and imagination ("fantasy") that are often associated with intellect and measured intelligence.[127] (See the entries under Openness in Table 1.3.) We note in Section 5.3 that personality can affect performance on tests of fluid intelligence. Personality traits also affect acquired skills and knowledge (i.e., crystallized intelligence).[128] A general pattern is higher correlation of personality tests with tests of crystallized knowledge (e.g., achievement tests).[129] For many personality traits and for measures of cognition that are based on fluid intelligence, the correlations are close 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

[122] See Bowen, Chingos, and McPherson (2009a) and Geiser and Santelices (2007). However, there is a potential problem with restriction on the range in many of these studies.
[123] See ACT, Incorporated (2007).
[124] See Kuncel and Hezlett (2007).
[125] See McHenry, Hough, Toquam, Hanson, and Ashworth (1990).
[126] Schmidt and Hunter (1983, 1998); Hartigan and Wigdor (1989); and McHenry, Hough, Toquam, Hanson, and Ashworth (1990).
[127] McCrae and Costa (1997a) and Noftle and Robins (2007).
[128] See Chamorro-Premuzic and Furnham (2005) for an extended discussion of this topic.
[129] See, e.g., Borghans, Golsteyn, Heckman, and Humphries (2011).

Table 1.3 The Big Five Domains and Their Facets

Big Five Personality Factor	American Psychology Association 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) Straightforwardness (not demanding) Altruism (warm) Compliance (not stubborn) Modesty (not show-off) Tender-mindedness (sympathetic)	Empathy Perspective taking Cooperation Competitiveness	Irritability* Aggressiveness Willfulness

Continued

Table 1.3 The Big Five Domains and Their Facets—continued

Big Five Personality Factor	American Psychology Association Dictionary Description	Facets (and Correlated Trait Adjective)	Related Traits	Childhood Temperament Traits
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, 1983).

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

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

order to execute a less familiar behavior.[130] Some authors (e.g., Gray, 2004) also use the terms "cognitive control" and "self-control" interchangeably, though self-control is traditionally considered a personality trait rather than an aspect of cognition. Although tasks requiring executive function are related to questionnaire measures of self-control, the size of these associations is only about $r = 0.11 - 0.14$ (Duckworth and Schulze, 2009).

A region of the brain called the dorsolateral prefrontal cortex (PFC), in conjunction with the nearby region called the anterior cingulate cortex (ACC), are now understood as responsible for "executive control" over lower order processes.[131] 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, and Durston, 2005). Specific executive functions attributed to the PFC include abstract reasoning, planning, decision making, working memory (the ability to keep the facts of a problem at hand), attention, conflict monitoring, task switching, and inhibition of prepotent (i.e., dominant, habitual) impulses. Although many functions have been attributed to the PFC, Miller (2000) notes 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, and Miyake, Friedman, Emerson, Witzki, and Howerter, 2000). This distinction is important because working memory is highly related to performance on measures of fluid intelligence.[132] Being able to access all of the data about a problem is helpful in solving it. Thus, working memory is a common component of the constructs of both executive function and general intelligence.[133]

Although the construct of executive function demonstrates the inadequacy of terms such as "cognitive" and "noncognitive," many personality traits nevertheless are conceptually and empirically easily distinguished from general cognitive ability. Most personality traits are in fact 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 ability entails more than intelligence. Personality traits, however defined, do matter, and they have independent predictive power from standard measures of intelligence.

[130] Matsumoto and Tanaka (2004).
[131] Notably, the volume of dorsolateral prefrontal cortex (PFC) is correlated with Big Five Conscientiousness (DeYoung et al., 2010).
[132] Carpenter, Just, and Shell (1990) and Heitz, Unsworth, and Engle (2005).
[133] Friedman, Miyake, Corley et al. (2006).

5.3. Operationalizing the Concepts

Intelligence tests are routinely used in a variety of settings including business, education, civil service, and the military.[134] Psychometricians attempt to use test scores to measure a factor (a component of T in the notation of Section 4). The working hypothesis in the intelligence testing business is that specific tests measure only a single component of T 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.[135]

5.3.1 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.[136] Other pioneers in intelligence testing include Cattell (1890) and Galton (1883), both of whom developed tests of basic cognitive functions (e.g., 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 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 on formal education and cultural exposure. 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.

[134] Siegler (1992) provides a detailed overview of the different types of applications of psychological testing.

[135] See Roberts, Markham, Matthews, and Zeidner (2005) for a more complete history of intelligence testing.

[136] 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.

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.[137] The test taker must choose the missing part.[138] 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 exposure to such abstract puzzles (Blair, 2006). Our view is that to varying degrees, IQ and achievement tests reflect fluid intelligence, crystallized intelligence, and personality factors, such as motivation, to succeed on the test. We offer evidence on the effect of motivation on test scores below in Section 5.6.

5.4. 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 (e.g., Webb, 1915), even the most parsimonious personality models incorporate more than one factor. The most widely accepted taxonomy of personality traits is the Big Five.[139] The Big Five factors are obtained from conventional factor analysis using a version of Eq. (1.19) in which 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 1.3 presents the Big Five factors that were discussed 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). The acronym is 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 are the contexts in which the trait is predictive. In parentheses in the third column of Table 1.3, we have included a strongly related trait adjective. In the fourth column of Table 1.3, we present

137 See John and Srivastava (1999) for a discussion of the Raven test.
138 See Fig. A1 in Section A5 of the Web Appendix.
139 See Goldsmith et al. (1987) for an historical overview of the development of the Big Five.

other traits in each family. In the fifth column, we relate the Big Five to children's temperament traits studied by developmental psychologists.

Temperament is the term used by developmental psychologists to describe the behavioral tendencies of infants and children.[140] 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.[141] 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, whereas personality is studied by adult personality psychologists. However, the past decade has seen some convergence of these two research traditions, and there is evidence that temperamental differences observed during the preschool years anticipate adult personality and interpersonal functioning decades later (e.g., Caspi, 2000; Newman, Caspi, Moffitt, and Silva, 1997; Shiner and Caspi, 2003). The traits displayed in column 5 of Table 1.3 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.[142] 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 have been proposed that group lower order traits into higher order dimensions; several of these taxonomies resemble the Big Five (e.g., John, Caspi, Robins, and Moffitt, 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

[140] 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.

[141] 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.

[142] 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 (e.g., 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 (e.g., 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 the absence of large-sample studies of child temperament.

interested in documenting changes in personality over the full life cycle. Developing the required measures is an active area of research.

5.5. 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 1.4 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). In addition to these controversies, the facet-level organization of any given Big Five factor is subject to debate and controversy.

Recent research suggests that the rush to accept the Big Five may be premature.[143] The first studies of the Big Five were based primarily on English-speaking samples. Although the Big Five structure appears to replicate across many cultures (McCrae and Costa, 1997b), studies of more diverse cultures show that taxonomies known as the Big Six (Ashton et al., 2004) or the Multi-Language Seven (ML7; Saucier, 2003), may better represent the personality domain. Although they add one or two dimensions to the Big Five and shift the meaning of the Big Five slightly, they are, however, not very different from the Big Five.[144]

One of the most stinging criticisms of the Five-Factor model is that it is atheoretical (Block, 1995). It is derived from factor analysis of a variety of measures without any firm biological underpinnings. Although research is under way on determining the neural substrates of the Big Five (see Canli, 2006, and DeYoung 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 any scientifically grounded theory.

Some psychologists suggest that the categories are too crude to be useful. Estimates based on the Big Five factors obscure relationships between specific facets of the Big Five and outcomes.[145] Given that each Big Five factor is a composite of distinct facets, the predictive validities are diluted when analyses consider only factor-level aggregate scores. For instance, Paunonen and Ashton (2001) compared Big Five Conscientiousness and Openness to Experience with two related facets, need for achievement and

[143] This discussion draws heavily on Roberts (2006).
[144] See Roberts (2006) for a description of these shifts.
[145] Hough (1992), Hough and Oswald (2000), and Paunonen and Ashton (2001).

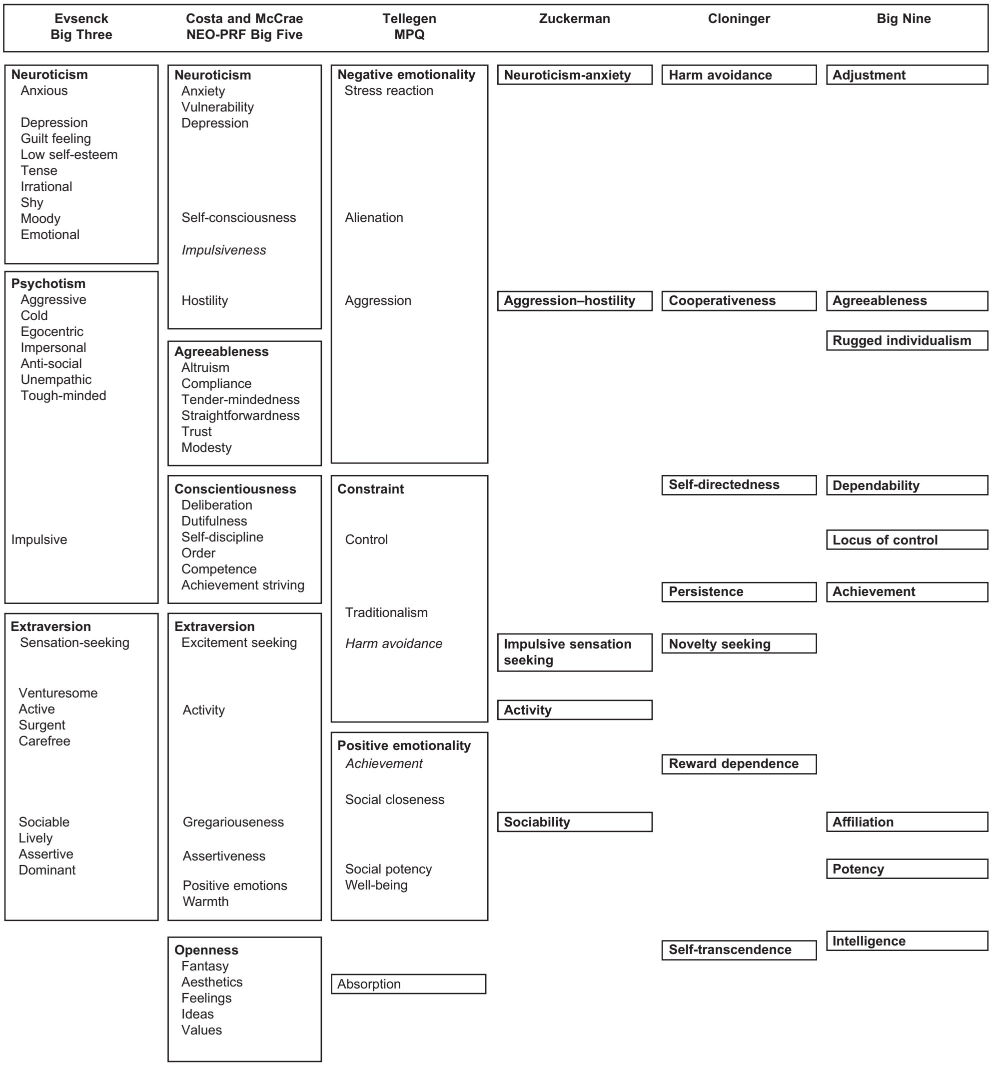

Figure 1.4 Competing Taxonomies of Personality.

Source: Reproduced from Bouchard and Loehlin (2001), with kind permission from Springer Science and Business Media.

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 about motivation. In the notation of Section 3, ψ parameterizes preferences and goals. The omission of motivation (i.e., 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 were 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 10 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, Wood, and Webb (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 distinguished in the literature.[146] 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, Quilty, and Peterson, 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, and Goldberg, 2005), and still others suggest that it is a blend of Conscientiousness, Extraversion, and perhaps Neuroticism (Revelle, 1997). Figure 1.4 shows in italics facets of the Big Five whose classifications are under question. One reason for the proliferation of measures is the variety of alternative methodologies for verifying tests discussed in Section 4, which are not guaranteed to produce the same taxonomies.

5.5.1 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 since they are collected in large-sample longitudinal studies used by economists.[147] They are not part of the traditional Big Five typology. However, they can be related to it.

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

[146] See, e.g., Carroll (1993).

[147] See, e.g., NLSY79-based studies Heckman, Stixrud, and Urzua (2006) and some of the models in Heckman, Humphries, Urzua, and Veramendi (2011). The German Socioeconomic Panel (GSOEP) also collects these measures.

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 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, and Locke, 2005; Judge, Erez, Bono, and Thoresen, 2002; Judge and Hurst, 2007) 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, and discriminant validity. There is a negligible gain in predictive power of adding components of the construct beyond the predictive power of the construct itself. 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 1.3, associated aspects of core self-evaluations with the Big Five factors of Neuroticism and Emotional Stability.

5.5.2 Relating the Big Five to Measures of Psychopathology

Extreme manifestations of personality traits may be a form of mental illness. Thus, a very conscientious person may be viewed as an obsessive-compulsive person. It is of interest to consider how psychopathology may be characterized using the Big Five.

Psychopathology is defined by the APA dictionary as "patterns of behavior or thought processes that are abnormal or maladaptive." Used interchangeably with the terms mental illness and mental disorder, psychopathology is primarily 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. Very recently, however, several attempts have been made to integrate taxonomies of psychopathology and normal personality into a single framework. In particular, a compelling argument can be made for conceptualizing and measuring mental disorders as extreme variants of personality traits (see Krueger and Eaton, 2010, and ensuing commentary). This approach is quite revolutionary in the study of psychopathology in at least two ways. First, it takes a dimensional as opposed to categorical characterization of mental disorders. By a dimensional approach, psychologists mean that traits lie on an underlying continuum and are not discrete valued. Second, the recent research relies on structural validity (e.g., evidence of convergent and discriminant validity) rather than historical path dependency (e.g., diagnoses that persist because they are familiar to clinicians who learned about them during their training).

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,[148] which are 10 personality disorders that are more chronic and, generally, less impairing of overall functioning. Research has documented that Big Five Neuroticism is a nonspecific 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). The direction of causality is difficult to ascertain in typical cross-sectional studies (Bagby et al., 1997; Cloninger, Svrakic, Bayon, and Przybeck, 1999; Gunderson, Triebwasser, Phillips, and Sullivan, 1999). Twin studies demonstrate that the shared variance in mental disorders and personality traits is predominantly genetic, that is, common genetic antecedents give rise to certain mental disorders and personality traits.

More research has examined relations between Axis II disorders and normal personality variation. Several authors have proposed a Big Four taxonomy (the Big Five minus Openness to Experience). Watson, Clark, and Chmielewski (2008) proposed that a fifth factor called Oddity is 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, 2002, and Widiger, Trull, Clarkin, Sanderson, and Costa, 2002). For instance, Widiger, Trull, Clarkin, Sanderson, and Costa (2002) suggest that all Axis II personality disorders can be "translated as maladaptively extreme variants of the 30 facets" of the NEO-PI-R. More recently, Samuel and Widiger (2008) completed a meta-analytic review of the relationships between facets of the NEO-PI-R and Axis II personality disorders, which we reproduce in Table 1.4. Notably, personality disorders relate to multiple facets spanning more than one Big Five factor.

5.6. 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 personality traits of the test taker, as well as their motivation to perform.[149] A smart child unable to sit still during an exam or uninterested in exerting much effort can produce spuriously low scores on an IQ test.

[148] Axis II also includes mental retardation.

[149] It is likely that performance on personality tests also depends 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.

Table 1.4 Mean Correlations of Psychopathological Measures with the Big Five Traits

	FFM Facet	Paranoid	Schizoid	Schizotypal	Antisocial	Borderline	Histrionic	Narcissistic	Avoidant	Dependent	Obsessive
N	Anxiousness	**0.27**	0.13	**0.27**	0.00	**0.38**	0.00	0.02	**0.41**	**0.39**	0.16
	Angry hostility	**0.41**	0.19	**0.29**	**0.27**	**0.48**	0.08	**0.23**	**0.29**	0.18	0.10
	Depressiveness	**0.35**	**0.28**	**0.39**	0.12	**0.50**	−0.06	0.03	**0.53**	**0.41**	0.09
	Self-consciousness	**0.29**	**0.23**	**0.32**	0.02	**0.35**	−0.11	−0.03	**0.56**	**0.42**	0.13
	Impulsiveness	0.15	0.00	0.17	**0.27**	**0.34**	0.17	0.14	0.14	0.17	−0.07
	Vulnerability	**0.22**	0.14	**0.25**	0.04	**0.39**	0.01	−0.01	**0.40**	**0.43**	0.03
E	Warmth	**−0.28**	**−0.42**	**−0.28**	−0.13	**−0.20**	**0.26**	−0.07	**−0.35**	−0.03	−0.07
	Gregariousness	**−0.20**	**−0.48**	**−0.25**	0.02	−0.12	**0.35**	0.04	**−0.42**	−0.03	−0.16
	Assertiveness	−0.08	**−0.22**	−0.13	0.06	−0.09	**0.27**	0.19	**−0.39**	**−0.21**	−0.01
	Activity	−0.08	**−0.25**	−0.13	0.02	−0.10	**0.25**	0.09	**−0.29**	−0.12	0.03
	Excitement seeking	−0.01	**−0.21**	−0.04	**0.25**	0.06	**0.27**	0.16	**−0.23**	−0.06	−0.12
	Positive emotions	**−0.27**	**−0.38**	**−0.26**	−0.09	**−0.26**	**0.23**	−0.02	**−0.39**	−0.15	−0.09
O	Fantasy	0.00	−0.05	0.14	0.10	0.13	0.16	0.11	0.00	0.05	−0.09
	Aesthetics	−0.05	−0.06	0.07	0.00	0.05	0.10	0.04	−0.03	0.01	0.01
	Feelings	−0.02	−0.17	0.03	−0.02	0.09	0.18	0.05	−0.04	0.05	0.01
	Actions	−0.10	−0.13	−0.06	0.10	−0.03	0.12	0.04	**−0.20**	−0.13	−0.12
	Ideas	−0.03	0.00	0.09	0.04	−0.01	0.04	0.07	−0.05	−0.12	0.03
	Values	−0.05	−0.05	0.01	0.08	0.05	0.04	−0.01	−0.05	−0.04	−0.09
A	Trust	**−0.45**	**−0.28**	**−0.31**	**−0.22**	**−0.29**	0.05	**−0.20**	**−0.29**	−0.07	−0.08
	Straightforwardness	**−0.24**	−0.09	−0.16	**−0.37**	**−0.21**	−0.10	**−0.31**	−0.06	0.00	0.04
	Altruism	**−0.21**	−0.19	−0.15	**−0.24**	−0.18	0.02	**−0.20**	−0.12	0.03	0.04
	Compliance	**−0.27**	−0.08	−0.13	**−0.32**	**−0.27**	−0.12	**−0.26**	−0.02	0.10	0.01
	Modesty	−0.06	0.08	0.05	−0.17	0.03	−0.16	**−0.37**	**0.20**	0.16	0.02
	Tender-mindedness	−0.18	−0.11	−0.05	−0.19	−0.09	0.02	−0.17	−0.02	0.09	0.00
C	Competence	−0.13	−0.13	−0.18	**−0.21**	**−0.29**	−0.01	0.01	**−0.23**	**−0.25**	0.19
	Order	0.00	−0.02	−0.06	−0.18	−0.10	−0.05	−0.03	−0.03	−0.06	**0.25**
	Dutifulness	−0.10	−0.08	−0.10	**−0.29**	**−0.22**	−0.08	−0.10	−0.09	−0.08	**0.25**
	Achievement striving	−0.07	−0.13	−0.13	−0.19	−0.19	0.04	0.02	−0.19	−0.16	**0.25**
	Self-discipline	−0.14	−0.12	−0.18	**−0.25**	**−0.29**	−0.04	−0.09	**−0.22**	**−0.23**	**0.21**
	Deliberation	−0.09	−0.02	−0.10	**−0.38**	**−0.27**	−0.16	−0.13	−0.01	−0.06	**0.24**

Notes: All values larger than $r = 0.04$ are significant at $p < 0.05$; correlations larger than 0.20 are marked in boldface type.
Source: Reproduced from Samuel and Widiger (2008).

It is sometimes claimed that IQ tests measure maximal performance, that is, that IQ scores reflect the application of the maximal capacity of the person to the test.[150] The analysis of Section 3 suggests that IQ scores should be standardized for effort. A series of studies conducted over the past 40 years support this concern.

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.[151] 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 in which motivation to perform well was maximal. Raver and Zigler (1997) present further evidence on this point. Table 1.5 summarizes evidence that extrinsic incentives can substantially improve performance on tests of cognitive ability, especially among low-IQ individuals.[152]

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. Men with lower levels of Conscientiousness 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. They already operate at a high level even without 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.[153] 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.

[150] A leading psychometrician, Carroll (1993), does not accept the notion that IQ captures maximal effort.
[151] The incentives for invoking effort vary across studies.
[152] The studies do not include direct measures of personality traits.
[153] An ERP is an electrophysiological response of characteristic form and timing to a particular category of stimuli.

Table 1.5 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 and 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 and 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 and 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 higher out of possible 51 points on a posttest given under standard conditions higher than at baseline; control group dropped 2.75 points. On a second posttest with incentives, expand control groups increased 7.17 and 6.25 points	

Continued

Table 1.5 Incentives and Performance on Intelligence Tests—continued

Study	Sample and Study Design	Experimental Group	Effect Size of Incentive (in Standard Deviations)	Summary
Clingman and Fowler (1976)	Within subjects study of 72 first and second graders assigned randomly to contingent reward, noncontingent reward, or no reward conditions	M&Ms given for right answers in contingent condition; M&Ms given regardless of correctness in noncontingent condition	Only among low-IQ (<100) subjects was there an effect of the incentive. Contingent reward group scored about 0.33 standard deviations higher on the Peabody Picture Vocabulary test than did no reward group	"… contingent candy increased the IQ scores of only the "low-IQ" children. This result suggests that the high and medium-IQ groups were already functioning at a higher motivational level than children in the low-IQ 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 to be consistent with previous sample descriptions.	Motivation was optimized without giving test-relevant information. Gentle encouragement, easier items after items were missed, and so on.	At baseline (in the fall), there was a full standard deviation difference (10.6 points and SD was about 9.5 in this sample) between scores of children in the optimized vs. standard conditions. The nursery group improved their scores, but only in the standard condition.	"… performance on an intelligence test is best conceptualized as reflecting three distinct factors: (a) formal cognitive processes; (b) informational achievements that reflect the content rather than the formal properties of cognition, and (c) motivational factors that 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)

Continued

Table 1.5 Incentives and Performance on Intelligence Tests—continued

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

Continued

Table 1.5 Incentives and Performance on Intelligence Tests—continued

Study	Sample and Study Design	Experimental Group	Effect Size of Incentive (in Standard Deviations)	Summary
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) $F(1,93) = 2.76$, $p < 0.05$	Two reasons why incentive did not produce dramatic increase: few or no unmotivated subjects among college volunteers; 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 two 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: an increase in *g* due to schooling at an intensive charter school; an increase in knowledge or crystallized intelligence; an increase in motivation due to the change in IQ test format; and/or an increase in motivation due to experience at high-performing school.

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 overreacts 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 investment theory (1971) anticipates recent findings that knowledge and specific complex skills depend not only on the 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 that affect the acquisition of knowledge. 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. Figures 1.5 and 1.6, 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 measures.[154] We adjust by Rotter and Rosenberg in Fig. 1.5 and by the Big Five in Fig. 1.6.[155] A substantial portion of the variances in both AFQT scores and

[154] AFQT and DAT scores are highly correlated ($r = 0.76$). See Borghans, Golsteyn, Heckman, and Humphries (2011); Kilburn, Hanser, and Klerman (1998); Sticht (1995); and Wang (1993).

[155] The Big Five are not available in the NLSY79 data that have AFQT scores.

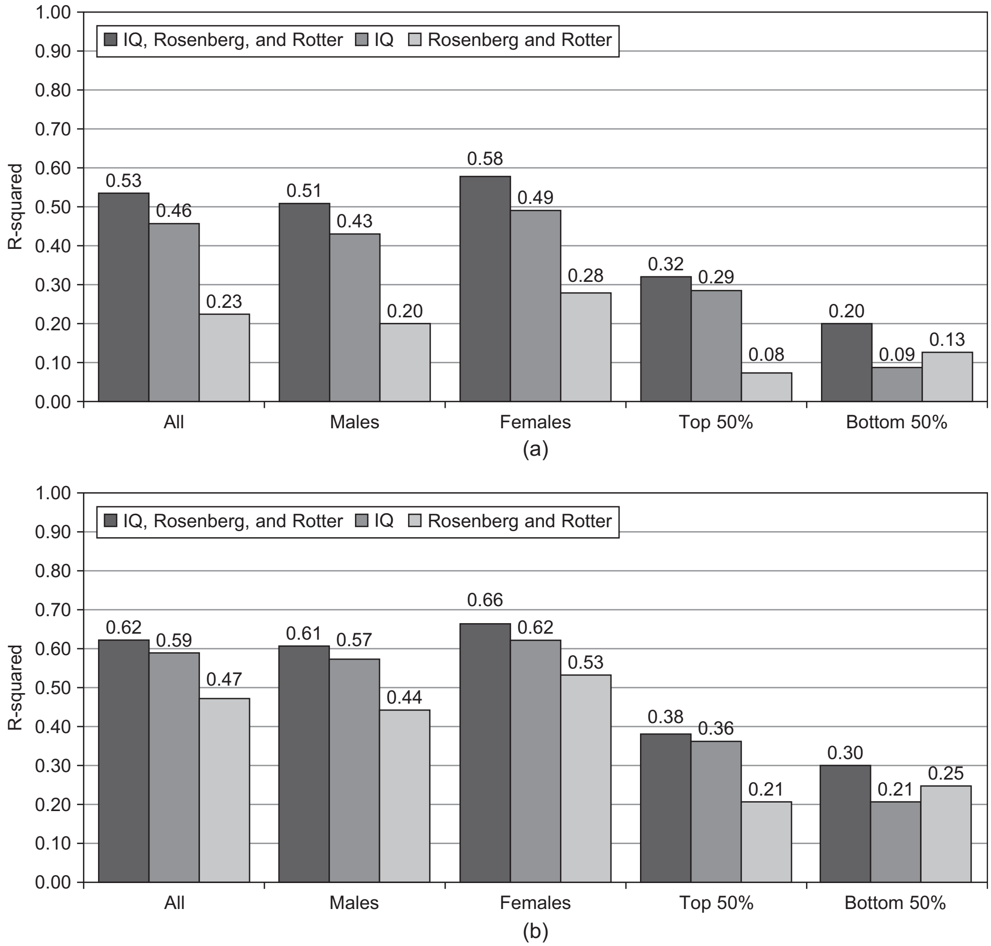

Figure 1.5 AFQT Decomposed by IQ, Rosenberg, and Rotter: (a) Not Controlling for Background Characteristics. (b) Controlling for Background Characteristics.

Notes: The data come from the NLSY79. Rotter was administered 1979. The ASVAB and Rosenberg were 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, Mathematical Knowledge, and Paragraph Comprehension ASVAB subtests. 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 ninth grade. Sample excludes the military oversample. Background variables include race and sex dummies, 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 cross-sectional sample of the NLSY79.
Source: Borghans, Golsteyn, Heckman, and Humphries (2011).

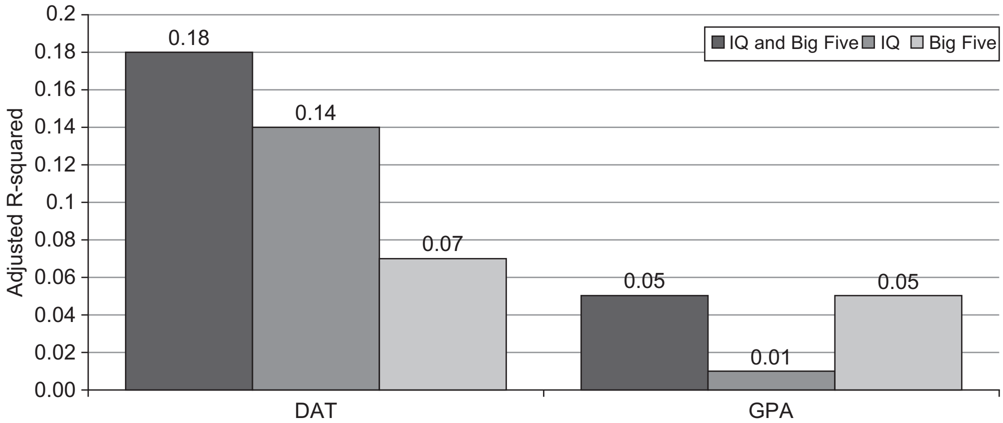

Figure 1.6 DAT Scores and GPA Decomposed by IQ and Personality.

Notes: Data is from Stella Maris, a high school in the Netherlands. Students were administered part of a Raven's IQ test and personality questions based on the Big Five. DAT and GPA are from high-school records. Source: Borghans, Golsteyn, Heckman, and Humphries (2011).

DAT scores are explained by personality factors.[156] 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 intelligence used by Herrnstein and Murray in *The Bell Curve* (1994) to predict a variety of social and economic outcomes is substantially affected by personality measures. We discuss evidence about personality and standardized achievement tests further in Section 7.

5.7. The Evidence on the Situational Specificity Hypothesis

Since the publication of Mischel's (1968) book, psychologists have addressed the situational specificity hypothesis, that is, the hypothesis that situations help explain variations across people in actions, effort, and behavior.[157] Boiled down to its essence, this hypothesis says little more than that situations affect actions and efforts in a nonlinear fashion, that is, that in Eqs (1.13)–(1.15), situational variables enter in a nonlinear fashion. This interaction effect gives rise to the Mischel and Shoda (1995) "if–then" relationship.

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

[157] For an early symposium in psychology on the person-situation debate, see Endler and Magnusson (1976).

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 behaviors, 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, and Angleitner (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 0.29 to 0.51.[158] 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.

A recent summary of the evidence on the person-situation debate is provided in a series of papers in the *Journal of Research in Personality* (January, 2009, Vol. 43) that offers a retrospective on the controversy. Virtually, all papers in that special issue acknowledge the existence of stable personality traits whose manifestations are affected by situations and incentives. The editors summarize the main message of the collected papers 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 are more than just error variance and that not all behavior is simply a function of the situation."*
>
> ***(Lucas and Donnellan, 2009, p. 147)***

[158] 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

Measures of personality predict a wide range of life outcomes that economists study. However, with our current knowledge, it is difficult to relate them to economic preference parameters except, of course, when the traits are the parameters. Since personality psychologists define traits as "relatively" stable, person-specific determinants of behavior, preferences are the natural counterpart of these traits in economics. Preferences are unaffected by changes in constraints. Although personality might relate to preferences, the exact link is unclear. Do preferences generate measured personality? Does personality generate preferences? Or, are both generated from other, deeper, motivation parameters that are as yet unknown? The model in Section 3 links preferences to measured personality. This section reviews the empirical evidence linking preferences and personality and discusses the conceptual differences between the two.

Overall, the links between measures of personality and preferences are largely unexplored. However, some evidence suggests that social preferences can be linked to the Big Five. The links between traditional preferences, such as risk aversion and time discounting, and personality, remain largely unknown. Personality measures might allow economists to broaden the dimensions of preferences and could potentially resolve some apparent inconsistencies in observed choices that arise from commonly used preference specifications in economics.

6.1. Evidence on Preference Parameters and Corresponding Personality Measures

The aspects of preferences that receive the most attention in the economics literature—time discounting, risk aversion, leisure preference, and social preferences[159]—appear to have analogues in the literature in psychology. Table 1.6 presents the definitions of commonly used preferences, tasks, and survey questions that have been used to measure them, and an overview of how they relate to measures of personality. The table includes measures, as well as latent factors (see Section 4).

Since the 1960s, psychologists have used experiments to elicit time preference and risk preference (see, e.g., Mischel et al., 2010, and Slovic, 1962). A recent example is the Balloon Analogue Risk Task (BART) (Lejuez et al., 2002), 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. In addition to the experimental measures, it is tempting to try to map preferences to more vaguely defined traits, but the precise mapping has not yet been made. Still, some speculation is useful. Time preference

[159] For a definition of these concepts and a discussion of measurement of preferences, see Table 1.6 and Sections A6.1 and A6.2 in the appendix.

Table 1.6 Measures of Standard Preference Parameters and Analogous Measures in the Psychology Literature

Preference	**Survey Questions and Experiments Used to Elicit Preference**	**Overview of Relationship to Personality Measures**
Time preference: Preference over consumption in different time periods	*Delay discounting*: A participant is given a series of choices for whether he would prefer to receive smaller payments sooner versus larger payments later. The amounts and times vary across choices. The choices can be over hypothetical payoffs or real-stakes payoffs (see, e.g., Dohmen et al., 2011). *Marshmallow task*: A participant (usually a child) is given a marshmallow. The experimenter leaves the room and tells the participant that he will receive a second marshmallow if he resists consuming the marshmallow until the experimenter returns. The length of time that the participant waits is a measure of short-term discounting (see, e.g., Mischel et al., 2010). *Example survey question*: "How patient are you on a scale from 1 to 10?" (see GSOEP, 2008).	*Conceptual relationships*: Conscientiousness, Self-Control, Affective Mindfulness, Consideration of Future Consequences, Elaboration of Consequences, Time Preference. *Empirical relationships*: Conscientiousness, Self-Control, Affective Mindfulness, Elaboration of Consequences, Consideration of Future Consequences (Daly, Delaney, and Harmon, 2009). Extraversion, Time Preference (Dohmen, Falk, Huffman, and Sunde, 2010). Agreeableness, Inhibitive Side of Conscientiousness (Anderson, Burks, DeYoung, and Rustichini, 2011).

Continued

Table 1.6 Measures of Standard Preference Parameters and Analogous Measures in the Psychology Literature—continued

Preference	Survey Questions and Experiments Used to Elicit Preference	Overview of Relationship to Personality Measures
Risk aversion: Preference over different states of the world	*Lottery choice task*: A participant is given a series of choices between a safe amount of money and a lottery. The lottery remains the same across choices, whereas the safe amount varies. The lowest safe amount for which the participant prefers the lottery is a measure of risk aversion. The choices can be over hypothetical payoffs or real-stakes payoffs (see, e.g., Dohmen et al., 2011). *Devil's Task (Slovic's risk task)*: A participant sequentially chooses between ten "switches" or urns associated with hidden payoffs. The participant is told that nine of the switches are associated with a reward and one of them results in a loss of all previous winnings. Once a participant chooses a switch, he cannot flip the same switch again. The participant can elect to stop picking switches at any time. The number of switches chosen is a measure of risk aversion (see, e.g., Slovic, 1966). *Balloon Analogue Risk Task (BART)*: The participant is given a computerized task in which he is presented with a series of "balloons" that can be inflated by "pumping" the balloon. The participant receives potential earnings each time he pumps a balloon. At any point, the participant can stop pumping, realize the potential earnings, and move to the next balloon. After a threshold number of pumps each balloon "explodes," and the participant receives nothing. The threshold varies across balloons, and participants are not told the distribution of thresholds (see, e.g., Lejuez, Aklin, Zvolensky, and Pedulla, 2003). *Example survey question*: "How willing are you to take risks, in general?" (see, e.g., Dohmen et al., 2011).	*Conceptual relationships*: Impulsive Sensation Seeking, Balloon Analogue Risk Task. *Empirical relationships*: Sensation Seeking (Zuckerman, 1994; Eckel and Grossman, 2002). Openness (Dohmen, Falk, Huffman, and Sunde, 2010). Neuroticism, Ambition, Agreeableness (Borghans, Golsteyn, Heckman, and Meijers, 2009). Balloon Analogue Risk Task (Lejuez, Aklin, Zvolensky, and Pedulla, 2003). Neuroticism, Inhibitive Side of Conscientiousness (Anderson, Burks, DeYoung, and Rustichini, 2011).

Continued

Table 1.6 Measures of Standard Preference Parameters and Analogous Measures in the Psychology Literature—continued

Preference	Survey Questions and Experiments Used to Elicit Preference	Overview of Relationship to Personality Measures
Leisure: Preference over consumption and leisure	*Payments for working*: The participant is given a choice to work at different wages. Their reservation wage is their preference for leisure. The choices can be over hypothetical payoffs or real-stakes payoffs (see, e.g., Borghans, Meijers, and ter Weel, 2008).	*Conceptual relationships*: Achievement Striving, Endurance, Industriousness. *Empirical relationships*: Inconsistent with psychological measures of leisure preferences (Borghans, Meijers, and ter Weel, 2008).
Altruism: Unconditional kindness *Inequity aversion:* Value of equality in payoffs	*Dictator game*: A "proposer" has the option to transfer part of an endowment to a "responder." The responder passively receives any transfer. The transfer is used as a measure of pure altruism (see, e.g., Fehr and Schmidt, 2006).	*Conceptual relationships*: Warmth, Gregariousness, Tender-Mindedness, Hostility (opposite). *Empirical relationships*: Neuroticism, Agreeableness (Ashton, Paunonen, Helmes, and Jackson, 1998; Osiński, 2009; Bekkers, 2006; Ben-Ner and Kramer, 2011).
Trust: Willingness to make oneself vulnerable to opportunistic individuals	*Trust game*: An "investor" receives an endowment and can decide to transfer some of it to a "trustee." The amount transferred increases in value. The trustee can then decide to transfer some back to the investor but has no monetary incentive to do so. The amount the investor transfers to the trustee is used a measure of trust (see, e.g., Fehr and Schmidt, 2006). *Example survey question*: "In general, one can trust people" (see, e.g., Dohmen, Falk, Huffman, and Sunde, 2008).	*Conceptual relationships*: Trust. *Empirical relationships*: Neuroticism, Agreeableness, Openness, Conscientiousness (Dohmen, Falk, Huffman, and Sunde, 2008).

Continued

Table 1.6 Measures of Standard Preference Parameters and Analogous Measures in the Psychology Literature—continued

Preference	Survey Questions and Experiments Used to Elicit Preference	Overview of Relationship to Personality Measures
Reciprocity: The way in which one person responds to another's actions. *Positive reciprocity*: Tendency to reward kind actions. *Negative reciprocity*: Tendency to punish others for unkind actions.	*Ultimatum game*: A "proposer" offers part of an endowment to a "responder." The responder can choose to accept the offer in which case both players keep the payoffs, or the responder can choose to reject the offer in which case the players receive nothing. The responder's choice is a measure of reciprocity (Fehr and Schmidt, 2006). *Trust game*: See above description. The trustee's action is used as a measure of reciprocity. *Gift exchange game*: An "employer" proposes a wage and an amount of desired effort to a potential "worker." The worker can either reject the proposal so that no one receives anything or can accept the proposal and choose any amount of effort. The employer receives a payment proportional to the worker's effort net of the wage. The workers action is used as a measure of reciprocity (see, e.g., Fehr and Schmidt, 2006). *Example survey question (positive reciprocity)*: "If someone does me a favor, I am prepared to return it." (see, e.g., Dohmen, Falk, Huffman, and Sunde, 2008). *Example survey question (negative reciprocity)*: "If I suffer a serious wrong, I will take revenge as soon as possible, no matter the cost." (see, e.g., Dohmen, Falk, Huffman, and Sunde, 2008).	*Conceptual relationships*: Warmth, Gregariousness, Hostility (opposite). *Empirical relationships*: Neuroticism, Agreeableness, Conscientiousness (Dohmen, Falk, Huffman, and Sunde, 2008).

likely relates to Conscientiousness, Self-control, and Consideration of Future Consequences. Risk Aversion is likely related to Openness to Experience and impulsive sensation seeking, a trait proposed by Zuckerman, Kolin, Price, and Zoob (1964), 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."[160]

Preferences for leisure may be related to several personality measures. The Big Five includes an 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." Industriousness has been proposed as one of six facets of Conscientiousness (Roberts, Chernyshenko, Stark, and Goldberg, 2005) and is plausibly related to the preference for leisure.

Social preferences also have conceptual 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.

Despite this intuitive mapping of preferences to traits, the empirical evidence supporting such mappings is weak. The few studies investigating empirical links typically report only simple regressions or correlations without discussing any underlying model. Some use survey and self-report measures similar to those used by psychologists rather than elicited preferences. The last column of Table 1.6 gives an overview of papers investigating these links.

The evidence relating personality to time preferences is mixed. Using data from an experiment involving college students, Daly, Delaney, and Harmon (2009) find that 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. Dohmen, Falk, Huffman, and Sunde (2010) measure time preferences experimentally. Although time preference is related to cognition, Openness to Experience is the only Big Five trait that explains any variation in time preference. Figure 1.7 reports correlations between experimental measures of time preference, Big Five factors, and measures of cognition. In that figure, only cognitive measures are correlated with time preference.[161]

Dohmen, Falk, Huffman, and Sunde (2010) find that Openness to Experience and Agreeableness are related to risk aversion. Figure 1.7 reports correlations between risk aversion, the Big Five, and measures of cognition for a sample of Germans. Of the

[160] See Zuckerman (1994).

[161] Figure A2 in Section A6 of the Web Appendix displays correlations among the survey measures in the GSOEP.

Time preference								
–	Risk tolerance							
0.01	0.14***	**O**						
0.07	–0.07	0.11***	**C**					
0.07	0.08	0.42***	0.11***	**E**				
0.08*	–0.09**	0.08***	0.29***	0.05	**A**			
0.06	–0.03	0.19***	0.06**	–0.01	–0.01	**N**		
–0.12***	0.18***	0.16***	–0.15***	0.15***	–0.14***	–0.06**	Symbol test	
–0.10**	0.19***	0.15***	–0.13***	0.12***	–0.18***	–0.08**	0.40***	Word test

Figure 1.7 Pairwise Correlations between Time Preference (Impatience), Risk Tolerance, Personality, and Cognitive Ability for Males and Females from GSOEP.

*Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; ***statistically significant at the 1% level.* ***O****, Openness to Experience;* ***C****, Conscientiousness;* ***E****, Extraversion;* ***A****, Agreeableness;* ***N****, Neuroticism. The value in each box is the pairwise correlation. Darker-shaded boxes have lower p-values. The measures of the Big Five are based on three 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 and risk tolerance were elicited through a real-stakes experiment. Source: The data come from Dohmen, Falk, Huffman, and Sunde (2010), available online. The calculations were conducted by the authors of this Handbook chapter.*

Big Five, Openness to Experience and Agreeableness are correlated with risk aversion. There is little evidence connecting risk aversion and sensation seeking, but Eckel and Grossman (2002) include it as a control in a study of risk aversion and find no statistically significant effect. However, Bibby and Ferguson (2011) find that sensation seeking is associated with a lottery measure of risk tolerance.[162] They also find that people who

[162] Bibby and Ferguson report this as a measure of loss aversion, but it is more akin to a measure of risk tolerance.

are better at processing emotional information and who are less extraverted are more susceptible to framing effects when making risky decisions. Borghans, Golsteyn, Heckman, and Meijers (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 ambition, a trait that may involve investment in uncertain opportunities. Further, Agreeableness is positively associated with risk aversion.

Figure 1.8 displays a related analysis by Anderson, Burks, DeYoung, and Rustichini (2011), who find that both cognitive ability and Agreeableness are positively associated with delay acceptance elicited from a real-stakes experiment in a sample of truck driver trainees.[163]

When they separately regress delay acceptance on Neuroticism, Agreeableness, Extraversion, Conscientiousness, cognitive skill, race, marital status, age, and education, none of the personality traits are statistically significant at the 10% level. However, when they split Conscientiousness into an inhibitive side (moral scrupulousness and cautiousness) and a proactive side (the need for achievement), they find that only the inhibitive side is positively associated with delay acceptance ($\beta = 0.13$, $p < 0.10$). This result highlights the importance of examining facets of the Big Five when considering the relationship between preferences and personality.

As shown in Fig. 1.8, Anderson, Burks, DeYoung, and Rustichini (2011) find that of the Big Five, only Neuroticism is positively associated with risk aversion but only for lotteries over gains, not losses. In a separate regression controlling for Neuroticism, Agreeableness, Extraversion, Conscientiousness, cognitive skill, race, marital status, age, and education, risk aversion is positively associated with both Neuroticism ($\beta = 0.15$, $p < 0.01$) and the inhibitive side of Conscientiousness ($\beta = 0.10$, $p < 0.10$).

The links between social preferences and the Big Five traits are better established. Ben-Ner and Kramer (2011) find that Extraversion is associated with higher giving in a dictator game. Dohmen, Falk, Huffman, and Sunde (2008) use an experimentally validated survey measure of trust and find that Conscientiousness and Neuroticism are negatively associated with trust, whereas Agreeableness and Openness to Experience are positively associated with trust. Agreeableness and Conscientiousness are associated with more positive reciprocity and less negative reciprocity, whereas Neuroticism is associated with more negative reciprocity.

In sum, although many measures of personality and preferences seem conceptually related, the empirical associations are not uniform across studies, and often the measures of preference are uncorrelated with intuitively similar personality traits. Nevertheless, in several studies, Neuroticism is associated with risk aversion, and facets of Conscientiousness are associated with delay acceptance. Some evidence suggests that considering facets of the Big Five

[163] They do not use a measure of Openness to Experience to separate out its influence from that of cognitive ability.

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

Figure 1.8 Pairwise Correlations between Risk Acceptance, Delay Acceptance, Cognitive Ability, and Personality in a Sample of US Truckers.

*Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; ***statistically significant at the 1% level. **C**, Conscientiousness; **E**, Extraversion; **A**, Agreeableness; **N**, Neuroticism. The value in each box is the pairwise correlation. Darker-shaded boxes have lower p values. Delay Acceptance and Risk Acceptance for Gains and Losses come from real-stakes experiments. Cognitive Skill is the first factor from a Raven's Progressive Matrix test, a numeracy test, and the Hit 15. The facets of the Big Five were constructed from the Multidimensional Personality Questionnaire. The sample consists of 1,065 trainee truck drivers in the United States.*
Source: Adapted from Anderson, Burks, DeYoung, and Rustichini (2011).

might help establish a mapping between personality and preferences. However, the empirical links between preference parameters and personality traits depend on the data used.

6.2. Mapping Preferences into Personality

Despite some plausible, empirical and conceptual links between preferences and traits, a precise mapping between the measures is not yet available. In Section 3, we argued that measured personality is generated by underlying preference parameters and constraints.

However, the preferences measured by economists are often chosen to ensure identification on particular types of data on choices and may not be sufficiently rich. Further, studies documenting relationships between preferences and traits typically only study correlations without being motivated by an underlying model. Hence, causal claims are, at this stage, largely premature. There are two main reasons for the disconnection between measures of personality and measures of preferences.

First, economists typically study marginal rates of substitution, measured over relevant ranges via observed choices. Personality psychologists typically do not study these trade-offs and often do not study choice behavior. Most approaches to measuring preferences in economics, whether observational or experimental, use some variation of revealed preference given observed choices. In contrast, psychologists typically use surveys to elicit preferences, information, or "typical" actions. Some questions elicit how respondents would feel about a given outcome, without presenting an alternative outcome. Although such questions may elicit some (unspecified) feature of preferences, it is not clear what is being measured. The difference in approach makes it intrinsically difficult to compare economic and psychological measures.

Second, traditional preference parameters may not span the entire space of human decisions measured by psychologists. Time, risk, social, and leisure preferences do not capture the only trade-offs in life. Although 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 typically make strong simplifying assumptions to make their models tractable and to secure identification. The estimated parameters are used to build models, evaluate policy, and create counterfactuals. The most widely used specifications of trade-offs are through parameterizations assuming separability and assume that marginal rates of substitution are summarized by one or two parameters. 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 policy analyses.

6.3. Do Measured Preference Parameters Predict Real-World Behavior?

One test of the stability of measured preferences is whether they predict behavior in other contexts. Several recent studies have investigated whether risk preferences predict behavior. For example, Dohmen et al. (2011) use an experimentally validated measure of risk preference in 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 a survey question about "willingness to take risks in general." However, the observed relationship might arise because both the self-reported behaviors and questions about willingness to take risk are noisy contemporaneous survey measures. Barsky, Juster, Kimball, and Shapiro (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, measures of risk tolerance only explain a small fraction of the variation in risky behaviors.

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 laboratory and the field and argue that this does not arise because people behave inconsistently, but because experimenters are not controlling for relevant aspects of the choice situation. Their work is a rehash of the old person-situation debate. Falk and Heckman (2009) present a different interpretation of the value of experiments. We discuss the evidence on this question below.

6.4. Integrating Traits into Economic Models

Behavioral economics has incorporated some aspects of personality psychology to investigate how standard models of preferences can be improved to better reflect reality. Behavioral economics has highlighted many so-called anomalies, ways in which standard preferences do not accurately describe 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. Below we discuss some of the now-standard examples, such as loss aversion, hyperbolic discounting, and reciprocity. These are not anomalies with respect to rationality but are examples that challenge 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 nonstandard preferences.

Second, behavioral economists have introduced the concept of bounded rationality. They discuss behaviors for which there is no reasonable preference specification that can rationalize a behavior. 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 think of these as mental constraints, or traits, along the lines of the models discussed in Section 3. However, these examples 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 high level of generality used in many quarters of behavioral economics tends to make it difficult to identify the parameters in the data commonly used by economists (see the discussion in Hansen, 2005).

6.4.1 Traits as Constraints

Preference measurements that do not account for all of the constraints that agents face might be biased. In the model of personality in Section 3, we describe how agents act based on both preference parameters and productive traits that embody constraints. The marginal rate of substitution is typically identified through price variation. However, the true price ratio might also depend partly on the unobserved traits of the individual. Failure to account for the traits that arise from constraints could lead 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 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 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 (e.g., de Wit, Flory, Acheson, McCloskey, and Manuck, 2007; Dohmen, Falk, Huffman, and Sunde, 2010), but it is less helpful in explaining why smarter individuals also have lower discount rates when choosing between relatively simple cash sums (Funder and Block, 1989) and between noncash rewards (such as smaller versus larger candy bars, as in Mischel and Metzner, 1962).[164] A meta-analysis by Shamosh and Gray (2007) of 24 studies in which both IQ and discount rates were measured shows that 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 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.

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

Measures of cognitive constraints also relate to measured risk preference. There is an inverse relationship between cognitive ability and risk aversion, where higher IQ people have higher risk tolerance (Dohmen, Falk, Huffman, and Sunde, 2010).[165] Reference dependence can lead subjects to be susceptible to framing because they will perceive two identical lotteries differentially when one is framed as a loss and the other is framed as a gain. Some evidence suggests that individuals with higher cognitive ability and education are less risk averse. Burks, Carpenter, Goette, and Rustichini (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 can better translate their preferences into choices between lotteries.

Borghans, Golsteyn, Heckman, and Humphries (2011) 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.

6.4.2 Traits as Preferences

Some aspects of traits may be more naturally believed as aspects of preferences than as constraints. For example, Openness to Experience might relate to a preference for learning, and Extraversion might reflect a preference for social interactions. The distinction between preferences and constraints often seems tautological. One way of incorporating personality into preferences is by modifying functional forms, which fall into two broad and sometimes overlapping categories. First, some of the domains that are traditionally treated as fundamentally different, such as risk and time preference, social and risk preference, and leisure and time preference, may be closely related and generated from a common set of psychological traits. Second, nonseparabilities could confound measures of trade-offs. The literature on addiction presents an interesting class of nonseparable models,[166] as does the literature on exotic preferences in economics.[167]

6.4.2.1 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 facilitates identification given sparse data, and if it is a sensible specification of preferences, it gives a convenient description of behavior. However, one or two parameters may not describe behavior well. Conversely, some of the concepts analyzed separately in the literature may be governed by the same parameters.

[165] The two cognitive ability tests used by Dohmen, Falk, Huffman, and Sunde (2010) were coding speed and vocabulary tests.

[166] See Becker and Murphy (1988).

[167] See Epstein and Zin (1989), Hansen (2005), and Hansen and Sargent (2008).

In discussing the concept of time discounting, Frederick, Loewenstein, and O'Donoghue (2002) argue that time preference has three dimensions: *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. There are multiple interpretations of this assertion.

First, the trade-off between different time periods might be described by several parameters. Second, impulsivity, inhibition, and compulsivity might reflect constraints, that is, something that affects shadow prices of consumption in different time periods. Third, the relevant trade-off might not be between different time periods but, for example, in the case of impulsivity might be between various levels of sensation seeking, a behavior which is also related to risk seeking.

Like time preference, risk preference may depend on multiple parameters. 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 function. See Starmer (2000) for a review of the literature on departures from expected utility. When psychologists started measuring risk-taking behavior, they were puzzled by the large variance across domains (see the discussion of situational specificity in Section 2). More recently, 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, and 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 et al. (2011) find that a survey question on willingness to take risks within a domain predicts self-reported behaviors within each domain. Einav, Finkelstein, Pascu, and Cullen (2010) also find that there are domain-specific components of risk-taking behavior. Domain specificity might arise because sensation seeking, enjoyment of risk per se, is an important aspect of risk preferences.[168]

Ambiguity aversion, the disutility from model uncertainty, might help explain some apparent inconsistencies. Ambiguity aversion is measured as the trade-off between lower expected return and higher model uncertainty. Ambiguity aversion explains Ellsberg's paradox: people tend to prefer an urn with a 0.5 probability of winning to an urn with an unknown probability in which they are allowed to choose which side to bet on. One version of preferences over ambiguity is due to Gilboa and Schmeidler (1989).

[168] 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.

They specify max-min preferences, where the agent maximizes an expected utility function that has been minimized with respect to the prior probabilities, that is,

$$U(X_1, X_2, \ldots, X_K) = \min_{(\pi_1, \pi_2, \ldots, \pi_K)} \pi_1 u(X_1) + \pi_2 u(X_2) + \cdots + \pi_K u(X_K)$$

$$\text{subject to} \qquad \sum_{j=1}^{K} \pi_j = 1.$$

Borghans, Golsteyn, Heckman, and Meijers (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.

There is no consensus on how social preferences govern choices. Social preferences refer to any explanation for nonselfish behavior, usually as measured in a dictator game in which people have to divide a sum between themselves and another person. Typically, more than 60% of proposers give positive amounts, averaging 20% of the sum. A variation of this game is the classic ultimatum game in which a giver divides a sum between him or herself and another subject, the receiver, and 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%. 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. Many studies seek to find deeper traits that govern these behaviors, such as preferences over the utility of one's own consumption compared to 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, that captures the preference that people often care about other people's utility or consumption. See Meier (2007) for a review. Fehr and Schmidt (1999) analyze inequality aversion. People dislike inequality rather than valuing the consumption or utility of agents per se.

Caring and altruism have 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 prosocial or antisocial manner depending on the behavior of others with whom they interact. People exhibit 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 in order to reward or punish others. Falk and Fischbacher (2006) develop a theory of reciprocity in which utility depends on the kindness of others, which is a function not only of the outcome from another person's action but also of the perceived intentions. Reciprocity then reflects how much value a person puts on rewarding kindness. Economists could model 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 in which the proposer offers a wage to a responder, who then subsequently chooses a level of effort. However, List (2009) argues that the importance of fairness preferences may have been overstated in the literature and that many of the observed results are due to concerns over either reputation or scrutiny by experimenters. Several studies have shown that observed reciprocity fades over a longer time frame than the short duration of lab experiments (Gneezy and List, 2006; Hennig-Schmidt, Rockenbach, and Sadrieh, 2010; Kube, Maréchal, and Puppe, 2006). Andreoni's (1995) warm glow model of altruism suggests that people do not care about others, but value 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 n:

$$\begin{aligned} &U_n(X_1, \dots, X_j, \dots, X_n, \dots, X_K) = \\ &\quad X_n - \alpha_n \frac{1}{K-1} \sum_{j \neq n} \max\{\,|X_j - X_n|, 0\} - \beta_n \frac{1}{K-1} \sum_{j \neq n} \max\{\,|X_n - X_j|, 0\}, \end{aligned}$$

where the weights satisfy $\beta_n \leq \alpha_n$ and $0 \leq \beta_n \leq 1$. People place a higher weight on own consumption compared to others', but they place value on inequality in a situation.

People appear to be more tolerant of inequality if they believe that it represents a difference earned through effort rather than from differences in exogenously given talent (see Tausch, Potters, and Riedl, 2010, for a review). This finding 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.

Some aspects of preferences appear to be multidimensional. However, many preference parameters are correlated. For example, the social preference of "trust" relates to risk aversion and reciprocity. Altmann, Dohmen, and Wibral (2008) measure trust as the willingness to give money to an investor in a trust game in which he will only be repaid if the investor decides to return the favor. In this game, one can think of trust as the belief about how own actions affect those of others. They find that trust and positive reciprocity are positively related. Using the German Socio-Economic Panel (GSOEP), Dohmen, Falk, Huffman, and Sunde (2008) 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. In situations involving trust, it seems natural that trust is closely related to risk and ambiguity aversion and that a person who is more prone to accept uncertainty is also more likely to trust others. Altmann, Dohmen, and Wibral (2008) also find that people who are less risk averse are also more willing to trust. However, they do not measure which beliefs the agents hold.

Care has to be taken in distinguishing trust from risk aversion. Kosfeld, Heinrichs, Zak, and Fehr (2005) find that people who receive oxytocin exhibit more trusting behavior in a real-stakes trust game. However, oxytocin does not make subjects more generous, suggesting that trust is not simply altruism. In addition, oxytocin does not affect people's decision over risky outcomes when playing against a computer rather than a human. Together, 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.

6.4.2.2 Preference Specifications and Their Consequences

The most restrictive version of the additively separable model suggests that the marginal rate of substitution between any two goods does not depend on the consumption of other goods outside of the pair. Browning, Hansen, and Heckman (1999) present ample evidence against this assumption. Apparent inconsistencies can arise if nonseparability is ignored. Further, estimates will suffer from omitted variable bias.[169]

The additively separable intertemporal model imposes the requirement that the intertemporal elasticity of substitution is the same as the relative risk aversion parameter. However, Barsky, Juster, Kimball, and Shapiro (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 the same stimulus: for example, an increase in the size of a reward generally decreases the time discount factor but increases the discount rate when rewards are probabilistic.[170] This is evidence against the standard intertemporally separable model of risk aversion.

[169] See Section A6.4 in the Web Appendix for a discussion of additive separability and its implications.

[170] 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, and Shapiro (1997) report that their estimates of time preference and risk tolerance are independent.

One type of nonseparability is between goods and the state or time period. The additively separable model allows for this type of dependence, represented by the subscript v on the utility function. Although 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 than would be predicted by exponential discounting. This is the motive 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,

$$U_v(X_{v,} X_{v+1}, \dots) = u(X_v) + \delta\beta u(X_{v+1}) + \delta\beta^2 u(X_{v+2}) + \cdots.$$

The consequence of these preferences is that the trade-off between period v and period $v+1$ is not evaluated the same way from the perspective of period $v-1$ and period v, leading to time inconsistency.[171] Other possibilities are that the discount rates change with age. Hyperbolic and age-dependent discounting make use of the subscript v 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 might account for what future agents might do. Further, discount rates appear to vary inversely with the size of reward and vary with the type of reward offered.[172]

As previously noted, the expected utility form for risk preferences does not explain risk preferences over small stakes (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) suggest that people are loss averse, that is, that losses weigh higher than gains in the utility function. This would imply that people have state-dependent preferences, which can be expressed as

$$U_n(X_{1,} X_2, \dots X_n, \dots X_K) = \pi_1 u(X_1 - X_n) + \pi_2 u(X_2 - X_n) + \cdots + \pi_K u(X_K - X_n),$$

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

The concepts of loss aversion, reference-dependence, and endowment effects (Thaler, 1980, and 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

171 This specification originates in the work of Phelps and Pollak (1968).

172 See Green, Fry, and Myerson (1994); Chapman, Nelson, and Hier (1999); Kirby (1997); Chapman and Coups (1999); Estle, Green, Myerson, and Holt (2007); Bickel, Odum, and Madden (1999); and Bonato and Boland (1983).

how he trades it off against other goods. In certain experiments, List (2003) has shown that this effect disappears when agents have market experience.

Reference-dependence has also been demonstrated in dictator games. In the standard dictator game, the first player, the "dictator," is given a positive endowment while the second player receives nothing. The game has a single move. The first player can choose to transfer some or none of his endowment to a second player. At this point the game ends. The second player can take no action. Numerous studies have shown that most first players transfer a positive amount, even though they have no monetary incentive to do so. List (2007) and Bardsley (2008) modified the standard dictator game by giving the second player an endowment and allowing the first player to transfer a positive amount to the second player or to take part of the second player's endowment. With this modification, most first players did not transfer positive amounts to the second player.

Experimental measures of social preferences vary greatly across studies. Levitt and List (2007) and List (2009) argue that the degree of scrutiny in the laboratory as opposed to in the real world may make subjects behave more prosocially (Bandiera, Barankay, and Rasul, 2005).[173] 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 between and within socioeconomic groups. Marginal rates of substitution depend on other factors such as education, age, cultural values, etc.[174] 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. Although most studies view life outcomes as the result of choices governed by exogenous preferences, and hence infer preferences from outcomes, initial conditions might 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 on predictive power. When measuring preferences, functional forms are chosen in an attempt to minimize approximation error subject to identification. However, economists typically consider preferences over a limited range of fundamental attributes. Time, risk, and social preferences may not be the right dimensions of choice over which parameters are stable. Each of these domains seems to be guided by multiple parameters, and some of these parameters seem to matter for each of the domains. Personality psychology may help in guiding economists as where to look for more fundamental parameters. However, the potential is largely unexploited.

[173] See, however, Falk and Heckman (2009).

[174] See the evidence in Browning, Hansen, and Heckman (1999).

6.5. Summary of Section 6 and Some Concluding Thoughts

Table 1.7 summarizes the main papers relating economic preference parameters to psychological measurements. The lack of any close correspondence between the traits of personality psychology and the parameters of economics suggests great opportunities for both fields. Each can learn from the other.

However, the different emphases of interest in the two fields may present a challenge to integration. The greater emphasis on prediction, intervention, and causality in economics, compared to the greater emphasis on description in personality psychology, may lead to a new choice-motivated set of psychological traits that supplant traditional Big Five measures. Developments in this direction are discussed in Ferguson, Heckman, and Corr (2011). Personality psychologists may well adopt a more choice-oriented set of parameters to supplant the description-oriented Big Five parameters. Moreover, both the conventional psychological traits and economic preference parameters may be stable manifestations of deeper parameters, connected to human motivation and goal-seeking that remain to be discovered.

7. THE PREDICTIVE POWER OF PERSONALITY TRAITS

This section discusses the empirical evidence on the power of personality in predicting life outcomes. A growing body of evidence suggests that personality measures—especially those related to Conscientiousness and, to a lesser extent, Neuroticism—predict a wide range of outcomes. The predictive power of any particular personality measure tends to be less than the predictive power of IQ, but in some cases rivals it.

For three reasons, summarizing the large literature on the predictive power of personality on outcomes is a daunting task. First, the measures 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 Big Five categories. When this is not possible, we discuss the measures used and how they relate to the Big Five measures.

Second, different studies use different measures of predictive power. Many studies report only simple correlations or simple standardized regression coefficients.[175] Such estimated relationships do not control for other factors that may influence outcomes. This is particularly problematic for estimated relationships between personality measures and other outcomes that do not control for cognition, situation, or the effect of other personality measures. Where possible, we report both simple and partial correlations.

[175] Standardized regressions produce regression coefficients of outcomes divided by their standard deviations regressed on explanatory variables divided by their standard deviations. This produces correlation coefficients in bivariate regressions and partial correlation coefficients in multivariate regressions. See, e.g., Goldberger (1968).

Table 1.7 Links among 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, and 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 ($p < 0.05$).

Continued

Table 1.7 Links among Personality Traits and Preferences—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
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 ($p < 0.05$).

Continued

Table 1.7 Links among Personality Traits and Preferences—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Burks, Carpenter, Goette, and Rustichini (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 US 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 that 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 ($p < 0.10$).

Continued

Table 1.7 Links among Personality Traits and Preferences—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Dohmen et al. (2011)	Outcome(s): *experimental risk measure*—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*—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% 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.

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 that are used in psychology. This method 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.[176]

For example, Heckman, Stixrud, and Urzua (2006) report the effects of percentile changes in cognitive and personality measures 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 approach to measuring predictive power is increasingly being applied by economists.[177]

Third, many studies do not address the question of causality, that is, 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 working on the relationship between personality and outcomes 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. As discussed in Section 4, using early measures of personality traits to predict later outcomes raises problems of its own. We delineate how each study surveyed addresses causality.

7.1. An Overview of the Main Findings

Before presenting a detailed survey of the effects of personality and cognition on a variety of outcomes, it is useful to have an overview of the main findings. One principal finding of our survey, consistent with the claims of the early psychologists cited in Section 2, is that Conscientiousness is the most predictive Big Five trait across many outcomes. However, other personality measures predict some outcomes.

[176] The slope versus variance explained distinction is an old one. However, the use of slopes as measures of "importance" is problematic in general because of the arbitrariness in the scales of the dependent and independent variables (see Goldberger, 1968). This arbitrariness is resolved in the measure used in the recent literature by mapping quantiles into quantiles. This literature is nonparametric. The measure is clear in its choice of units but the economic significance is still questionable. A better measure would relate costs of a change in the independent variable to the benefits.

[177] See, e.g., Piatek and Pinger (2010).

Measures of personality predict a range of educational outcomes. Of the Big Five, Conscientiousness best predicts overall attainment and achievement. Other traits, such as Openness to Experience, predict finer measures of educational attainment, such as attendance and course difficulty selected. Traits related to Neuroticism also affect educational attainment, but the relationship is not always monotonic. Conscientiousness predicts college grades to the same degree as SAT scores. Personality measures predict performance on achievement tests and, to a lesser degree, performance on intelligence tests.

Personality measures also predict a variety of labor market outcomes. Of the Big Five traits, Conscientiousness best predicts overall job performance but is less predictive than measures of intelligence. However, Conscientiousness predicts performance and wages across a broad range of occupational categories, whereas the predictive power of measures of intelligence decreases with job complexity. In addition, traits related to Neuroticism (e.g., locus of control and self-esteem) predict a variety of labor market outcomes, including job search effort. Many traits predict sorting into occupations, consistent with the economic models of comparative advantage discussed in Section 3. Personality traits are valued differentially across occupations.

All Big Five traits predict some health outcomes. However, Conscientiousness is the most predictive and better predicts longevity than intelligence or background. Personality measures predict health both through the channel of education and by improving health-related behavior, such as smoking.

The evidence on the effect of personality measures on crime suggests that traits related to Conscientiousness and Agreeableness are important predictors of criminality. These findings are consistent with the possibility that personality is related to social preferences as discussed in Section 6.

The survey presented in the text, even though extensive, is not fully comprehensive. We place additional material in the Web Appendix.

7.2. 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 test scores.

7.2.1 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 personality 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 1.8 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, and Hughes, 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. However, Openness to Experience 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 et al., 1996). Thus, Openness to Experience may proxy for intelligence. However, as Fig. 1.9 illustrates, controlling for measures of crystallized intelligence and fluid intelligence does not affect the coefficients on the Big Five within the third sample.[178] 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 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. This complicates the interpretation of the estimated effects of personality on schooling in older samples.

Nevertheless, the components of Openness to Experience representing an intrinsic interest in ideas and learning may affect aspects of educational achievement 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 1 year later ($\beta = 0.30$, $p < 0.05$), 2 years later ($\beta = 0.31$, $p < 0.05$), and 3 years later ($\beta = 0.26$, $p < 0.10$) but did not have higher grades in a standardized set of courses.[179] Similarly, of the Big Five, Openness to Experience is most consistently associated with fewer contemporaneously measured school absences in seventh grade ($r = -0.31$, $p < 0.01$), tenth grade

[178] Table A7 in Section A7 of the Web Appendix presents the full results from this regression. Table A8 in Section A7 of the Web Appendix presents analogous results for high-school graduation.

[179] Wong and Csikszentmihalyi (1991).

Table 1.8 The Relationship Between Years of Educational Attainment and Big Five Traits

Source	Sample	Timing of Measurement and Outcome	Controls	Metric	Results	
Goldberg, Sweeney, Merenda, and Hughes (1998)	Representative sample of US 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, father's occupational status	Standardized regression coefficient (β)	Openness	0.14^{***}
					Conscientiousness	0.05^{***}
					Extraversion	-0.07^{***}
					Agreeableness	-0.07^{**}
					Neuroticism	-0.09^{***}
German Socio-Economic Panel GSOEP (2004–2008), own calculations	Representative sample of Germans aged 21–94 ($N = 2{,}381$)	The Big Five were measured three 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% level; ***statistically significant at the 1% level.

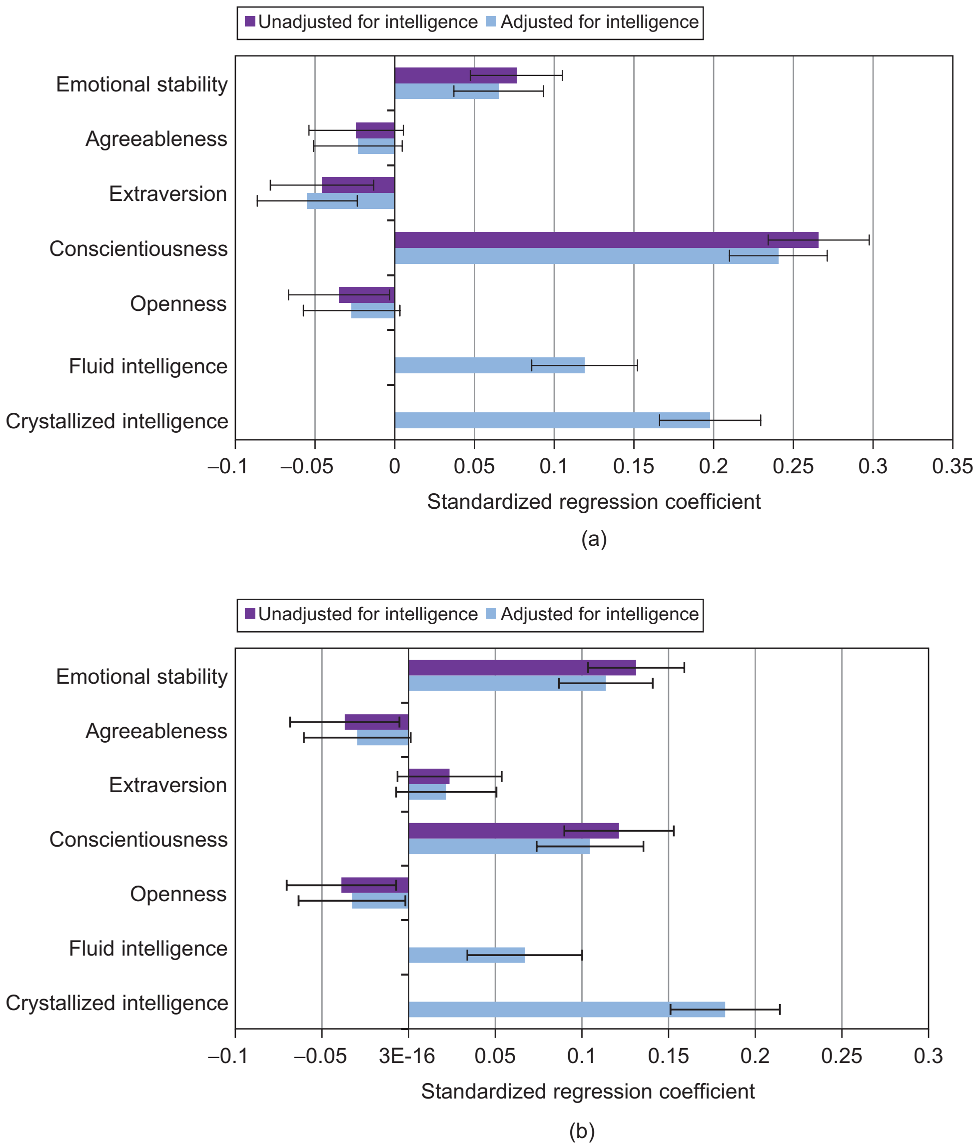

Figure 1.9 Association of the Big Five and Intelligence with Years of Schooling in GSOEP. (a) Males. (b) Females.

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

($r = -0.19$, $p < 0.01$), and twelfth grade ($r = -0.27$, $p < 0.01$; Lounsbury, Steel, Loveland, and Gibson, 2004). Still, interest in learning is not the whole story. Using prospective data, Lleras (2008) finds that controlling for cognitive ability, three behaviors associated with Conscientiousness (completing homework, working hard, arriving promptly to class) in tenth grade predicted educational attainment 10 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 1.10, taken from Heckman, Humphries, Urzua, and Veramendi (2011), 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 those of high-school dropouts.[180]

Supporting the evidence from the GED program that personality plays an important role in explaining 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, and Tremblay, 1997; Kelly and Veldman, 1964; Whisenton and Lorre, 1970).[181] Table 1.9 presents findings from three more recent studies examining the relationship between locus of control, a trait related to Emotional Stability, and high-school graduation. Although 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 a 4.5–6.8% point increase in graduating from high school. Two of the studies control for cognitive ability and find that doing so reduces the association to only 1.4–1.5%. However, the measures of cognitive ability (course grades and AFQT score) are partly determined by locus of control as discussed later in this section.

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. They correct for the influence of schooling on

[180] See the discussion of the GED program in Heckman, Humphries, and Mader (2011).
[181] See Section 5.4 for a discussion of the links between these personality facets and the Big Five traits.

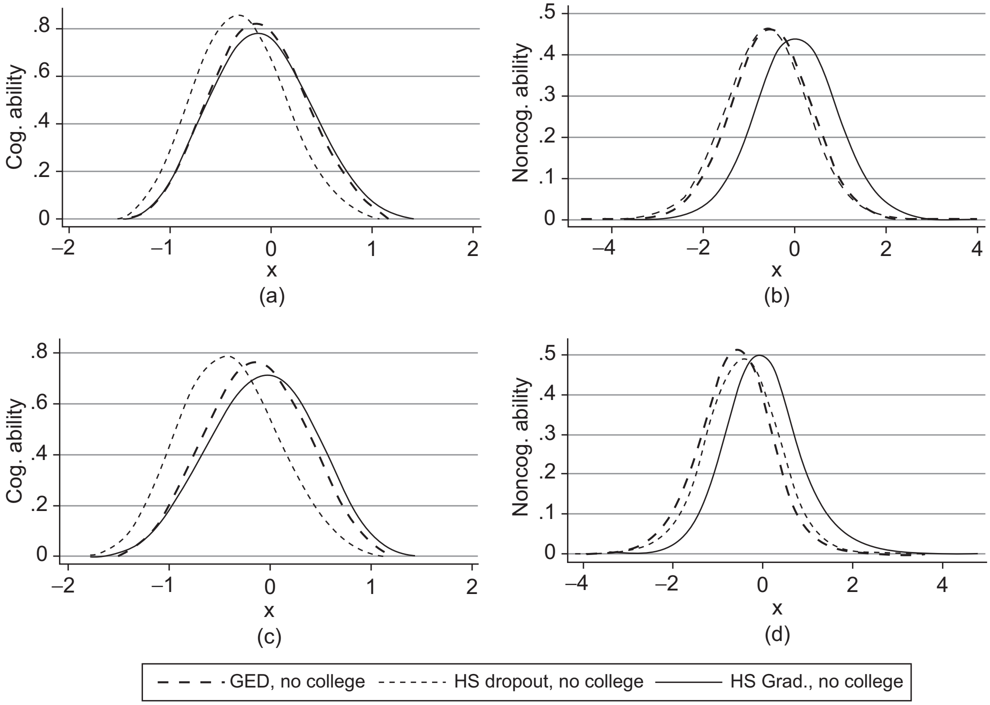

Figure 1.10 Distribution of Cognitive and Noncognitive Skills by Education Group. (a) Female Cognitive Ability (no college sample). (b) Female Noncognitive Ability (no college sample). (c) Male Cognitive Ability (no college sample). (d) Male Noncognitive Ability (no college sample).

Notes: The data is from the National Longitudinal Study of Youth 1979 (no college sample, all ethnic groups). 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, and Veramendi (2011).

personality. They 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 1.11 shows that better adolescent personality traits—as measured by locus of control and self-esteem (traits related to Neuroticism)—increases the probability of graduating from, and stopping at, high school for males at the lowest quantiles of the personality distribution. However, at higher quantiles, the probability of stopping education at high-school graduation is decreasing in measured personality because such students continue on to college. As discussed in Section 3, the effects of traits on outcomes need not be monotonic, but they can be, as Fig. 1.12 shows, where both higher cognitive and personality traits have strong

Table 1.9 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 United States (N = 1,394)	Locus of control was measured in grades 10 or 11	(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) Locus of control (2)	4.6*** 1.5
Coleman and DeLeire (2003)	Nationally representative sample of students in the United States (N = (1) 13,720 and (2) 12,896).	Locus of control was measured in grade 8.	(1) Race, gender; (2) race, gender, eighth-grade math score, eighth-grade reading score, eighth-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) Locus of control (2)	6.8 1.4**

Notes: The numbers in the "Controls" column indicate the controls used in different specifications. The numbers preceding the estimate reported in the "Results" column indicate the model used as defined in the "Controls" column.
*Statistically significant at 10% level; **statistically significant at 5% level; ***statistically significant at 1% level.

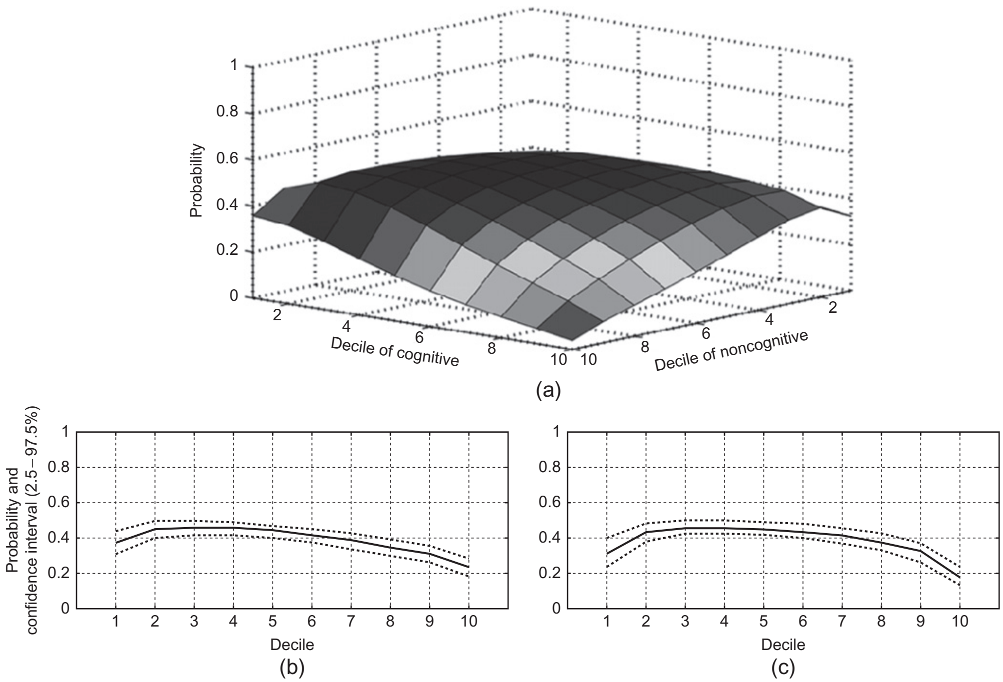

Figure 1.11 Probability of Being a High-School Graduate at Age 30 and Not Going on to Further Education, Males. (a) By Decile of Cognitive and Noncognitive Factors. (b) By Decile of Cognitive Factor. (c) By Decile of Noncognitive Factor.

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).

effects on graduating from a 4-year college at all deciles. Moving from the lowest decile to the highest decile in the measured personality distribution increases the probability of graduating from college more than a similar change in the cognitive trait 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.[182]

Cunha, Heckman, and Schennach (2010) use a dynamic factor model 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

[182] See the nonmonotonicity in Figure 1.11.

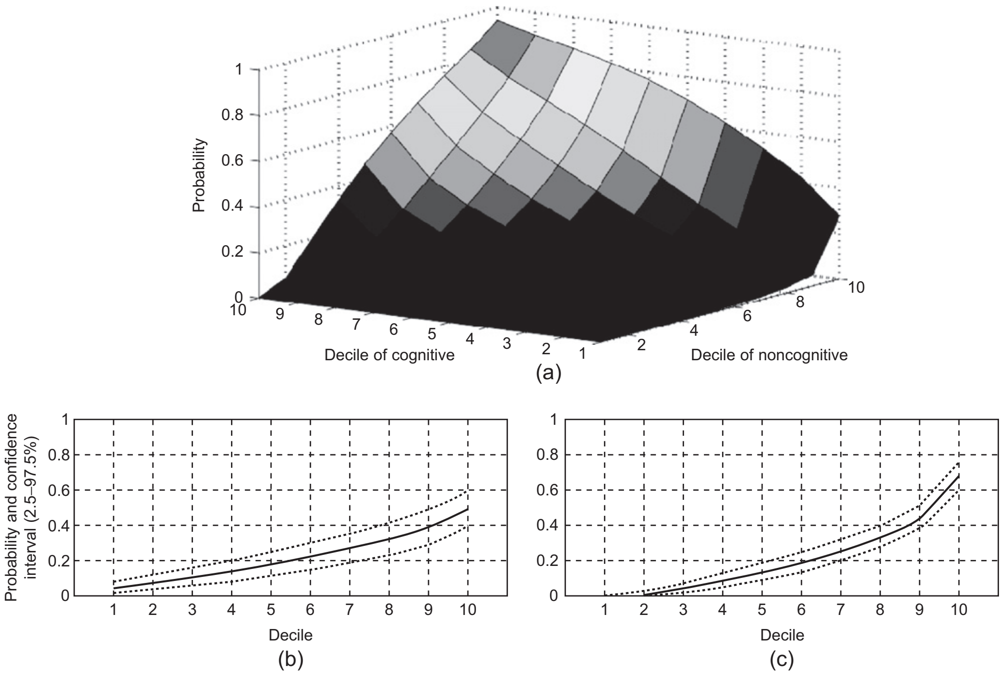

Figure 1.12 Probability of Being a 4-year College Graduate or Higher at Age 30, Males. (a) By Decile of Cognitive and Noncognitive Factors. (b) By Decile of Cognitive Factor. (c) By Decile of Noncognitive Factor.

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 (b) and (c) are evaluated at the mean of the trait not being varied.
Source: Heckman, Stixrud, and Urzua (2006, Figure 21).

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 variation.

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 measures of intelligence and demographic variables, antisocial behavior, but not attention measured in childhood, predicts high-school completion, where antisocial behavior is negatively associated with completion. Similarly, 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, Brendgen, Larose, and Tremblay (2005) examine 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 important in determining how many total years of education individuals complete in their lifetimes. 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 but should use the probability of attainment at different grades. Attention and early aggression, traits related to Conscientiousness and Agreeableness, are also predictive.

7.2.2 Course Grades

Conscientiousness is the most robust Big Five predictor of course grades, in terms of raw and partial correlations. Poropat (2009) conducted a meta-analysis of Big Five personality traits and course grades in primary, secondary, and post-secondary education, presented in Fig. 1.13. Associations between grades and Conscientiousness are almost as large as

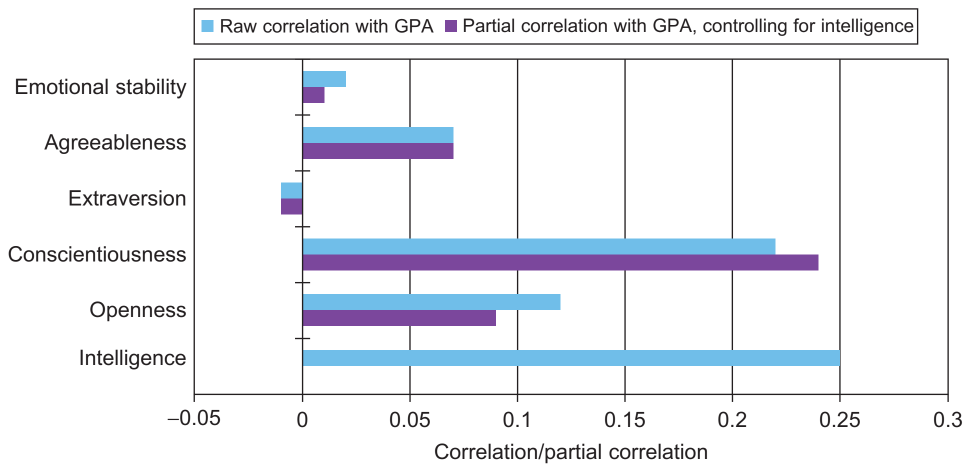

Figure 1.13 Correlations of the Big Five and Intelligence with Course Grades.

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).

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.

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 studies that do not account for halo effects. 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). Similarly, Duckworth, Tsukayama, and May (2010) 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 1.14 shows that associations between course grades and personality and cognitive ability and grades are generally stronger in the primary grades, a pattern consistent with censoring.[183] A notable exception to this trend is Conscientiousness, which has the same association with course grades at all levels. If censoring on cognitive and

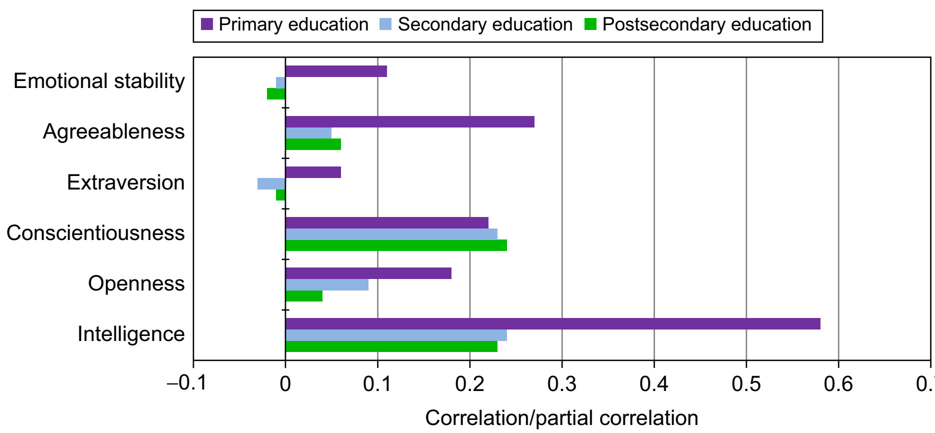

Figure 1.14 Correlations with Course Grades by Level of Education.

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).

[183] The estimated predictive validity diminishes by grade due to censoring. Censoring was not accounted for in the meta-analysis in Poropat (2009), presumably because norms for variance in representative samples are generally unavailable for personality measures (Duckworth, 2009).

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.[184] Consistent with this possibility, 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).[185]

Overall, the empirical evidence suggests that Conscientiousness may be as predictive as cognitive ability in predicting and possibly causing higher course grades. Why? Even intelligent students might not enjoy the work (Wong and Csikszentmihalyi, 1991). 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.[186]

7.2.3 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 expects, therefore, that traits related to Conscientiousness predict achievement test scores. Ample empirical evidence shows that aspects of personality predict both metrics of performance, although studies using standardized achievement tests are less common than studies using grades. As shown in Section 5, two traits related to Neuroticism, locus of control and self-esteem, explain a considerable portion of the variance of the Armed Forces Qualification Test (AFQT), an achievement test that is often used as a measure of pure intelligence in studies in economics. Similarly, Fig. 1.15 shows that in samples from three New York City middle schools, controlling for IQ, Openness to Experience is associated with Standardized Achievement Test Scores.

Roy 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). Similarly, 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 4 years later (Alexander, Entwisle, and Dauber, 1993).

[184] Cameron and Heckman (1998).

[185] The estimated correlation was corrected for truncation.

[186] Credé and Kuncel (2008); Lubbers, Van der Werf, Kuyper, and Hendriks (2010); Noftle and Robins (2007); Valiente, Lemery-Chalfant, and Castro (2007); and Valiente, Lemery-Chalfant, Swanson, and Reiser (2008).

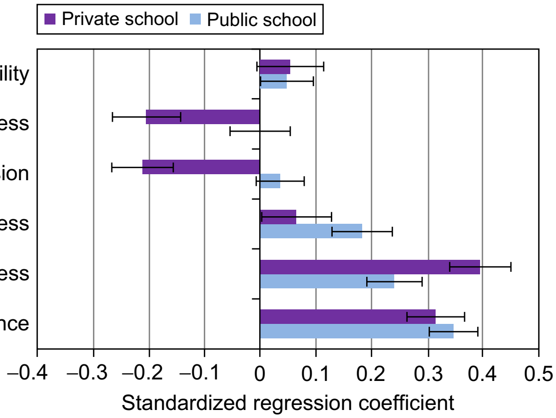

Figure 1.15 Associations with Standardized Achievement Test Scores.

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: Duckworth (2011).

More recently, in a sample of preschool children from low-income homes, parent and teacher ratings of effortful control, a facet of Conscientiousness, predicted 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 6 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 a preschool test of delay of gratification 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 require 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 et al., 2008; Ponitz, McClelland, Matthews, and Morrison, 2009). Performance on this brief task predicts later performance on standardized achievement tests (McClelland et al., 2007).

Perhaps most conclusively, Duncan et al. (2007) analyzed six large, longitudinal data sets 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 traits 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), such as locus of control, are less important for test performance.

7.2.4 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 grade point average and achievement test scores.[187] 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."[188] 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, might depend on a particular teacher's judgment.

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

[187] The correlations were even higher when the test and grades were based on similar subject matter. They use the data from the National Education Longitudinal Study (NELS).

[188] Willingham, Pollack, and Lewis (2002, p. 2).

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, and 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, Chingos, and McPherson (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. Therefore, it is not surprising that Frey and Detterman (2004) found a correlation of $r = 0.82$ ($p < 0.01$) between SAT scores and performance on the ASVAB, an aptitude and achievement test developed for the US 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, fourth-grade self-control was a stronger predictor of ninth-grade GPA ($\beta = 0.40$, $p < 0.001$) than was fourth-grade IQ ($\beta = 0.28$, $p < 0.001$). In contrast, fourth-grade self-control was a weaker predictor of ninth-grade standardized test scores ($\beta = 0.11$, $p < 0.05$) than was fourth-grade IQ ($\beta = 0.64$, $p < 0.001$). 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. Similarly, 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 1.10 presents results showing that Conscientiousness and SAT scores are similarly predictive of college GPA. However, in each of the studies, Conscientiousness was measured in college, which presents problems for a causal interpretation of this evidence due to the potential for reverse causality.

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

Table 1.10 The Predictive Power of Conscientiousness Relative to SAT Scores for College GPA

Source	Sample	Timing of Measurement and Outcome	Controls	Metric	Results	
Conard (2005)	University students in the United States ($N = 186$).	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 United States ($N = 10{,}472$).	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 United States ($N = 465$).	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 United States ($N = 444$).	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 United States ($N = 201$).	GPA and SAT were provided by the College'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 the 10% level; **statistically significant at the 5% level; ***statistically significant at the 1% level.

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).

7.3. Labor Market Outcomes

"Eighty percent of success is showing up."

Woody Allen, as quoted in Safire (1989).

It is intuitive that personality traits affect labor market outcomes. Showing up is required for completing a task. However, precisely quantifying the direct effects of personality is more difficult.[189] 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.[190] 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 require completing work on a schedule and involve intelligence to varying degrees. Therefore, it is 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 1.16 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. However, Conscientiousness 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 semiskilled or unskilled laborers (Schmidt and Hunter, 2004). In contrast, the importance of Conscientiousness does not vary much with job complexity (Barrick and Mount, 1991), suggesting that it pertains to a wider spectrum of jobs. Causality remains an open question, as it does in most of the literature in psychology. The raw correlations presented in Fig. 1.16 do not account for reverse causality, and the authors do not clearly delineate when the measures of personality were taken.

[189] Allen admits that his estimate is partially based on the fact that "80" has better cadence than "70" (Safire, 1989).

[190] Bowles, Gintis, and Osborne (2001b) discuss evidence on the association between personality traits and labor market outcomes.

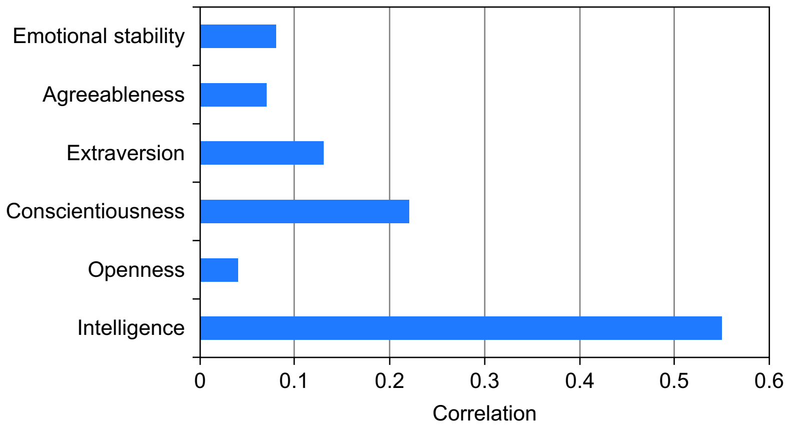

Figure 1.16 Associations with Job Performance.

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 statistically significant with a lower bound on the 90% credibility value of 0.10. The value for IQ is a raw correlation.
Source: 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. However, accounting for reverse causality is particularly important because strong evidence suggests that 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 measures of personality measured well before individuals enter the labor market and find that locus of control and self-esteem, two facets of Emotional Stability, predict wages (Judge and Hurst, 2007; Drago, 2008; Duncan and Dunifon, 1998). Table 1.11 presents results from the structural model of 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 levels. In general, their measure of noncognitive ability (personality) affects wages to a similar degree across all education levels, whereas cognitive ability tends to have little effect for GED recipients, high-school dropouts, and college dropouts.

However, more recent evidence suggests that personality affects wages mostly through the channel of educational attainment. In Section 7.1, we presented evidence that personality measures (along with measurements of cognition) are strong predictors of educational attainment. Heckman, Humphries, Urzua, and Veramendi (2011) estimate

Table 1.11 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)
Two-year college degree	.039	.368	.206	.279
	(.138)	(.209)	(.096)	(.145)
Four-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. The sample excludes 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).

a model of sequential educational choice and find that personality, as measured by participation in adolescent risky behaviors, primarily affects age 30 earnings through its effects on education. They find that given educational attainment, the effects of personality variables on a variety of outcomes are weak.[191] Further highlighting the role of traits in explaining outcomes by education level, Fig. 1.17 shows that GED recipients—who have lower levels of noncognitive skills but comparable levels of cognitive skills (see the previous section)—have lower wages, lower total wage income, and work fewer hours relative to high-school graduates, when controlling for ability. Other studies by Heckman, Stixrud, and Urzua (2006) and Cattan (2011), using other measures of personality traits, find that the traits affect earnings above and beyond their effects on education and the effects of education on earnings. Resolving these disparate findings is an important topic for future research.

These various studies have shown that personality is associated with wages, but do not explain *why* they are associated other than suggesting that the relationship occurs through the channel of educational attainment. Other mechanisms might be absenteeism, self-employment, and unemployment. Indeed, controlling for basic demographics, employment

[191] See Heckman, Humphries, Urzua, and Veramendi (2011) for a discussion of why their results differ from Heckman, Stixrud, and Urzua (2006).

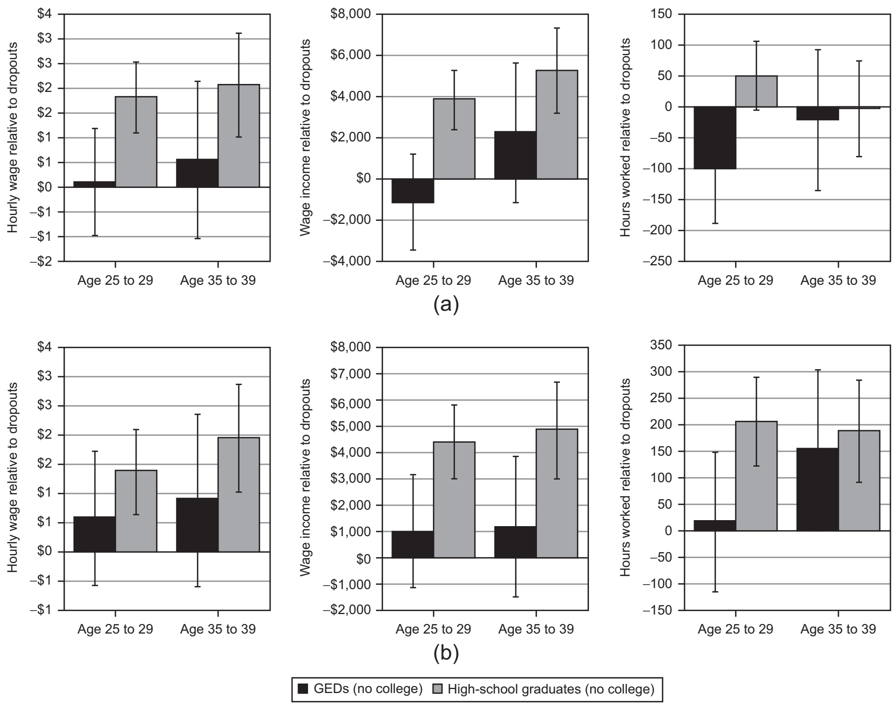

Figure 1.17 Ability-Adjusted Economic Gaps Relative to Dropouts: GEDs and High-School Graduates for (a) Males and (b) Females.

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, Heckman, and Mullen (2004) as implemented in Carneiro, Heckman, and Masterov (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/h and less than \$100/h, and individuals working less than 4000 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) as analyzed by Heckman, Humphries, and Mader (2011).

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. However, the study uses contemporaneous measures of personality and absenteeism.[192]

[192] All other Big Five traits were not statistically significant at the 10% level.

Personality plays a role outside of formal employer-employee relationships. Self-employed workers, with either very low or high levels of risk aversion, a trait related to dimensions of personality as discussed in Section 6, tend to remain self-employed for a shorter time, suggesting that they are less suited to self-employment (Caliendo, Fossen, and Kritikos, 2010).[193]

Personality could directly affect the duration of unemployment spells. Gallo, Endrass, Bradley, Hell, and Kasl (2003) find that an internal locus of control is associated with a higher probability of reemployment. A couple of studies have explicitly incorporated locus of control into standard job search models. For example, Caliendo, Cobb-Clark, and Uhlendorff (2010) examine whether a higher locus of control increases the perceived marginal benefit of exerting search effort so that people with a greater 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 a 5.3% increase in the number of job applications submitted ($p < 0.01$), controlling for demographic characteristics and past employment history.[194] Although the measures were contemporaneous, the respondents became unemployed near the time that the locus of control was measured, potentially limiting the role of reverse causality. Similarly, McGee (2010) proposes a model in which people with a higher locus of control believe that search effort has a higher return. His model predicts that those with an internal locus of control search more intensely but have higher reservation wages so that the effect on the hazard rate of leaving unemployment is ambiguous. In line with his predictions, he finds that a one–standard deviation increase in locus of control, measured before market entry, is associated with a 1.3% increase in the reservation wage ($p < 0.01$) and a 20% increase in the time spent searching for a job per week ($p = 0.14$).[195] Those with moderate levels of locus of control have the highest hazard rates for leaving unemployment. Consistent with the interpretation that locus of control affects beliefs (not productivity), locus of control has no effect on reemployment wages when controlling for reservation wages.

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 nonpecuniary benefits associated with particular occupations. Supporting this notion, Conscientiousness (Barrick and Mount, 1991, and Ham, Junankar, and Wells, 2009), locus of control and self-esteem (Heckman, Stixrud and Urzua, 2006) predict sorting

[193] Caliendo, Fossen, and Kritikos (2010) use measures of risk-aversion from 2004 and employment status from 2000 to 2005, assuming that risk aversion did not change over this period.

[194] The associations were partially mediated when controlling for the Big Five, suggesting that locus of control overlaps with the Big Five as discussed in Section 5.

[195] The effect on the reservation wage is higher for people looking for their first jobs.

into occupations. However, these studies use relatively broad occupational categories that might obscure more nuanced influences of personality. Analyzing 18 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 the probability of being a manager ($p < 0.01$) and a 2.9% decrease in being a business professional ($p < 0.01$). A standard-deviation increase in internal locus of control is associated with 2.8% decrease in the probability of 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 ($p < 0.01$).[196,197]

Furthermore, the value of cognitive ability and personality differs by occupation just as it does by education. Cattan (2011) estimates a structural model of comparative advantage along the lines discussed in Section 3 and finds that different skills are valued differently, depending on the occupation. Accounting for selection, a one-standard deviation increase in adolescent sociability (related to Extraversion) leads to a 7% increase in the wages of managers ($p < 0.01$), a 4% increase in the wages of sales and service workers ($p < 0.01$), but leads to a 2% ($p < 0.05$) decrease in the wages of professionals and has no significant impact on the wages of blue-collar and clerical workers. Self-esteem and locus of control are positively valued in all occupations, but the magnitudes also depend on the occupation. The effects of traits are not uniform on wages across occupations even after controlling for schooling.

Personality affects not only the occupational selection but also the type of compensation scheme selected within an 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, are 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, and Sunde (2009) find that in a German sample self-reported positive reciprocity is positively associated with income, employment, and working overtime. Negative reciprocity tends to operate in the opposite direction. As discussed in Section 6, these measures of social preference relate to personality.[198]

[196] The data for occupational categories came from 2001 to 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 potentially important.

[197] They find other statistically significant results at the 5% and 10% levels, which we omit for brevity.

[198] Agreeableness and Conscientiousness are associated with greater positive reciprocity and lower negative reciprocity, whereas Neuroticism is associated with greater negative reciprocity (Dohmen, Falk, Huffman, and Sunde, 2008).

In sum, there are good theoretical reasons as well as some empirical evidence that personality affects labor market outcomes through channels other than education. Conscientiousness and Neuroticism are associated with job performance and wages to a similar but lesser degree than cognitive ability. Personality traits are more important for people with lower levels of job complexity or education level, whereas 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 types of compensation. Table A10 in Web Appendix A7 summarizes a variety of studies that associate personality with labor market outcomes.

7.4. Personality and Health[199]

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.[200] Consistent with Hippocrates' ideas, recent 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 research shows that personality measures predict longevity. Roberts, Kuncel, Shiner, Caspi, and Goldberg (2007) review evidence from 34 different studies on the predictive validity of the Big Five personality traits, relative to that of cognitive ability and socioeconomic status. 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 correlation coefficients that can be compared across studies. As shown in Fig. 1.18, Conscientiousness was a stronger predictor of longevity than any other Big Five trait and a stronger predictor than either IQ or socioeconomic status.[201] In general, traits related to Conscientiousness, Openness to Experience, and Agreeableness are associated with longer lives, whereas those related to Neuroticism are associated with shorter life spans.[202] However, the magnitudes of the relationships vary across studies and not all results are replicable. Although the specific channels through which personality affects longevity and health are largely unknown, several studies provide some clues.

[199] This section is a summary of Pietro Biroli's extensive discussion of personality and health that is presented in Web Appendix A7.1.
[200] See Hampson and Friedman (2008) and Friedman (2007) for a brief historic review.
[201] The timing of the measurements of personality relative to the outcomes varies by study.
[202] See Martin, Friedman, and Schwartz (2007); Kern and Friedman (2008); Mroczek and Spiro (2007); Boyle et al. (2005); Schulz, Bookwala, Knapp, Scheier, and Williamson (1996); and Kubzansky, Sparrow, Vokonas, and Kawachi (2001).

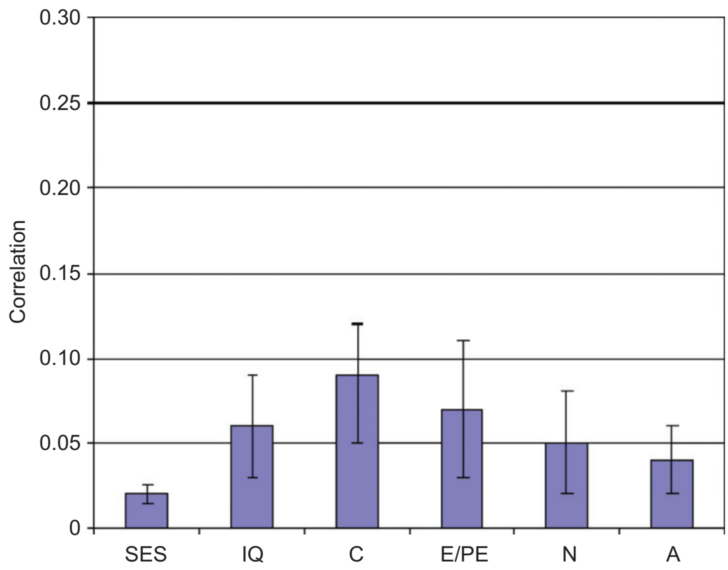

Figure 1.18 Correlations of Mortality with Personality, IQ, and Socioeconomic Status (SES).

Notes: The figure represents results from a meta-analysis of 34 studies. Average effects (in the correlation metric) of low socioeconomic status (SES), low IQ, low Conscientiousness (C), low Extraversion/Positive Emotion (E/PE), Neuroticism (N), and low Agreeableness (A) on mortality. Error bars represent standard error. The lengths of the studies represented vary from 1 year to 71 years.
Source: Roberts, Kuncel, Shiner, Caspi, and Goldberg (2007).

Personality may affect health-related behavior, such as smoking, diet, and exercise. For example, Hampson, Goldberg, Vogt, and Dubanoski (2007) find that high scores of teacher assessments of Extraversion, Agreeableness, and Conscientiousness during elementary school predict overall health behaviors during midlife (less smoking, more exercise, better self-rated health) and indirectly affect health through educational attainment.[203] The effects 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, Luyckx, and Mroczek, 2010). As Fig. 1.19 illustrates, Heckman, Stixrud and Urzua (2006) find that their personality factor affects the probability of daily smoking for males. The gradient is steepest at the high and low quantiles of the distribution.

[203] Conti, Heckman, and Urzua (2010a,b) and Heckman, Humphries, Urzua, and Veramendi (2011) present evidence on the causal relationship between education and health, and also survey the previous literature on this question.

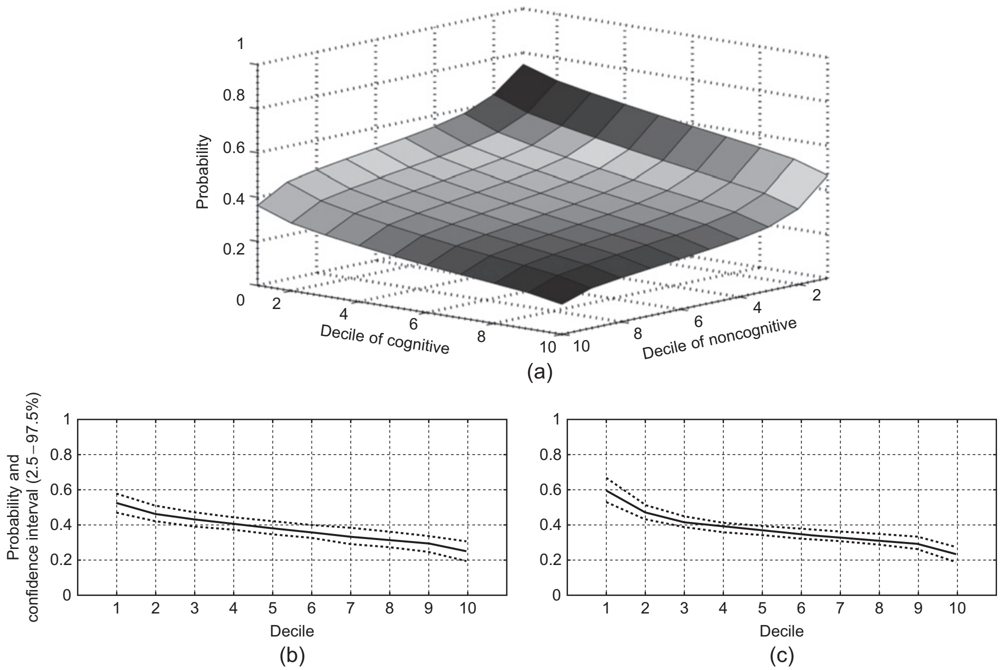

Figure 1.19 Probability of Daily Smoking by Age 18 for Males. (a) By Decile of Cognitive and Noncognitive Factors. (b) By Decile of Cognitive Factor. (c) By Decile of Noncognitive Factor.

Notes: The data are simulated from the estimates of the model of Heckman, Stixrud, and Urzua (2006) and their NLSY79 sample. They 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).

Few studies 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.[204] Some studies investigate the mechanisms by which personality affects health by considering how initial endowment of traits and health affect midlife 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

[204] Pesonen et al. (2008); Ryden et al. (2003); Sell, Tooby, and Cosmides (2009); and Hoffman, Fessler, Gneezy, and List (2010).

social interactions at age 70. They find that Neuroticism is associated with worse health for women but better health for men. Most studies do not account for the possibility that health and personality exhibit dynamic complementarities over the life cycle.

Several studies have controlled for reverse-causality by using structural models to estimate the life-cycle evolution of health. Using a structural model, Conti and Heckman (2010) estimate the causal relationship between personality traits, initial health endowments, and schooling choices and postschooling outcomes. Children 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 50% of the difference in poor health, depression, and obesity at age 30, observed between the educated and less educated. Figure 1.20 shows that for males, personality and health endowments are more predictive than are cognitive endowments, whereas for females, all three are roughly equally predictive. Using similar methods, Savelyev (2010) finds that both Conscientiousness (measured in youth) and higher education increase survival through age 80, but these traits serve as substitutes for each other so that effects of education are strongest 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 even more so in some cases.

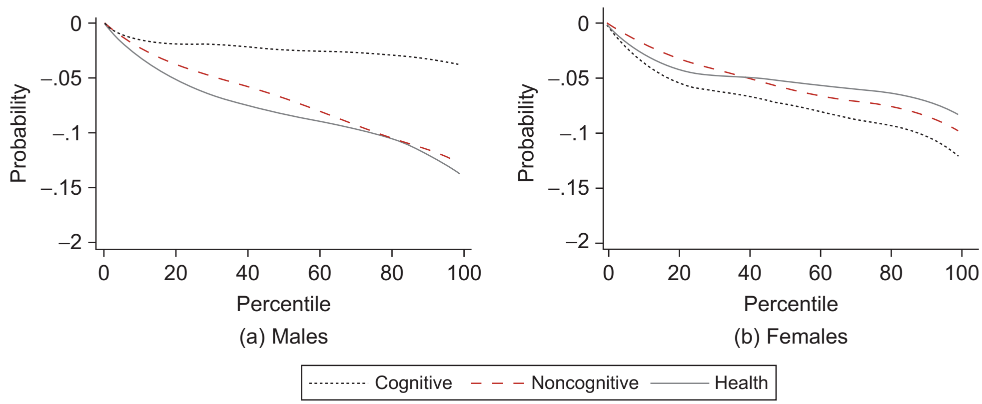

Figure 1.20 Effects of Cognitive, Noncognitive, and Health Endowments on Self-rated Health (A Lower Number Corresponds to a Better Outcome). (a) Males. (b) Females.

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 means.
Source: Conti and Heckman (2010).

7.5. Crime[205]

Few studies have examined the relationship between the Big Five and criminal behavior. The available evidence suggests that Big Five Conscientiousness and Agreeableness are important protective factors against criminal activity. Figure 1.21 illustrates that in a sample of at-risk youth, boys who had committed severe delinquent behaviors were more than three quarters 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, and Moffitt, 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 nonverbal" (Gottfredson and Hirschi, 1990, p. 90). 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

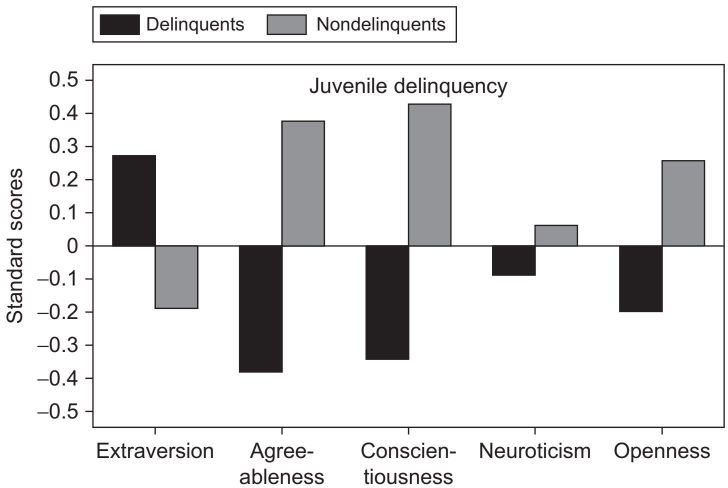

Figure 1.21 Juvenile Delinquency and the Big Five.

Notes: Delinquents are those who have committed at least one of the following: breaking and entering, strong-arming, or selling drugs. Nondelinquents 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, and Moffitt (1994).

[205] This section summarizes the more comprehensive survey of the literature on personality and crime prepared by Amanda Agan. See Web Appendix Section A7.2 for her survey.

of self-control explained between 10% and 16% of the variance in contemporaneously measured theft, assault, drug use, and vandalism (Vazsonyi, Pickering, Junger, and Hessing, 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[206] and sensation seeking and impulsive behavior were 0.27 and 0.13, respectively, when controlling for peer behavior and measures of risk appraisal (Horvath and Zuckerman, 1993).

Self-control might not be the entire story. Negative emotionality—a tendency toward depression likely related to 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 blacks ($p < 0.05$) in one sample (Caspi et al., 1994) to $r = 0.22$ ($p < 0.01$) in another sample (Agnew, Brezina, Wright, and Cullen, 2002). None of these studies control for cognitive ability or address causality.

An emerging literature investigates the causal effects of education on crime. Heckman, Stixrud, and Urzua (2006) estimate a causal model of personality and education accounting for reverse causality. They find that both cognitive traits and noncognitive traits, as captured by locus of control and self-esteem, are affected by schooling.[207] These traits in turn are approximately equally predictive of criminal activity.[208]

Using changes in compulsory schooling laws as an instrument, Lochner and Moretti (2004) and Machin, Marie, and Vujić (2010) find that years of education are negatively associated with criminal activities in the United States and United Kingdom, respectively. In a structural model of skill production, 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

In this section, we review empirical 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 ontogeny (programmed developmental processes common to all persons) and sociogeny (shared socialization processes), and show how aspects of personality, such as sensation seeking, evolve as the brain develops.

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

[207] We discuss this work in Section 8.

[208] Their measure of prediction is the effect of decile improvements of cognition and personality traits on the probability of being in jail.

Second, we show how personality changes through external forces that operate through alterations in normal biology, such as brain lesions and chemical interventions. Third, and most relevant for policy, we show that education, early childhood interventions, and parental investment can affect personality throughout the life cycle. We also discuss the less-abundant evidence on the malleability of preferences.

8.1. 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 measures of a trait over time. *Rank-order change*, in contrast, refers to changes in the ordinal ranking of a trait in a population and is measured by rank correlations among longitudinal measures. One commonly held view is that rank-order, as well as 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, and Roberts and Mroczek, 2008).

During the early years of life, mean-level changes in measured 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 measured 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 1.22 shows that people typically become more socially dominant, conscientious, and emotionally stable (nonneurotic) across the life cycle, whereas social vitality and Openness to Experience rise early in life and then decrease in old age.[209] Surprisingly, after childhood, the greatest mean-level change in most measured personality traits takes place not during adolescence but rather in young adulthood.

[209] Figure A3 in Section A9 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.

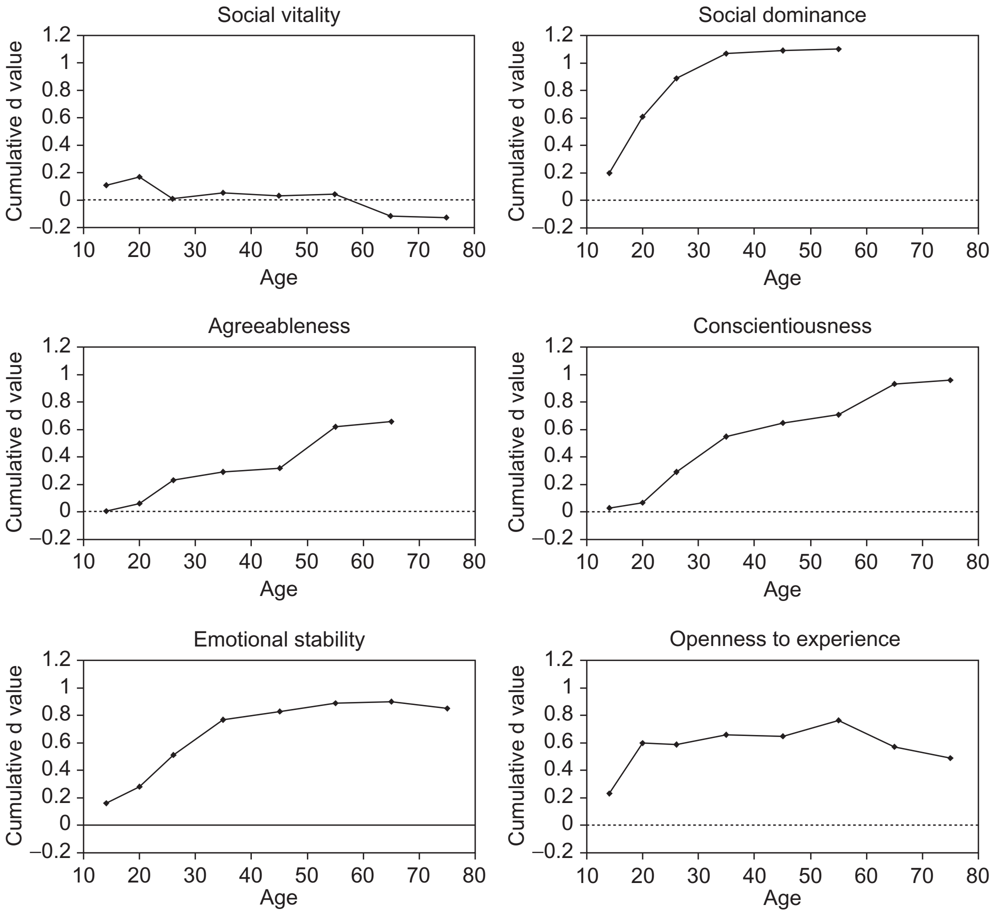

Figure 1.22 Cumulative Mean-Level Changes in Personality across the Life Cycle.

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

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). Figure 1.23a shows mean-level changes in cognitive skills using a longitudinal analysis, and Figure 1.23b shows mean-level changes using a cross-sectional analysis.[210] As

[210] Cross-sectional estimates of mean-level change are biased by cohort effects (e.g., 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.

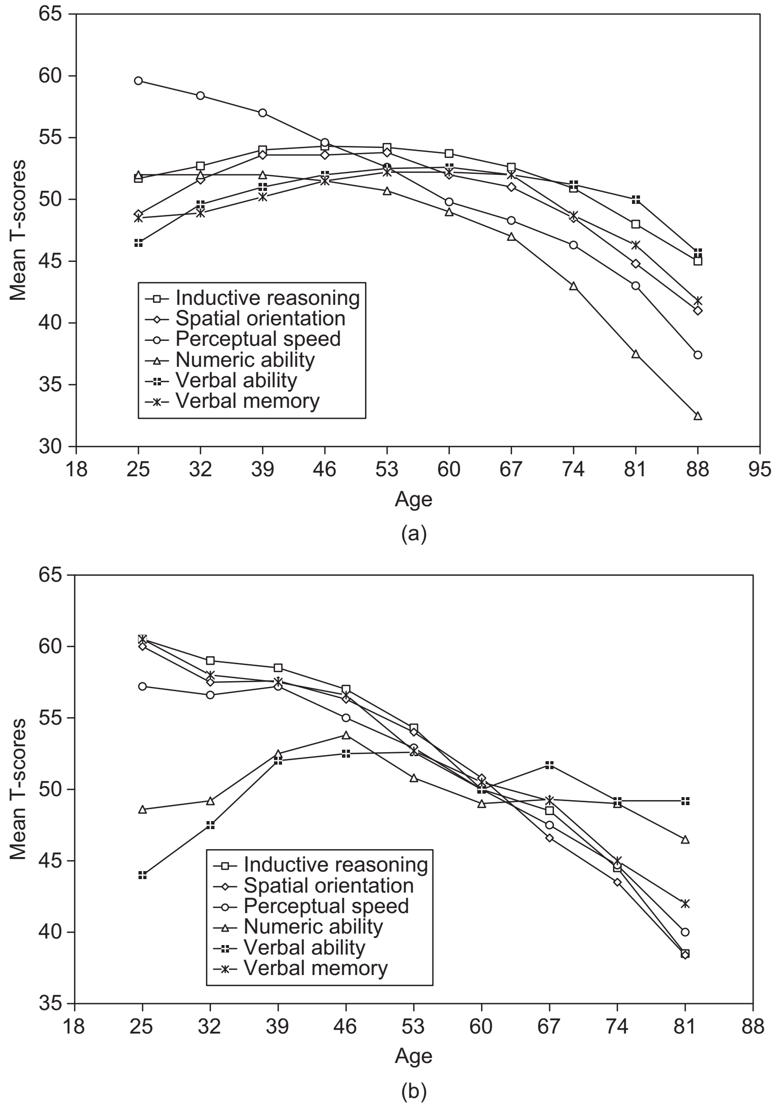

Figure 1.23 (a) Longitudinal Analysis and (b) Cross-Sectional Analysis of Mean-Level Change in Cognitive Skills over the Life-Span.

Note: T-scores on the y-axis are standardized scores with a mean of 50 and a standard deviation of 10.
Source: Figures taken from Schaie (1994), used with permission of the publisher.

schematically illustrated in Fig. 1.24, fluid intelligence decreases and crystallized intelligence increases 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.

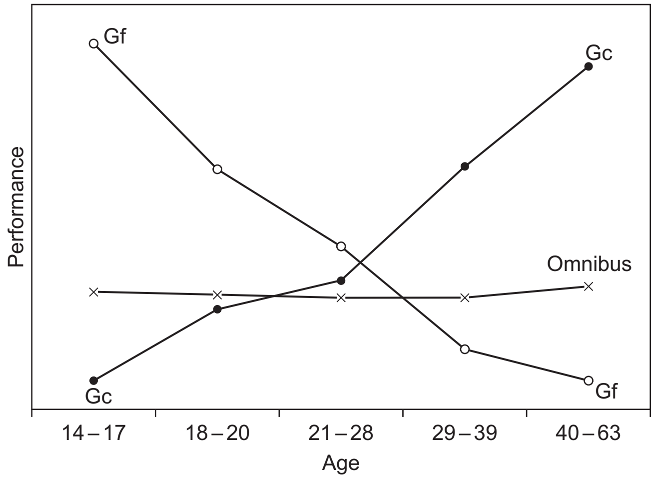

Figure 1.24 Fluid Intelligence Decreases and Crystallized Intelligence Increases across the Life-Span.

Source: Figure from Horn (1970), used with permission of Elsevier.

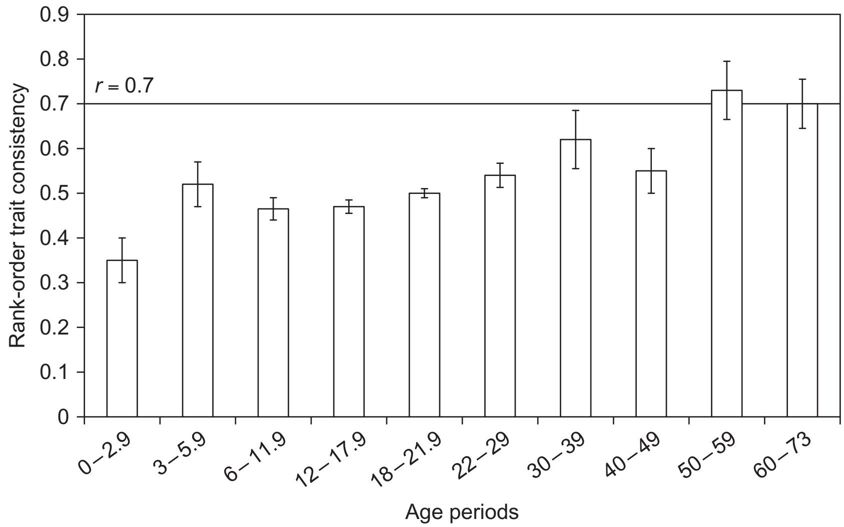

Figure 1.25 Rank-Order Stability of Personality over the Life Cycle.

Notes: The meta-analysis 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.

Rank-order stability in measured personality increases steadily over the life span. Figure 1.25 shows that 7-year test–retest stability estimates for personality plateau far from unity, at $r = 0.74$, which is about the same level as terminal stability estimates for IQ (Roberts and DelVecchio, 2000). However, measured personality does not reach

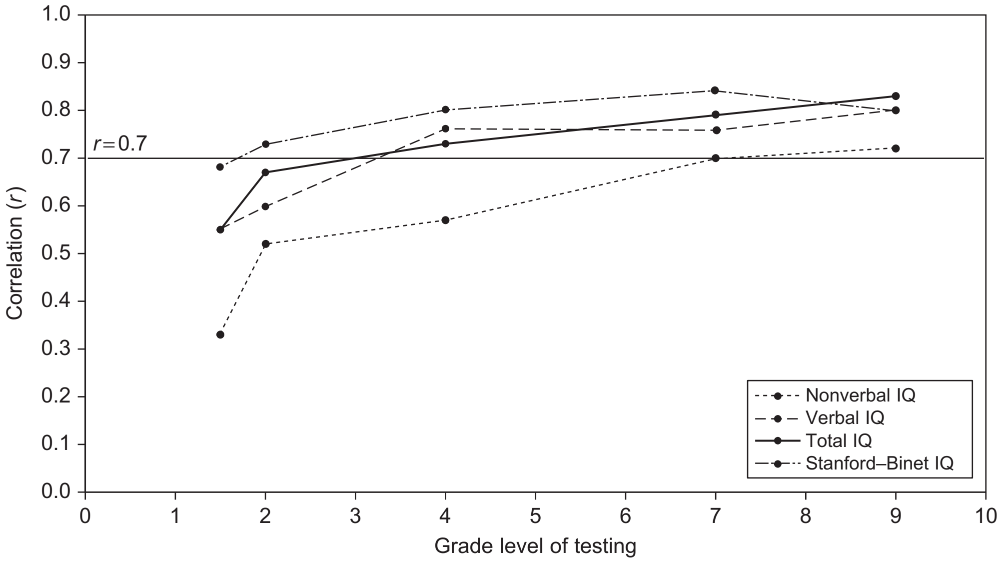

Figure 1.26 Rank-Order Stability of IQ across the Life Span.

Notes: The points represent 10-year, test–retest correlations over 10-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.

this plateau until at least age 50, whereas IQ reaches this plateau by age six or eight (Hopkins and Bracht, 1975, and Schuerger and Witt, 1989). Figure 1.26 shows rank-order stability of IQ over broad age ranges.

8.2. Evidence on Ontogenic and Sociogenic Change

A useful dichotomy contrasts *normative change*, defined as changes that are caused either by biological programming (ontogenic) or by predictable changes in social roles (sociogenic), with *nonnormative change*, encompassing both intentional change, caused by deliberate, self-directed efforts, deliberately chosen changes in social roles and atypical life events (trauma, for example).[211]

If, as McCrae and colleagues have claimed, normative changes reflect genetically programmed processes, then investment should not affect change. The current literature in psychology claims that genetic factors are largely responsible for *stability* in personality in

[211] "Normative" in this context refers to what most people, or average persons, experience. If most people deliberately do something that causes change, it would be normative. But that seems unlikely. Therefore, most deliberative change is nonnormative, but logically this is not necessarily true.

adulthood, whereas environmental factors are mostly responsible for *change* (Blonigen, Hicks, Krueger, Patrick, and Iacono, 2006, and Plomin and Nesselroade, 1990).[212,213]

In a longitudinal study of twins surveyed at age 20 and then again at age 30, about 80% 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 measured personality was mostly attributed to environmental factors. Helson, Kwan, John, and Jones (2002), for example, document the substantial influence that social roles and cultural milieu can have on personality development. Their analysis 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, 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, and historical and cultural milieu.[214] Increases in the heritability of IQ from childhood (about 40%) to adulthood (estimates range from 60–80%) 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, and Lykken, 1993; Plomin, DeFries, Craig, and McGuffin, 2002).[215] Heritability estimates for Big Five traits are relatively stable across the life cycle at about 40–60% (Bouchard and Loehlin, 2001).[216] Behavioral genetics studies typically estimate the effect of common parental environments on adult measures of outcomes to be near zero, but Turkheimer, Haley, Waldron, D'Onofrio, and Gottesman (2003) find estimates from such studies to be biased downward by the overrepresentation of middle- and upper-class families. Among poor families, Turkheimer et al. find that 60% of the variance in IQ is accounted for by shared environment. In addition, he finds that heritability estimates are much smaller than they are for affluent families, whereas among affluent families, the contribution of heritability is much larger. Krueger and colleagues have

[212] Plomin and the essays in the December issue of *Monographs for the Society for Research in Child Development* (Kovas et al., 2007) extend this analysis to childhood.

[213] 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.

[214] We note here that while genes remain constant through the life cycle, the expression of genes is determined, in part, by experience.

[215] 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 on conditions in the maternal womb. See also Rutter (2006b) and an emerging literature on epigenetics.

[216] 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.

recently demonstrated that other moderators also influence the heritability of traits (see Krueger, South, Johnson, and Iacono, 2008).[217]

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, 2006a). 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 common.[218]

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).[219] 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).[220]

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, 2006a). 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). Lizzeri and Siniscalchi (2008) develop an economic model of differential parenting of siblings.

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) speculates that typical biological (ontogenic) development explains the surge of risk taking in adolescence followed by the decline in adulthood. Figure 1.27 illustrates

[217] It is important to note that shared environment is not the same as environment. Children may be treated individually by parents.

[218] 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 et al., 2003; Moffitt, Caspi, and Rutter, 2005; and Caspi et al., 2002).

[219] 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, e.g. Some caution must be taken in relying on the claims in this literature. Blair, Gamson, Thorne, and Baker (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.

[220] 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).

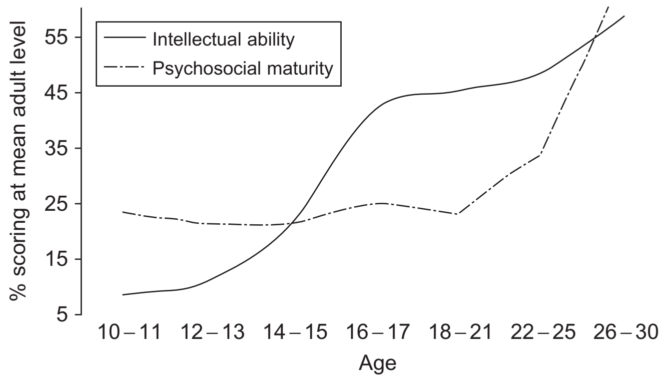

Figure 1.27 Proportion of Individuals in Each Age Group Scoring at or Above the Mean for 26- to 30-Year Olds on Indices of Intellectual and Psychosocial Maturity.

Source: From Steinberg, Cauffman, Woolard, Graham, and Banich (2009) submitted for publication.

his conjecture about how basic intellectual ability and psychosocial maturity (related to, e.g., impulsivity, risk perception, sensation seeking, future orientation) evolve over the life cycle.[221] He argues that intellectual ability matures more rapidly than psychosocial maturity. In his model, increases in adolescent risk taking are 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 to the development of the brain's cognitive control system, specifically improvements in the prefrontal cortex that 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 partially depends on the presence of peers, which corresponds to aspects of the situation (*h* in the framework of Section 3). This example highlights the difficulty in disentangling situational and biological changes in personality.

What factors 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 (e.g., getting a job for the first time or 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 measured personality. If, on the other hand, these social roles are not assumed synchronously, we will observe rank-order changes.

One difficulty with many of the studies that address this question is the problem of reverse causality. Changes in personality may drive social role changes rather than the other way around.

[221] Spear (2000a,b) also finds that sensation seeking reaches its peak in adolescence.

8.3. External Changes to Biology

The previous subsection discusses the difficulty in disentangling biological changes in personality from environmental or situational effects. In this subsection, we provide some evidence on causal changes in personality due to external forces that either damage parts of the brain or abruptly alter the chemistry of the brain.

8.3.1 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 but preserved his problem-solving abilities (Damasio, Grabowski, Frank, Galaburda, and Damasio, 2005). Since then, there have been many more case studies of patients with brain damage. For example, Mataró 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 antisocial behavior, suggesting that personality is malleable in different dimensions, even through brain damage. The effects of brain damage are persistent. After 5 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 might allow people to assign and store value to particular outcomes in a way that is useful for decision making. The amygdala is believed to signal impulsive emotional responses to immediate environmental stimuli, such as reacting quickly to a snake. In contrast, the VMPC is believed 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. These findings are consistent with McClure, Laibson, Loewenstein, and Cohen's β/δ system that describes hyperbolic discounting

(McClure, Laibson, Loewenstein, and Cohen, 2004). However, some recent research in neuroscience challenges this theory and presents empirical evidence that contradicts β/δ theory (Monterosso and Luo, 2010).

Further experiments involving these parts of the brain highlight why attempts to separate cognitive and noncognitive traits might be futile. For example, Bechara and Damasio (2005) study the performance of patients with lesions in the VMPC in a seemingly cognitive task. Participants in their experiment were given the Iowa Gambling Task, in which they repeatedly chose between four decks of cards that represented lotteries of different value, unknown to the participant at the onset. Throughout the experiment, the authors also measured skin conductance responses (SCRs), a known physiological reflection of emotion. By trial and error, participants without lesions learned to choose the "better" decks of cards with lower short-term payoffs but higher average payoffs. The normal participants also showed emotional activity both when picking their card and when receiving the rewards or penalties. 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. 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 future consequences. These findings suggest that emotion helps to guide decision making. Numerous other studies show the role of the amygdala in signaling emotions and its relationship to cognition and behavior (Phelps, 2006).

8.3.2 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 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, and Fehr, 2005). As discussed in Section 5, the Big Five traits are linked to personality disorders. 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 paroxetine might have a long-lasting impact through a biochemical change in the brain (Tang et al., 2009). Similarly, Knutson et al. (1998) find evidence that paroxetine can diminish hostile behavior through a decrease in general negative effect.

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

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. A growing body of evidence suggests that education, parental investment, and interventions can causally affect personality traits. More than just ontogenic and sociogenic processes are at work. A major contribution of economics to the literature in psychology is to develop and apply a framework to investigate how investment, including education, work experience, and self help, changes traits. We discuss the evidence on trait changes through these mechanisms, using the theoretical framework introduced in Section 3.8 as a guide. In all of the models considered in this subsection, the development of traits arises from purposive actions of agents and not just from exogenous biological processes.

The empirical literature has not estimated the investment model (1.16) in Section 3.8 in its full generality. It focuses on estimating productivity functions (1.1) specified in terms of traits θ. Due to data limitations, there is no empirical work yet to report that standardizes for effort or for situation. To simplify the notation, we keep h implicit.

Denote the productivity traits at age v by θ^v. Substituting for actions in terms of their determinants, the performance on task j at age v is

$$P_j^v = \phi_j^v(\theta^v, e_j^v),\ j \in \{1, \ldots, J\}, v \in \mathcal{V} \tag{1.22}$$

where e_j^v is effort devoted to task j at time v. For simplicity, break θ^v into cognitive, μ, and personality, π, components:

$$\theta^v = (\theta_\mu^v, \theta_\pi^v),$$

using the notation of Section 3.[222] e_j^v depends on preferences, rewards and information.

The vector of productivity traits evolves via a simplified version of (1.16):

$$\theta^{v+1} = \eta^v(\theta^v, IN^v, h^v),\ v \in \mathcal{V}. \tag{1.23}$$

IN^v is interpreted very broadly to include investment by parents, schools, work experience, and interventions. θ^0 is a vector of initial endowments. Some components of effort may be included in investment.

The productivity of investment can depend on the age at which it is made. A crucial feature of the technology that helps to explain many findings in the literature on

[222] See Eq. (1.15).

skill formation (see Cunha and Heckman, 2007, 2009) is *complementarity of traits with investment*:

$$\frac{\partial^2 \eta^v(\theta^v, IN^v, h^v)}{\partial \theta^v \partial (IN^v)'} \geq 0. \tag{1.24}$$

Technology (1.23) is characterized by *static complementarity* between period v traits and period v investment. The higher the θ^v, the higher the productivity of the investment. There is also *dynamic complementarity* if the technology determines period $v+1$ traits (θ^{v+1}). This generates complementarity between investment in period $v+1$ and investment in period s, $s > v+1$. Higher investment in period v raises θ^{v+1} because technology is increasing in IN^v, which in turn raises θ^s because the technology is increasing in θ^v, between v and s. This, in turn, increases $\frac{\partial \eta^s(\cdot)}{\partial IN^s}$ because θ^s and IN^s are complements, as a consequence of (1.24).

Dynamic complementarity explains the evidence that early nurturing environments affect the ability of animals and humans to learn. It explains why investments in disadvantaged young children are so productive (see Knudsen, Heckman, Cameron, and Shonkoff, 2006). 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^v is low.[223] Using dynamic complementarity, one can define *critical* and *sensitive* periods for investment. If $\frac{\partial \eta^v(\cdot)}{\partial IN^v} = 0$ for $v \neq v^*$, v^* is a critical period for that investment. If $\frac{\partial \eta^v(\cdot)}{\partial IN^v} > \frac{\partial \eta^{v'}(\cdot)}{\partial IN^{v'}}$ for all $v' \neq v$, v is a sensitive period.[224] The technology of skill formation is consistent with a body of evidence that shows critical and sensitive periods in human development for a variety of traits.[225]

Figure 1.28 shows how adult outcomes are shaped by *sequences* of investments over the life cycle. The importance of the early years depends on how easy it is to compensate for adverse early effects with later investment. The literature shows that resilience and remediation are possible, but are more costly later on.[226] 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.

For the purposes of policy analysis, it is important 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

[223] See the evidence in Cunha and Heckman (2007); Heckman (2007); Heckman (2008b); and in Cunha, Heckman, Lochner, and Masterov (2006).

[224] This expression is evaluated at common levels of the inputs on both sides of the expression.

[225] See the evidence summarized in Heckman (2008b); Cunha and Heckman (2009); and Cunha, Heckman, Lochner, and Masterov (2006).

[226] See Cunha and Heckman (2007, 2009); Cunha, Heckman, and Schennach (2010); and Heckman (2008b).

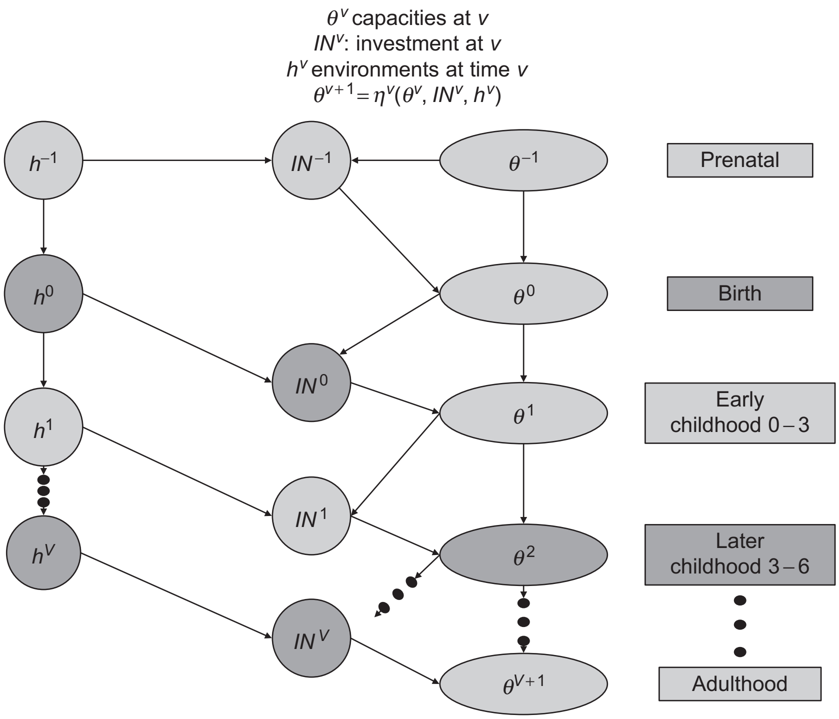

Figure 1.28 A Life Cycle Framework for Organizing Studies and Integrating Evidence: $V+1$ Period Life Cycle. θ^v: Capacities at v. IN^v: Investment at v. h^v: environments at time v. $\theta^{v+1} = \eta^v(\theta^v, IN^v, h^v)$.

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.

The most general empirical specification of the technology to date is that of Cunha, Heckman, and Schennach (2010). They allow for $Q+1$ different developmental stages in the life of the child: $q \in \{0,\ldots,Q\}$. Developmental stages may be defined over specific ranges of ages, $v \in \{0,\ldots,V\}$, so $Q \leq V$. They assume that each component of θ^v and IN^v can be represented by a scalar as can environment h^v.[227] Letting IN_k^v be investment in trait k at age v, they estimate a CES, stage-specific version of (1.23) for trait k at stage q:

$$\begin{aligned} \theta_k^{v+1} &= \left[\gamma_{\mu,k}^q(\theta_\mu^v)^{\sigma_k^q} + \gamma_{\pi,k}^q(\theta_\pi^v)^{\sigma_k^q} + \gamma_{IN,k}^q(IN_k^v)^{\sigma_k^q} + \gamma_{E,k}^q(h^v)^{\sigma_k^q}\right]^{\frac{1}{\sigma_k^q}}, \\ \gamma_{m,k}^q &\geq 0, \quad \sum_{m\in(\mu,\pi,IN,h)} \gamma_{m,k}^q = 1 \text{ for all } k \in \{\mu,\pi\} \text{ and } q \in \{0,\ldots,Q\}. \end{aligned} \tag{1.25}$$

[227] For them, environment is parental environment.

A main finding of Cunha, Heckman, and Schennach (2010) is that the elasticity of substitution σ_μ^q governing the acquisition of cognitive traits decreases with q. This is consistent with other evidence that shows the declining malleability of cognition with age, that is, cognitive deficits are easier to remedy at early ages than at later ages. At the same time, σ_π^q, associated with personality, stays roughly constant over q. This is consistent with evidence on the emergence of psychological maturity, as shown in Figure 27.[228]

Adjoined with measurement systems for productivity on tasks in period v (Eq. (1.22)), the econometric model is a "state space" model that accounts for errors in measurements and endogeneity of inputs. Cunha and Heckman (2008) and Cunha, Heckman, and Schennach (2010) estimate these models on panel data on the growth dynamics of individuals and show that accounting for measurement error and endogeneity is empirically important.

Cunha, Heckman, and Schennach (2010) estimate technology (1.25) using longitudinal data on the development of children with rich measures of parental investment and of child traits. They examine the estimated substitution parameters to examine the issue of the cost of remediating early disadvantage at later ages. Their findings shed light on the dynamic process of capability formation in a way that raw correlations 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 is consistent with the high rank stability of cognition over ages past 10–12.

This finding helps to explain the evidence on ineffective cognitive remediation strategies for disadvantaged adolescents documented in Cunha, Heckman, Lochner, and Masterov (2006); Knudsen, Heckman, Cameron, and Shonkoff (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.

8.4.1 Evidence of Change in Traits from Other Studies of Parental Investment

Cunha, Heckman, Lochner, and Masterov (2006) summarize a large literature on child development. Evidence from a substantial literature suggests that for intelligence, the enduring effects of environment are greater earlier in life. Duyme, Dumaret, and Tomkiewicz (1999) studied children with IQs below 86, who were adopted between the ages of four and six into stable homes. As measured in their adolescent years, children adopted into high-SES homes gained an average of 19.5 IQ points; children

[228] Cunha and Heckman (2008) estimate a linear version of the technology. Their specification rules out interaction and assumes that, over the feasible range, investment can perfectly substitute for skill deficits.

adopted into low-SES homes showed an average gain of 7.7 IQ points. In a study 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 6 months and markedly less so when adoption was delayed until later ages (Beckett 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. 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).[229]

8.4.2 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 the effect of schooling on cognitive and personality traits. 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 (NLSY79), which has measures on the components of the Armed Services Vocational Battery (ASVAB) that are used to create the Armed Forces Qualifying Test (AFQT), a widely used 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 oneself (Rosenberg, 1965). The relationship between these measures and the Big Five traits of Neuroticism is discussed in Section 5.

Different traits might be more responsive to investment at different ages. Figure 1.29 shows the causal effects of years of schooling attained on five components of the Armed Forces Qualifying Test (AFQT). Schooling in the high-school years has moderate but positive effects on the measures of cognition, consistent with previous research by

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

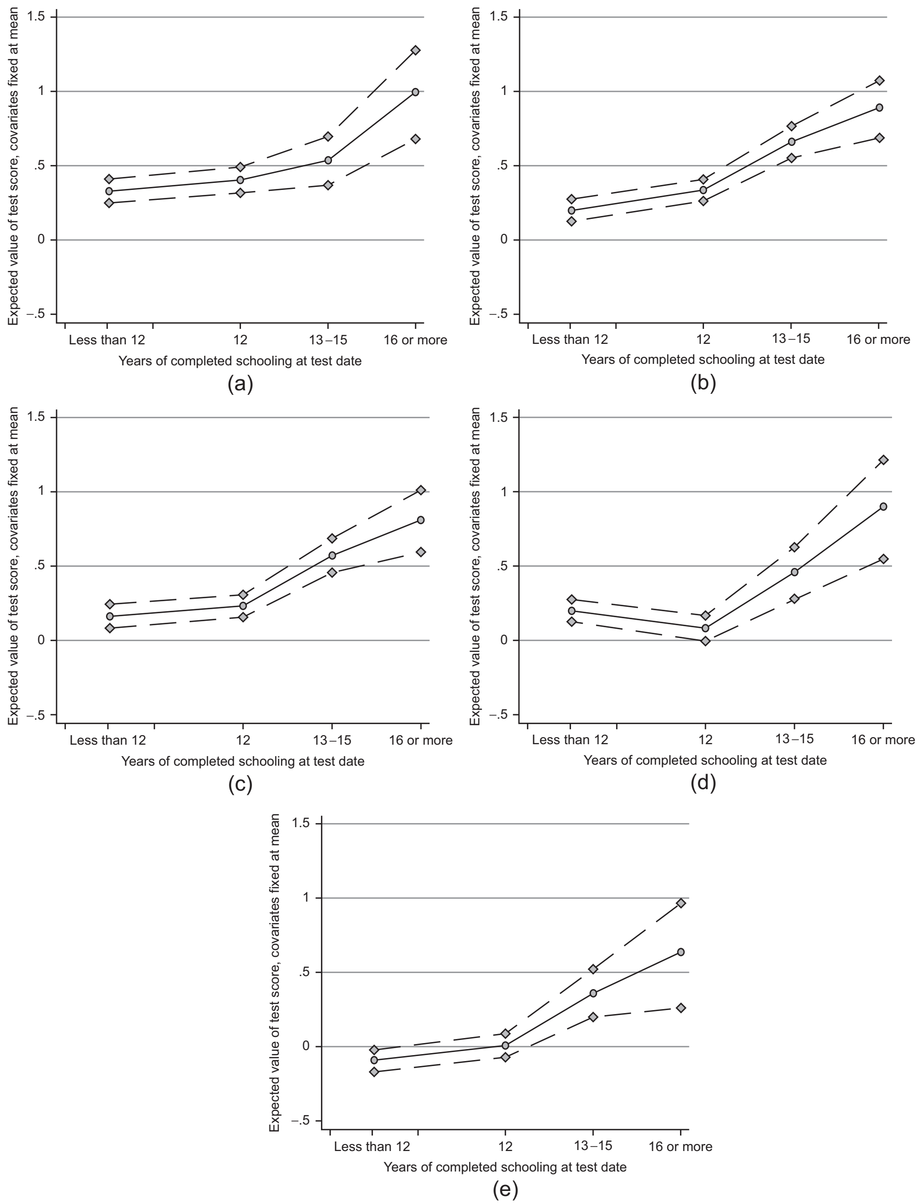

Figure 1.29 Causal Effect of Schooling on ASVAB Measures of Cognition. (a) Arithmetic Reasoning. (b) Word Knowledge. (c) Paragraph Comprehension. (d) Math Knowledge. (e) Coding Speed.

Notes: Effect of schooling on components of the ASVAB. The first four components are averaged to create males with average ability. We standardize the test scores to have within-sample mean zero and 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).

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. Figure 1.30 shows the causal effects of years of schooling attained on locus of control and self esteem. 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.[230]

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. Schooler and his colleagues show that complex (i.e., cognitively demanding) work increases intellectual functioning among adults and vice versa (Schooler, Mulatu, and Oates, 1999, and Kohn and Schooler, 1978).

8.4.3 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, and Savelyev, first draft 2008, revised 2011). 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 ended after 2 years of enrollment and both treatments and controls entered the same school. Participants were taught social skills in a "plan-do-review" sequence in which students planned a task, executed it, and then reviewed it with teachers and fellow students. They learned to work with others when problems arose.[231] 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. The program had a statistically significant rate of return of around 6–10% per annum for both boys and girls. These returns are above the post–World War II, pre-2008 meltdown, stock market returns to equity in US labor market that are estimated to be 5.8% per annum.[232] The Perry Preschool Program worked primarily through socioemotional channels. Figure 1.31 shows that the program improved scores on the California Achievement Test (CAT). However, the program did not have a lasting effect on IQ scores. This evidence is consistent with the discussions in Sections 5 and 7 that show that achievement test results are strongly dependent on personality traits (see Borghans, Duckworth, Heckman, and ter Weel, 2008, and Borghans, Golsteyn, Heckman, and Meijers, 2009). Indeed the personalities of participants improved. Participants had better direct measures of personal behavior (a weighted average of "absences

[230] See Hansen, Heckman, and Mullen (2004) for additional estimates of the causal effect of schooling on AFQT.
[231] Sylva (1997) describes the Perry Program as a Vygotskian program fostering personality traits.
[232] See DeLong and Magin (2009).

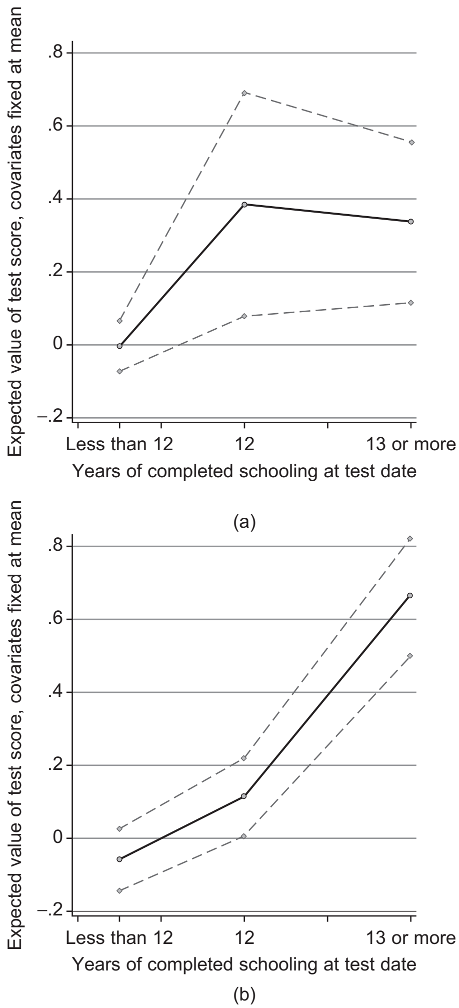

Figure 1.30 Causal Effect of Schooling on Two Measures of Personality. (a) Rotter Locus of Control Scale. (b) Rosenberg Self-Esteem Scale.

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).

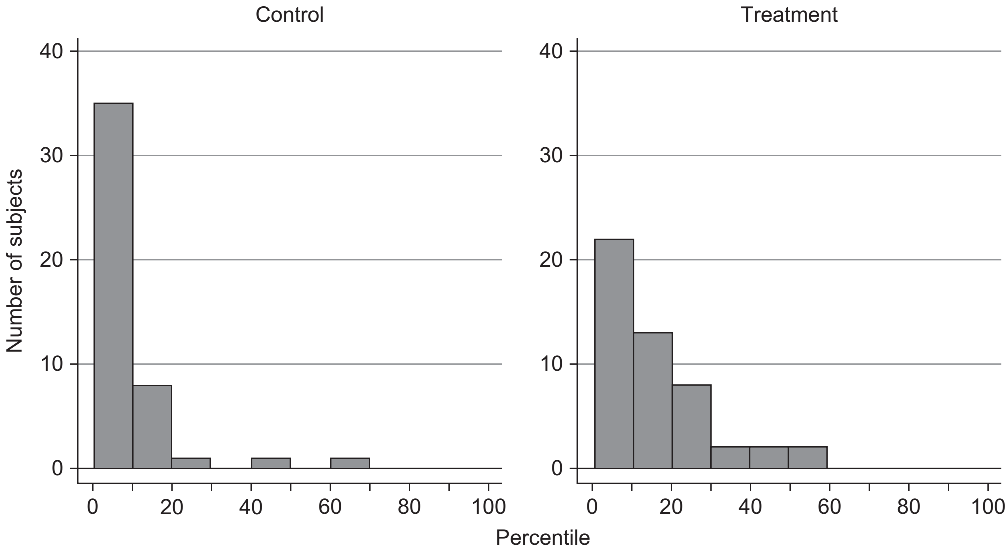

Figure 1.31 Perry Age 14 Total CAT Scores, by Treatment Group.
CAT, California Achievement Test. Treatment: $N = 49$; Control: $N = 46$. Statistically Significant Effect for Males and Females (p-values 0.009 and 0.021, respectively).

Source: Heckman, Malofeeva, Pinto, and Savelyev (first draft 2008, revised 2011).

and truancies," "lying and cheating," "stealing," and "swears or uses obscene words" measured by teachers in the elementary school years). Participants of both genders improved their externalizing behavior, which, as noted in Section 5, is related to Agreeableness and Conscientiousness. For girls, Openness to Experience (proxied by academic motivation) was also improved. Heckman, Malofeeva, Pinto, and Savelyev (first draft 2008, revised 2011) decompose the treatment effects of the Perry Program into components due to experientially induced changes of cognition (IQ), measures of personality, and residual factors. Improvements in personality led to beneficial changes in life outcomes.

Analyses of data from Project STAR, a program that randomly assigned kindergarteners and teachers to classes of different sizes, yields results similar 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 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 re-emerge later in life. People in better kindergarten classrooms had significantly higher earnings in early adulthood. Furthermore, kindergarten classroom quality also predicted better fourth- and eighth-grade behavior as measured by teacher-assessed effort, initiative, interest in the class, and disruptive behavior.[233] 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 traits, but a few programs did. 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 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 pro-social behavior, and improves teacher ratings of academic engagement.[234] PATHS is an exemplar of school-based social and emotional learning (SEL) programs, whose impact on both course grades ($d = 0.33$), where d is measured in units of standard deviations ("effect sizes") and standardized achievement test scores ($d = 0.27$), was recently documented in a meta-analysis of controlled studies involving over 270,000 children in kindergarten through college (Durlak, Weissberg, Dymnicki, Taylor, and Schellinger, 2011).[235] Similarly, a random assignment evaluation of *Tools of the Mind*, a preschool and early primary school curriculum, shows that in short-term follow-ups it improves classroom behavior as well as executive function, defined as higher level cognitive skills including inhibitory control, working memory, and cognitive flexibility (Barnett et al., 2008; Barnett, Yarosz, Thomas, and Hornbeck, 2006; Bodrova and Leong, 2001; Bodrova and Leong, 2007; Diamond, Barnett, Thomas, and Munro, 2007). Similar findings are reported for the Montessori preschool curriculum (Lillard and Else-Quest, 2006).

There is also evidence that targeted intervention efforts can improve aspects of Conscientiousness. These studies are typically more short term and, in contrast to the multifaceted curricula described above, are designed to isolate a particular mechanism for behavioral change. For instance, Rueda, Rothbart, McCandliss, Saccomanno, and Posner (2005) designed a set of computer exercises to train attention in children between 4 and 6 years of age. Children in the intervention group improved in performance on computer tasks of attention relative to children who instead watched interactive videos for a comparable amount of time. Similarly, Stevens, Fanning, Coch, Sanders, and Neville (2008) designed a 6-week computerized intervention and showed that it can improve selective auditory attention (i.e., the ability to attend to a target auditory signal in the face of an irrelevant, distracting auditory signal). Again, all of these programs have short-term follow-ups.

[233] These scales are based on more-detailed questionnaires. Only a subset of the sample has their behavioral measures.
[234] Bierman et al. (2010)
[235] 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).

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- to 15-year olds in Portugal by increasing motivation, self-esteem, and study skills. The program consists of one-on-one meetings with a trained staff member or meetings in small groups. The intervention was tailored to each participant's individual skill deficit. Overall, the program was successful and cost-effective, decreasing grade retention by 10% 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. Although the program does not require enrollment in the military, it stresses aspects of military discipline. The program features a 2-week assessment period, a 20-week residential program often conducted at a military base, and a 1-year mentoring program. Nine months after entry, participants in the program were 12% more likely to obtain a high-school diploma or GED, 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 is too short to know if the program has long-lasting effects. Although these studies show that adolescent personality can be improved through intervention, a couple of other studies show less-promising results (Rodríguez-Planas, 2010, and Holmlund and Silva, 2009).

Behncke (2009) provides some experimental evidence that short-term exogenous shocks to noncognitive skills affect test performance. She finds that giving words of encouragement before a diagnostic math test (an intervention that might boost short-term self-efficacy or self-esteem), was associated with 2.5% higher scores among all students ($p < 0.05$) and 8% higher scores among those with self-reported difficulties with math ($p < 0.01$). The result suggests that noncognitive skills can be shaped, at least 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 trial 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 experimental 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) report a decrease in Neuroticism for working women. Others show that marital and family experiences shape personality

(Helson and Picano, 1990, and Roberts, Helson, and Klohnen, 2002). However, these studies are all correlational in nature. None of these studies have the random assignment features of the Gottschalk study.

Personality may even be malleable at the end of life. Jackson, Hill, Payne, Roberts, and Stine-Morrow (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 laboratory 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 wait-listed control group. However, the elderly people were not followed after the program to determine whether the change was long lasting or whether important outcomes, like life expectancy, improved.

Table 1.12 summarizes the evidence on the effects of interventions that is discussed in this subsection. The evidence is consistent with effects of interventions, but there are woefully few causal studies with long-term follow-up.

8.4.4 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 (1980) 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 2 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.[236]

8.5. 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 mixed. Green, Fry, and Myerson (1994) and Harrison,

[236] Some evidence that further intervention can produce enduring change in nonclinical 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 2 years postintervention.

Table 1.12 The Effect of Interventions on Personality

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Barnett 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. *Posttreatment*—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% among all students ($p < 0.05$) and by 8% among students who reported difficulties with math ($p < 0.01$).

Continued

Table 1.12 The Effect of Interventions on Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Bierman 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 first grade. *Posttreatment*—Behavioral measures were taken again in the spring of third 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 with 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.

Continued

Table 1.12 The Effect of Interventions on Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Chetty et al. (2010)	Outcome(s): *noncognitive 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 fourth-grade students and 1,780 eighth-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 third grade.	Control Variables: wave fixed effects, student gender, free-lunch status, age, race, a quartic in the 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. *Posttreatment*—Age-relevant SAT tests and behavioral surveys were given in fourth and eighth grades. College quality and attendance was measured at age 19. Earnings were measured at age 27.	A 1-percentile improvement in kindergarten class quality increases an index of noncognitive skills by 0.15 percentiles in fourth grade ($p < 0.05$) and 0.13 percentiles in eighth grade ($p < 0.05$). Better classrooms were also associated with better life outcomes.

Continued

Table 1.12 The Effect of Interventions on Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Diamond, Barnett, Thomas, and Munro (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: *Posttreatment*—The tasks were given at the end of the second year of the program.	84% of students in Tools were successful in the Reverse Flanker task compared with 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 with the control group.
Durlak, Weissberg, Dymnicki, Taylor, and Schellinger (2011)	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.

Continued

Table 1.12 The Effect of Interventions on Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
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, and Savelyev (first draft 2008, revised 2011)	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.	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 improved Openness to Experience, as measured by academic motivation, for females ($p < 0.10$) but not for males. The program also generated a wide range of later-life outcomes primarily through improvements in noncognitive skills.

Continued

Table 1.12 The Effect of Interventions on Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
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 noncognitive 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. *Posttreatment*—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 1.2–1.4 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	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 seven SACD programs or left with their traditional	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	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

Continued

Table 1.12 The Effect of Interventions on Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
	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: seven different programs (ABC, CSP, LBW, PA, PATHS, 4Rs, SS) aimed to build Social and Character Development (SACD) compared to the "standard practice" programs at nontreated schools.	curriculum. The data were analyzed using hierarchical linear modeling (HLM).	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. *Posttreatment*—Data were collected near the end of the program in the spring of 2007.	($p < 0.10$), 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.

Continued

Table 1.12 The Effect of Interventions on Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Jackson, Hill, Payne, Roberts, and Stine-Morrow (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 60–94. 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 5 and 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 noncognitive skills including motivation, self-esteem, and study skills.	Data: EPIS database; 15,307 students of grade 7–9 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 grades 7 and 8. *During treatment*—Measures were taken each quarter that the students participate in the program through seven 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 ($p < 0.001$).

Continued

Table 1.12 The Effect of Interventions on Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
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 ninth 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 subpopulations: 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.

Continued

Table 1.12 The Effect of Interventions on Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Stevens, Fanning, Coch, Sanders, and Neville (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 score was 0.81 standard deviations higher for the participants than for the nonparticipants ($p < 0.01$). 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.

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, and Pauly (2007), de Wit, Flory, Acheson, McCloskey, and Manuck (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 1.13 below summarizes (but not exhaustively) the findings from a variety of recent economic studies on the heritability, malleability, and stability of preferences and personality.

8.6. Summary of Section 8

We have reviewed the evidence on change in personality over the life cycle. The evidence is strong that personality changes over the life cycle, both in terms of mean-level and rank-order change. The evidence on the source of the change is less clear-cut. Three competing visions of the source of change are discussed: (1) The ontogenic and sociogenic model that describes how biology and socialization produce changes in average traits. This approach does not explain why individuals develop with different trajectories; (2) The biological and pharmacological model that describes how alterations in the biology of the person can explain variations in personality and its evolution; and (3) The intervention/family influence model that describes how investment and environments generate changes. No study considers all three sources of development at the same time, largely due to data limitations. The evidence from the intervention and family influence studies suggests that interventions that target personality may be effective but much further evidence is required to specify the exact mechanisms through which the interventions work.

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 psychological constructs of personality into an economic framework? Can conventional models of preferences in economics characterize 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

Table 1.13 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 versus a certain payment. Explanatory Variable(s): *short-term gender environment*–whether the student was assigned to a co-ed or single-sex group during the experiment; *long-term gender environment*–whether the student attends a co-ed 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.	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 co-ed environments might exhibit more risk-taking in co-ed environments to try to impress girls.	Girls from co-ed high schools 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 co-ed experimental groups.

Continued

Table 1.13 The Heritability, Malleability, and Stability of Preferences and Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Burks, Carpenter, Goette, and Rustichini (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 personality traits, social closeness, social potency, and stress reaction.	Data: Collected by authors, administrative data from a human resources department; 1,063 trainee truckers from a US trucking company. Methods: probit, ordered probit, linear spline.	Controls: actual test performance, harm avoidance, education levels, ethnicity, sex, age, age squared, household income, before-test belief, and post-test belief 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 given the option to receive the results of their tests. Theory: People misjudge their own ability due to systematically biased noisy signals; value their self-assessed ability and avoid updating when the assessment is positive; or subconsciously 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 3.0 percentage point higher probability of demanding information about the IQ test ($p < 0.01$) and a 3.9 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 ($p < 0.05$).

Continued

Table 1.13 The Heritability, Malleability, and Stability of Preferences and Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Dohmen et al. (2011)	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 (GSOEP); 450 adults from Germany/22,019 people living in Germany Methods: interval regression, probit.	Controls: sex, age, height, parental education, 2002 household wealth, 2003 household income 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 ($p < 0.01$). Height is positively associated with a general willingness to take risks ($p < 0.01$). Mother and father's education is positively associated with willingness to take risks ($p < 0.01$). Stability of risk: The 6 measures of contextualized risk aversion are correlated with each other, ranging from 0.456–0.609.
Einav, Finkelstein, Pascu, and Cullen (2010)	Outcome(s): *risk preference*—order rankings of observed decisions to 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 next period.	Data: Administrative data; 12,752 employees of Alcoa, Inc. (2004) Methods: Spearman correlations, OLS.	Controls: the menu of benefits the employee faced, 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 0.164. Empirical results hold up under several robustness checks.

Continued

Table 1.13 The Heritability, Malleability, and Stability of Preferences and Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Kosfeld, Heinrichs, Zak, and Fehr (2005)	Outcome(s): *trust*—willingness to "invest" in a real-stakes two-player trust game by the investee (the first player); *risk preference*—real-stakes trust game played against a computer that randomly gives payoffs; *altruism*—the amount transferred back by the investee in the trust game. 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: RCT Timing of Measurements: The measures 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 (the second players in the trust game) do not show more altruistic behavior when given oxytocin.
Le, Miller, Slutske, and Martin (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: gender, age, education, and marital status. Timing of Measurements: The measures are contemporaneous.	Heritability of the risk measure is 0.192 ($p < 0.01$). Heritability of the conservative measure is 0.134 ($p < 0.01$).

Continued

Table 1.13 The Heritability, Malleability, and Stability of Preferences and Personality—continued

Author(s)	Main Variable(s)	Data and Methods	Causal Evidence	Main Result(s)
Sutter 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 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*—disutility if 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 8–17 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 and maturity.	20% of girls and boys behave selfishly. An increase in one year of age is associated with a 0.044 increase in the probability of having efficiency preferences for males ($p < 0.05$), but has no effect for females.

economic theory. The leading models of personality psychology are special cases of our model.[237]

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

   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 and how child temperament is related to adult traits. We link specific diagnoses of pathology with conventional measures of psychological traits.

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 the 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. Although 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 evidence on this question is not strong.

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 stable 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, evidence suggests that there are stable personality traits

[237] Freudian models of the unconscious would make the traits that govern behavior, and especially ψ, unknown to agents but nonetheless governing choices. A pure model of behaviorism would feature the effects of constraints on choices. Borghans, Duckworth, Heckman, and ter Weel (2008) develop such a model. We review it in the Web Appendix.

that predict a variety of behaviors in different situations. Personality is neither an ephemeral creation of situations nor is its manifestation invariant across situations. Moreover, 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 the life cycle is less ample. There is evidence that standard separable models of preferences are inadequate descriptions of choice behavior. There is little evidence on the stability or instability of preference parameters 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 have not yet attempted to establish the causal status of personality. However, there are a few experimental manipulations that establish the causal effect of personality. Recent studies in economics establish causal status of certain personality traits on outcomes for observational studies invoking assumptions that are inevitably subject to debate. Research in this area is likely to flourish in the coming years.

7. *Can personality be altered across the life cycle? Are interventions that change 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 very promising.

While personality psychology can enrich economics, the flow will likely be both ways. Economists have recently supplemented their traditional menu of preference parameters to account for a richer array of choices and actions. Personality psychologists have begun to use these new parameters. See Ferguson, Heckman, and Corr (2011). As personality psychologists shift their emphasis from the task of description to study issues related to causality, policy, and prediction, the Big Five may be replaced.

Also, economists and personality psychologists might eventually derive both the psychological traits and the economic preference parameters from a deeper set of motivation-oriented parameters. Conventional psychological traits and preference parameters may be manifestations of as yet undiscovered parameters rooted in human biology.

REFERENCES

Abbring, J.H., Heckman, J.J., 2007. Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation. In: Heckman, J., Leamer, E., (Ed.), Handbook of Econometrics. Elsevier, Amsterdam, 6B, 5145–5303.

Achenbach, T.M., McConaughy, S.H., Howell, C.T., 1987. Child/Adolescent Behavioral and Emotional Problems: Implications of Cross-Informant Correlations for Situational Specificity. Psychol. Bull. 101 (2), 213–232.

Ackerman, P.L., Heggestad, E.D., 1997. Intelligence, Personality, and Interests: Evidence for Overlapping Traits. Psychol. Bull. 121 (2), 219–245.

ACT, Incorporated, 2007. The ACT Technical Manual. ACT, Inc., Iowa City, IA.

Agnew, R., Brezina, T., Wright, J.P., Cullen, F.T., 2002. Strain, Personality Traits, and Delinquency: Extending General Strain Theory. Criminology 40 (1), 43–72.

Aigner, D.J., Hsiao, C., Kapteyn, A., Wansbeek, T., 1984. Latent Variable Models in Econometrics. In: Griliches, Z., Intriligator, M.D., (Ed.), Handbook of Econometrics. Elsevier, Amsterdam, 2, 1321–1393.

Akerlof, G.A., Dickens, W.T., 1982. The Economic Consequences of Cognitive Dissonance. Am. Econ. Rev. 72 (3), 307–319.

Alexander, K.L., Entwisle, D.R., Dauber, S.L., 1993. First-Grade Classroom Behavior: Its Short- and Long-Term Consequences for School Performance. Child Dev. 64 (3), 801–814.

Allport, G.W., Odbert, H.S., 1936. Traitnames. A Psycho-Lexical Study. Psychol. Monogr. 47, 171.

Altmann, S., Dohmen, T., Wibral, M., 2008. Do the Reciprocal Trust Less? Econ. Lett. 99 (3), 454–457.

American Council on Education, 2009. Technical Manual: 2002 Series Ged Tests. American Council on Education, GED Testing Service, Washington, DC.

American Psychological Association, 2007. APA Dictionary of Psychology. American Psychological Association, Washington, DC.

Anderson, J., Burks, S., DeYoung, C., Rustichini, A., 2011. Toward the Integration of Personality Theory and Decision Theory in the Explanation of Economic Behavior. Unpublished manuscript. Presented at the IZA Workshop: Cognitive and Non-Cognitive Skills, January 27, 2011. http://www.iza.org/conference_files/CoNoCoSk2011/rustichini_a3566.pdf.

Anderson, T.W., Rubin, H., 1956. Statistical Inference in Factor Analysis. In: Neyman, J., (Ed.), Proceedings of the Third Symposium on Mathematical Statistics and Probability, 5. University of California Press, Berkeley, pp. 111–150.

Andreoni, J., 1995. Warm-Glow Versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments. Q. J. Econ. 110 (1), 1–21.

Ardila, A., Pineda, D., Rosselli, M., 2000. Correlation between Intelligence Test Scores and Executive Function Measures. Arch. Clin. Neuropsychol. 15 (1), 31–36.

Aristotle, 1956. The Nicomachean Ethics. Harvard University Press, Cambridge, MA.

Ashton, M.C., Lee, K., Perugini, M., Szarota, P., de Vries, R.E., Di Blas, L., Boies, K., De Raad, B., 2004. A Six-Factor Structure of Personality-Descriptive Adjectives: Solutions from Psycholexical Studies in Seven Languages. J. Pers. Soc. Psychol. 86 (2), 356–366.

Ashton, M.C., Paunonen, S.V., Helmes, E., Jackson, D.N., 1998. Kin Altruism, Reciprocal Altruism, and the Big Five Personality Factors. Evol. Hum. Behav. 19 (1), 243–255.
Ayllon, T., Kelly, K., 1972. Effects of Reinforcement on Standardized Test Performance. J. Appl. Behav. Anal. 5 (4), 477–484.
Bagby, R.M., Bindseil, K.D., Schuller, D.R., Rector, N.A., Young, L.T., Cooke, R.G., Seeman, M.V., McCay, E.A., Joffe, R.T., 1997. Relationship between the Five-Factor Model of Personality and Unipolar, Bipolar, and Schizophrenic Patients. Psychiatry Res. 70 (2), 83–94.
Bandiera, O., Barankay, I., Rasul, I., 2005. Social Preferences and the Response to Incentives: Evidence from Personnel Data. Q. J. Econ. 120 (3), 917–962.
Bardsley, N., 2008. Dictator Game Giving: Altruism or Artefact? Exp. Econ. 11 (2), 122–133.
Barenbaum, N.B., Winter, D.G., 2008. History of Modern Personality Theory and Research. In: John, O.P., Robins, R.W., Pervin, L.A., (Ed.), Handbook of Personality: Theory and Research. Guilford Press, New York, pp. 3–26.
Barnett, W.S., Jung, K., Yarosz, D.J., Thomas, J., Hornbeck, A., Stechuk, R., Burns, S., 2008. Educational Effects of the Tools of the Mind Curriculum: A Randomized Trial. Early Child. Res. Q. 23 (3), 299–313.
Barnett, W.S., Yarosz, D.J., Thomas, J., Hornbeck, A., 2006. Educational Effectiveness of a Vygotskian Approach to Preschool Education: A Randomized Trial. National Institute for Early Education Research, Rutgers, NJ.
Báron, J.D., Cobb-Clark, D.A., 2010. Are Young People's Educational Outcomes Linked to Their Sense of Control? IZA Discussion Paper No. 4907.
Barrick, M.R., Mount, M.K., 1991. The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. Pers. Psychol. 44 (1), 1–26.
Barsky, R.B., Juster, F.T., Kimball, M.S., Shapiro, M.D., 1997. Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. Q. J. Econ. 112 (2), 537–579.
Baumeister, R.F., Bratslavsky, E., Muraven, M., Tice, D.M., 1998. Ego Depletion: Is the Active Self a Limited Resource? J. Pers. Soc. Psychol. 74 (5), 1252–1265.
Bechara, A., 2005. Decision Making, Impulse Control and Loss of Willpower to Resist Drugs: A Neurocognitive Perspective. Nat. Neurosci. 8 (11), 1458–1463.
Bechara, A., Damasio, A.R., 2005. The Somatic Marker Hypothesis: A Neural Theory of Economic Decision. Games Econ. Behav. 52 (2), 336–372.
Becker, G.S., 1964. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. National Bureau of Economic Research, distributed by Columbia University Press, New York.
Becker, G.S., Murphy, K.M., 1988. A Theory of Rational Addiction. J. Polit. Econ. 96 (4), 675.
Beckett, C., Maughan, B., Rutter, M., Castle, J., Colvert, E., Groothues, C., Kreppner, J., Stevens, S., O'Connor, T.G., Sonuga-Barke, E.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 Dev. 77 (3), 696–711.
Behncke, S., 2009. How Do Shocks to Non-Cognitive Skills Affect Test Scores? IZA Discussion Paper No. 4222.
Bekkers, R., 2006. Traditional and Health-Related Philanthropy: The Role of Resources and Personality. Soc. Psychol. Q. 69 (4), 349–366.
Ben-Ner, A., Kramer, A., 2011. Personality and Altruism in the Dictator Game: Relationship to Giving to Kin, Collaborators, Competitors, and Neutrals. Pers. Indiv. Differ. In press, Corrected Proof.
Benz, M., Meier, S., 2008. Do People Behave in Experiments as in the Field?—Evidence from Donations. Exp. Econ. 11 (3), 268–281.
Bergen, S.E., Gardner, C.O., Kendler, K.S., 2007. Age-Related Changes in Heritability of Behavioral Phenotypes over Adolescence and Young Adulthood: A Meta-Analysis. Twin Res. Hum. Genet. 10 (3), 423–433.
Bibby, P.A., Ferguson, E., 2011. The Ability to Process Emotional Information Predicts Loss Aversion. Pers. Indiv. Differ. In Press, Corrected Proof.
Bickel, W.K., Odum, A.L., Madden, G.J., 1999. Impulsivity and Cigarette Smoking: Delay Discounting in Current, Never, and Ex-Smokers. Psychopharmacology 146 (4), 447–454.

Bierman, K.L., Coie, J.D., Dodge, K.A., Greenberg, M.T., Lochman, J.E., McMahon, R.J., Pinderhughes, E., 2010. The Effects of a Multiyear Universal Social-Emotional Learning Program: The Role of Student and School Characteristics. J. Consult. Clin. Psychol. 78 (2), 156–168.
Binet, A., Simon, T., 1916. The Development of Intelligence in Children (the Binet-Simon Scale). Williams & Wilkins Co, Baltimore, MD.
Blair, C., 2006. How Similar Are Fluid Cognition and General Intelligence? A Developmental Neuroscience Perspective on Fluid Cognition as an Aspect of Human Cognitive Ability. Behav. Brain Sci. 29 (2), 109–160.
Blair, C., Gamson, D., Thorne, S., Baker, D., 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, C., Razza, R.P., 2007. Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten. Child Dev. 78 (2), 647–663.
Blais, A.R., Weber, E.U., 2006. A Domain-Specific Risk-Taking (DOSPERT) Scale for Adult Populations. Judgm. Decis. Mak. 1 (1), 33–47.
Block, J., 1995. A Contrarian View of the Five-Factor Approach to Personality Description. Psychol. Bull. 117 (2), 187–215.
Blonigen, D.M., Hicks, B.M., Krueger, R.F., Patrick, C.J., Iacono, W.G., 2006. Continuity and Change in Psychopathic Traits as Measured Via Normal-Range Personality: A Longitudinal-Biometric Study. J. Abnorm. Psychol. 115 (1), 85–95.
Bloom, D., Gardenhire-Crooks, A., Mandsager, C.L., 2009. Reengaging High School Dropouts: Early Results of the National Guard Youth ChalleNGe Program Evaluation. MDRC Report, New York.
Boake, C., 2002. From the Binet-Simon to the Wechsler-Bellevue: Tracing the History of Intelligence Testing. J. Clin. Exp. Neuropsychol. 24 (3), 383–405.
Bodrova, E., Leong, D.J., 2001. Tools of the Mind: A Case Study of Implementing the Vygotskian Approach in American Early Childhood and Primary Classrooms. International Bureau of Education, UNESCO, Geneva.
Bodrova, E., Leong, D.J., 2007. Tools of the Mind: The Vygotskian Approach to Early Childhood Education. Pearson Education, Inc, Upper Saddle River.
Bonato, D.P., Boland, F.J., 1983. Delay of Gratification in Obese Children. Addict. Behav. 8 (1), 71–74.
Booth, A.L., Nolen, P.J., 2009. Gender Differences in Risk Behaviour: Does Nurture Matter? IZA Discussion Paper No. 4026.
Borghans, L., Duckworth, A.L., Heckman, J.J., ter Weel, B., 2008. The Economics and Psychology of Personality Traits. J. Hum. Resour. 43 (4), 972–1059.
Borghans, L., Golsteyn, B.H.H., Heckman, J.J., and Humphries, J.E., 2011. Identification Problems in Personality Psychology, *Personality and Individual Differences* 51 (Special Issue on Personality and Economics), 315–320.
Borghans, L., Golsteyn, B.H.H., Heckman, J.J., Meijers, H., 2009. Gender Differences in Risk Aversion and Ambiguity Aversion. J. Eur. Econ. Assoc. 7 (2–3), 649–658.
Borghans, L., Meijers, H., ter Weel, B., 2008. The Role of Noncognitive Skills in Explaining Cognitive Test Scores. Econ. Inquiry 46 (1), 2–12.
Borkenau, P., Mauer, N., Riemann, R., Spinath, F.M., Angleitner, A., 2004. Thin Slices of Behavior as Cues of Personality and Intelligence. J. Personal. Soc. Psychol. 86 (4), 599–614.
Borsboom, D., Mellenbergh, G.J., van Heerden, J., 2003. The Theoretical Status of Latent Variables. Psychol. Rev. 110 (2), 203–219.
Bouchard, T.J., Loehlin, J.C., 2001. Genes, Evolution and Personality. Behav. Genet. 31 (3), 243–273.
Bowen, W.G., Chingos, M.M., McPherson, M.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 University Press, Princeton, NJ, pp. 1–19.
Bowen, W.G., Chingos, M.M., McPherson, M.S., 2009b. Test Scores and High School Grades as Predictors. In: Crossing the Finish Line: Completing College at America's Public Universities. Princeton University Press, Princeton, NJ, pp. 112–133.

Bowles, S., Gintis, H., 1976. Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life. Basic Books, New York.
Bowles, S., Gintis, H., Osborne, M., 2001a. The Determinants of Earnings: A Behavioral Approach. J. Econ. Lit. 39 (4), 1137–1176.
Bowles, S., Gintis, H., Osborne, M., 2001b. Incentive-Enhancing Preferences: Personality, Behavior, and Earnings. Am. Econ. Rev. 91 (2), 155–158.
Bowman, P.C., Matthews, C.V., 1960. Motivations of Youth for Leaving School. University of Chicago Development Project, Chicago IL.
Boyle, S.H., Williams, R.B., Mark, D.B., Brummett, B.H., Siegler, I.C., Barefoot, J.C., 2005. Hostility, Age, and Mortality in a Sample of Cardiac Patients. Am. J. Cardiol. 96 (1), 64–66.
Breuning, S.E., Zella, W.F., 1978. Effects of Individualized Incentives on Norm-Referenced IQ Test Performance of High School Students in Special Education Classes. J. Sch. Psychol. 16 (3), 220.
Bridgman, P.W., 1959. The Way Things Are. Harvard University Press, Cambridge.
Browning, M., Hansen, L.P., Heckman, J.J., 1999. Micro Data and General Equilibrium Models. In: Taylor, J.B., Woodford, M., (Ed.), Handbook of Macroeconomics. Elsevier, Amsterdam, 1A, 543–633.
Buck, R., 2003. Advanced Calculus. Waveland Press, Inc., Long Grove, IL.
Burks, S.V., Carpenter, J.P., Goette, L., Rustichini, A., 2009. Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment. Proc. Natl. Acad. Sci. 106 (19), 7745–7750.
Burks, S.V., Carpenter, J., Goette, L., Rustichini, A., 2010. Overconfidence Is a Social Signaling Bias. IZA Discussion Paper No. 4840.
Cahan, S., Cohen, N., 1989. Age Versus Schooling Effects on Intelligence Development. Child Dev. 60 (5), 1239–1249.
Caliendo, M., Cobb-Clark, D., Uhlendorff, A., 2010. Locus of Control and Job Search Strategies. IZA Discussion Paper No. 4750.
Caliendo, M., Fossen, F., Kritikos, A., 2010. The Impact of Risk Attitudes on Entrepreneurial Survival. Journal of Economic Behavior & Organization, 76 (1), 45–63.
Camerer, C.F., Fehr, E., 2004. Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists. In: Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., (Ed.), Foundations of Human Sociality: Experimental and Ethnographic Evidence from 15 Small-Scale Societies. Oxford University Press, New York, pp. 55–95.
Cameron, S.V., Heckman, J.J., 1998. Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males. J. Polit. Econ. 106 (2), 262–333.
Canli, T., 2006. Biology of Personality and Individual Differences. Guilford Press, New York.
Caplin, A., Leahy, J., 2001. Psychological Expected Utility Theory and Anticipatory Feelings. Q. J. Econ. 116 (1), 55–79.
Carneiro, P., Heckman, J.J., Masterov, D.V., 2005. Labor Market Discrimination and Racial Differences in Pre-Market Factors. J. Law Econ. 48 (1), 1–39.
Carpenter, P.A., Just, M.A., Shell, P., 1990. What One Intelligence Test Measures: A Theoretical Account of the Processing in the Raven Progressive Matrices Test. Psychol. Rev. 97 (3), 404–431.
Carpenter, J., Verhoogen, E., Burks, S., 2005. The Effect of Stakes in Distribution Experiments. Econ. Lett. 86 (3), 393–398.
Carroll, J.B., 1993. Human Cognitive Abilities: A Survey of Factor-Analytic Studies. Cambridge University Press, New York.
Casey, B.J., Tottenham, N., Liston, C., Durston, S., 2005. Imaging the Developing Brain: What Have We Learned About Cognitive Development? Trends Cogn. Sci. 9 (3), 104–110.
Caspi, A., 2000. The Child Is Father of the Man: Personality Continuities from Childhood to Adulthood. J. Personal. Soc. Psychol. 78 (1), 158–172.
Caspi, A., McClay, J., Moffitt, T.E., Mill, J., Martin, J., Craig, I.W., Taylor, A., Poulton, R., 2002. Role of Genotype in the Cycle of Violence in Maltreated Children. Science 297 (5582), 851–854.
Caspi, A., Moffit, T.E., Silva, P.A., Stouthamer-Loeber, M., Krueger, R.F., Schmutte, P.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, A., Shiner, R.L., 2006. Personality Development. In: Damon, W., Lerner, R.M., (Ed.), Handbook of Child Psychology. John Wiley & Sons, Inc., Hoboken, NJ, pp. 300–365.

Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A., Poulton, R., 2003. Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene. Science 301 (5631), 386–389.

Cattan, S., 2011. Heterogeneity and Selection in the Labor Market. PhD Thesis. Economics Department, University of Chicago.

Cattell, J.M., 1890. Mental Tests and Measurements. Mind 15 (59), 373.

Cattell, R.B., 1971. Abilities: Their Structure, Growth, and Action. Houghton Mifflin, Boston.

Cattell, R.B., 1987. Intelligence: Its Structure, Growth and Action. Elsevier, Amsterdam.

Cebi, M., 2007. Locus of Control and Human Capital Investment Revisited. J. Hum. Resour. 42 (4), 919–932.

Cervone, D., 2004. The Architecture of Personality. Psychol. Rev. 111 (1), 183–204.

Cervone, D., 2005. Personality Architecture: Within-Person Structures and Processes. Ann. Rev. Psychol. 56, 423–452.

Cervone, D., Pervin, L.A., 2009. Personality: Theory and Research. John Wiley and Sons, Inc., Hoboken, NJ.

Chamorro-Premuzic, T., Furnham, A., 2005. Personality and Intellectual Competence. L. Erlbaum Associates, Mahwah, NJ.

Chao, L.-W., Szrek, H., Sousa Pereira, N., Pauly, M.V., 2007. Time Preference and Its Relationship with Age, Health, and Longevity Expectations. Universidade do Porto Faculdade de Economia do Porto 0706.

Chapman, G.B., Coups, E.J., 1999. Time Preferences and Preventive Health Behavior: Acceptance of the Influenza Vaccine. Med. Decis. Mak. 19 (3), 307.

Chapman, G.B., Nelson, R., Hier, D.B., 1999. Familiarity and Time Preferences: Decision Making About Treatments for Migraine Headaches and Crohn's Disease. J. Exp. Psychol. Appl. 5 (1), 17.

Chesson, H., Viscusi, W.K., 2000. The Heterogeneity of Time-Risk Tradeoffs. J. Behav. Decis. Mak. 13 (2), 251–258.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Diane, S.W., Yagan, D., 2010. How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR. NBER Working Paper No. 16381.

Clausen, J.A., Gilens, M., 1990. Personality and Labor Force Participation across the Life Course: A Longitudinal Study of Women's Careers. Sociol. Forum 5 (4), 595–618.

Clingman, J., Fowler, R.L., 1976. The Effects of Primary Reward on the I.Q. Performance of Grade-School Children as a Function of Initial I.Q. Level. J. Appl. Behav. Anal. 9 (1), 19–23.

Cloninger, C.R., 1987. A Systematic Method for Clinical Description and Classification of Personality Variants: A Proposal. Arch. Gen. Psychiatry 44 (6), 573–588.

Cloninger, C.R., Svrakic, D.M., Bayon, C., Przybeck, T.R., 1999. Measurement of Psychopathology as Variants of Personality. In: Cloninger, C.R., (Ed.), Personality and Psychopathology. American Psychiatric Press, Washington, DC.

Cobb-Clark, D., Tan, M., 2009. Noncognitive Skills, Occupational Attainment, and Relative Wages. Forthcoming, Labour Econ.

Coleman, M., DeLeire, T., 2003. An Economic Model of Locus of Control and the Human Capital Investment Decision. J. Hum. Resour. 38 (3), 701–721.

Coller, M., Williams, M.B., 1999. Eliciting Individual Discount Rates. Exp. Econ. 2 (2), 107–127.

Conard, M.A., 2005. Aptitude Is Not Enough: How Personality and Behavior Predict Academic Performance. J. Res. Pers. 40 (2006), 339–346.

Conti, G., Heckman, J.J., 2010. Understanding the Early Origins of the Education-Health Gradient: A Framework That Can Also Be Applied to Analyze Gene-Environment Interactions. Perspect. Psychol. Sci. 5 (5), 585–605.

Conti, G., Heckman, J.J., Lopes, H.F., Piatek, R., 2010. Constructing Economically Justified Aggregates: An Application of the Early Origins of Health. Unpublished manuscript. Department of Economics, University of Chicago.

Conti, G., Heckman, J.J., Urzua, S., 2010a. Early Endowments, Education, and Health. Unpublished manuscript, University of Chicago, Department of Economics.

Conti, G., Heckman, J.J., Urzua, S., 2010b. The Education-Health Gradient. American Economic Review. 100 (2), 1–5.
Conti, G., Pudney, S., 2007. The Dynamics of Cognitive Development, Unpublished manuscript, Department of Economics, University of Essex.
Costa, P.T., McCrae, R.R., 1992a. Four Ways Five Factors Are Basic. Pers. Individ. Dif. 13 (6), 653–665.
Costa, P.T., McCrae, R.R., 1992b. Revised NEO Personality Inventory (NEO PI-R), NEO Five-Factor Inventory (NEO-FFI). Psychological Assessment Resources, Odessa, FL.
Costa, P.T., McCrae, R.R., 1994. Set Like Plaster? Evidence for the Stability of Adult Personality. In: Heatherton, T.F., Weinberger, J.L., (Ed.), Can Personality Change? American Psychological Association, Washington, DC, pp. 21–40.
Costa, P.T., McCrae, R.R., Siegler, I.C., 1999. Continuity and Change over the Adult Life Cycle: Personality and Personality Disorders. In: Cloninger, C.R., (Ed.), Personality and Psychopathology. American Psychiatric Press, Washington, DC.
Cragg, J.G., Donald, S.G., 1997. Inferring the Rank of a Matrix. J. Econom. 76 (1–2), 223–250.
Credé, M., Kuncel, N.R., 2008. Study Habits, Skills, and Attitudes: The Third Pillar Supporting Collegiate Academic Performance. Perspect. Psychol. Sci. 3 (6), 425–453.
Csikszentmihalyi, M., 1996. Creativity: Flow and the Psychology of Discovery and Invention. Harper Collins Publishers, New York.
Cudeck, R., MacCullum, R.C., 2007. Factor Analysis at 100: Historical Developments and Future Directions. Lawrence Erlibaum Associates, Mahwah, NJ.
Cunha, F., Heckman, J.J., 2007. The Technology of Skill Formation. Am. Econ. Rev. 97 (2), 31–47.
Cunha, F., Heckman, J.J., 2008. Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. J. Hum. Resour. 43 (4), 738–782.
Cunha, F., Heckman, J.J., 2009. The Economics and Psychology of Inequality and Human Development. J. Eur. Econ. Assoc. 7 (2–3), 320–364.
Cunha, F., Heckman, J.J., Lochner, L.J., Masterov, D.V., 2006. Interpreting the Evidence on Life Cycle Skill Formation. In: Hanushek, E.A., Welch, F., (Ed.), Handbook of the Economics of Education. North-Holland, Amsterdam, 1, 697–812.
Cunha, F., Heckman, J.J., Navarro, S., 2005. Separating Uncertainty from Heterogeneity in Life Cycle Earnings, the 2004 Hicks Lecture. Oxf. Econ. Pap. 57 (2), 191–261.
Cunha, F., Heckman, J.J., Schennach, S.M., 2010. Estimating the Technology of Cognitive and Noncognitive Skill Formation. Econometrica 78 (3), 883–931.
Daly, M., Delaney, L., Harmon, C.P., 2009. Psychological and Biological Foundations of Time Preferences. J. Eur. Econ. Assoc. 7 (2–3), 659–669.
Damasio, H., Grabowski, T., Frank, R., Galaburda, A.M., Damasio, A.R., 2005. The Return of Phineas Gage: Clues About the Brain from the Skull of a Famous Patient. In: Cacioppo, J.T., Berntson, G.G., (Ed.), Social Neuroscience: Key Readings. Key Reading in Social Psychology. Psychology Press, New York, pp. 21–28.
de Wit, H., Flory, J.D., Acheson, A., McCloskey, M., Manuck, S.B., 2007. IQ and Nonplanning Impulsivity Are Independently Associated with Delay Discounting in Middle-Aged Adults. Pers. Individ. Dif. 42 (1), 111–121.
Dee, T., West, M., 2008. The Non-Cognitive Returns to Class Size. NBER Working Paper No. 13994.
DeLong, J.B., Magin, K., 2009. The U.S. Equity Return Premium: Past, Present and Future. J. Econ. Perspect. 23 (1), 193–208.
Devlin, B., Daniels, M., Roeder, K., 1997. The Heritability of IQ. Nature 388 (6641), 468–471.
DeYoung, C.G., Hirsh, J.B., Shane, M.S., Papademetris, X., Rajeevan, N., Gray, J.R., 2010. Testing Predictions from Personality Neuroscience: Brain Structure and the Big Five. Psychol. Sci. 21 (6), 820–828.
DeYoung, C.G., Quilty, L.C., Peterson, J.B., 2007. Between Facets and Domains: Ten Aspects of the Big Five. J. Pers. Soc. Psychol. 93 (5), 880–896.
Diamond, A., Barnett, S., Thomas, J., Munro, S., 2007. Preschool Program Improves Cognitive Control. Science 318 (5855), 1387–1388.
Dickens, W.T., Flynn, J.R., 2001. Heritability Estimates Versus Large Environmental Effects: The IQ Paradox Resolved. Psychol. Rev. 108 (2), 346–369.

Digman, J.M., 1997. Higher-Order Factors of the Big Five. J. Pers. Soc. Psychol. 73 (6), 1246–1256.
Ding, X., Hartog, J., Sun, Y., 2010. Can We Measure Individual Risk Attitudes in a Survey? IZA Discussion Paper, 4807.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2008. Representative Trust and Reciprocity: Prevalence and Determinants. Econ. Inquiry 46 (1), 84–90.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2009. Homo Reciprocans: Survey Evidence on Behavioural Outcomes. Econ. J. 119 (536), 592–612.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are Risk Aversion and Impatience Related to Cognitive Ability? Am. Econ. Rev. 100 (3), 1238–1260.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences. J. Eur. Econ. Assoc. 9 (3), 522–550.
Drago, F., 2008. Self-Esteem and Earnings. IZA Discussion Paper, 3577.
Du, W., Green, L., Myerson, J., 2002. Cross-Cultural Comparisons of Discounting Delayed and Probabilistic Rewards. Psychol. Rec. 52 (4), 479.
Duckworth, A.L., 2007. Unpublished Dataset. Department of Psychology, University of Pennsylvania.
Duckworth, A.L., 2009. (Over and) Beyond High-Stakes Testing. Am. Psychol. 64 (4), 279–280.
Duckworth, A.L., 2011. Unpublished. Dataset on Personality as a Predictor of Grade Point Average and Achievement Test Scores. University of Pennsylvania, Department of Psychology.
Duckworth, A.L., Quinn, P.D., Tsukayama, E., 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. Department of Psychology, University of Pennsylvania.
Duckworth, A.L., Schulze, R. 2009. Jingle Jangle: A Meta-Analysis of Convergent Validity Evidence for Self-Control Measures. Manuscript. Department of Psychology, University of Pennsylvania.
Duckworth, A.L., Seligman, M.E.P., 2005. Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents. Psychol. Sci. 16 (12), 939–944.
Duckworth, A.L., Seligman, M.E.P., 2006. Self-Discipline Gives Girls the Edge: Gender in Self-Discipline, Grades, and Achievement Scores. J. Educ. Psychol. 98 (1), 198–208.
Duckworth, A.L., Tsukayama, E., May, H., 2010. Establishing Causality Using Longitudinal Hierarchical Linear Modeling: An Illustration Predicting Achievement from Self-Control. Soc. Psychol. Personal. Sci. 1 (4), 311–317.
Duncan, G.J., Dowsett, C.J., Claessens, A., Mugnuson, K., Huston, A.C., Klebanov, P., Pagani, L.S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., Japel, C., 2007. School Readiness and Later Achievement. Dev. Psychol. 43 (6), 1428–1446.
Duncan, G.J., Dunifon, R., 1998. Soft-Skills and Long-Run Labor Market Success. In: Polachek, S.W., (Ed.), Research in Labor Economics. JAI Press, Stamford, CT, 17, 123–149.
Duncan, G.J., Magnuson, K., 2010. The Nature and Impact of Early Achievement Skills, Attention Skills, and Behavior Problems. Working paper 2010 at the Department of Education, UC Irvine.
Dur, R., Non, A., Roelfsema, H., 2010. Reciprocity and Incentive Pay in the Workplace. J. Econ. Psychol. 31 (4), 676–686.
Durlak, J.A., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D., Schellinger, K.B., 2011. The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. Child Dev 82 (1), 405–432.
Duyme, M., Dumaret, A.-C., Tomkiewicz, S., 1999. How Can We Boost IQs Of "Dull Children?: A Late Adoption Study. Proc. Natl. Acad. Sci. 96 (15), 8790–8794.
Eckel, C.C., Grossman, P.J., 2002. Sex Differences and Statistical Stereotyping in Attitudes toward Financial Risk. Evol. Hum. Behav. 23 (4), 281–295.
Edlund, C.V., 1972. The Effect on the Behavior of Children, as Reflected in the IQ Scores, When Reinforced after Each Correct Response. J. Appl. Behav. Anal. 5 (3), 317–319.
Einav, L., Finkelstein, A., Pascu, I., Cullen, M.R., 2010. How General Are Risk Preferences? Choices under Uncertainty in Different Domains. NBER Working Paper No. 15686.
Endler, N., Magnusson, D., (Ed.), 1976. Interactional Psychology and Personality. Hemisphere Publishing Corp, New York, NY.

Epstein, S., 1979. The Stability of Behavior: I. On Predicting Most of the People Much of the Time. J. Personal. Soc. Psychol. 37 (7), 1097–1126.
Epstein, L.G., Schneider, M., 2003. Recursive Multiple-Priors. J. Econ. Theory 113 (1), 1–31.
Epstein, L.G., Zin, S.E., 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. Econometrica 57 (4), 937–969.
Estle, S.J., Green, L., Myerson, J., Holt, D.D., 2007. Discounting of Monetary and Directly Consumable Rewards. Psychol. Sci. 18 (1), 58.
Eysenck, H.J., 1991. Dimensions of Personality: 16, 5 or 3? Criteria for a Taxonomic Paradigm. Pers. Individ. Dif. 12 (8), 773–790.
Falk, A., Fischbacher, U., 2006. A Theory of Reciprocity. Games Econ. Behav. 54 (2), 293–315.
Falk, A., Heckman, J.J., 2009. Lab Experiments Are a Major Source of Knowledge in the Social Sciences. Science 326 (5952), 535–538.
Fehr, E., 2009. On the Economics and Biology of Trust. J. Eur. Econ. Assoc. 7 (2–3), 235–266.
Fehr, E., Gächter, S., 2000. Fairness and Retaliation: The Economics of Reciprocity. J. Econ. Perspect. 14 (3), 159–181.
Fehr, E., Schmidt, K.M., 1999. A Theory of Fairness, Competition, and Cooperation. Q. J. Econ. 114 (3), 817–868.
Fehr, E., Schmidt, K.M., 2006. The Economics of Fairness, Reciprocity and Altruism-Experimental Evidence and New Theories. In: Kolm, S.-C., Ythier, J.M., (Ed.), Handbook of the Economics of Giving, Altruism and Reciprocity. Elsevier, Amsterdam, 1, 615–691.
Feingold, A., 1982. The Validity of the Information and Vocabulary Subtests of the WAIS. J. Clin. Psychol. 38 (1), 169–174.
Ferguson, E., Heckman, J.J., Corr, P., 2011. Editorial: Personality and Economics: Overview and Proposed Framework. Personality and Individual Differences, 51 (Special Issue on Personality and Economics), 201–209.
Fergusson, D.M., Horwood, L.J., 1998. Early Conduct Problems and Later Life Opportunities. J. Child Psychol. Psychiatry 39 (8), 1097–1108.
Figner, B., Knoch, D., Johnson, E.J., Krosch, A.R., Lisanby, S.H., Fehr, E., Weber, E.U. 2010. Lateral Prefrontal Cortex and Self-Control in Intertemporal Choice. Nat. Neurosci. 13 (5), 538–539.
Finn, J.D., Pannozzo, G.M., Voelkl, K.E., 1995. Disruptive and Inattentive-Withdrawn Behavior and Achievement among Fourth Graders. Elem. Sch. J. 95 (5), 421–434.
Fleeson, W., 2001. Toward a Structure- and Process-Integrated View of Personality: Traits as Density Distributions of States. J. Personal. Soc. Psychol. 80 (6), 1011–1027.
Fleeson, W., Noftle, E., 2008. The End of the Person–Situation Debate: An Emerging Synthesis in the Answer to the Consistency Question. Soc. Personal. Psychol. Compass 2 (4), 1667–1684.
Flynn, J.R., 2007. What Is Intelligence? Beyond the Flynn Effect. Cambridge University Press, New York.
Foster, J.E., Sen, A.K., 1997. On Economic Inequality. Oxford University Press, New York.
Frederick, S., Loewenstein, G., O'Donoghue, T., 2002. Time Discounting and Time Preference: A Critical Review. J. Econ. Lit. 40 (2), 351–401.
Freud, S., 1909. Five Lectures on Psycho-Analysis (the Standard Edition). W. W. Norton & Company, New York.
Frey, M.C., Detterman, D.K., 2004. Scholastic Assessment or g? The Relationship between the Scholastic Assessment Test and General Cognitive Ability. Psychol. Sci. 15 (6), 373–378.
Friedman, H.S., 2007. Personality, Disease, and Self-Healing. In: Friedman, H.S., Silvers, R.C., (Ed.), Foundations of Health Psychology. Oxford University Press, New York, pp. 172–199.
Friedman, H.S., Kern, M.L., Reynolds, C.A., 2010. Personality and Health, Subjective Well-Being, and Longevity. J. Personal. 78 (1), 179–216.
Friedman, N.P., Miyake, A., Corley, R.P., Young, S.E., DeFries, J.C., Hewitt, J.K., 2006. Not All Executive Functions Are Related to Intelligence. Psychol. Sci. 17 (2), 172–179.
Funder, D.C., 2008. Persons, Situations, and Person-Situation Interactions. In: John, O.P., Robins, R.W., Pervin, L.A., (Ed.), Handbook of Personality: Theory and Research. Guilford Press, New York, pp. 568–580.

Funder, D.C., 2009. Persons, Behaviors and Situations: An Agenda for Personality Psychology in the Post-war Era. J. Res. Personal. 43 (2), 120–126.
Funder, D.C., Block, J., 1989. The Role of Ego-Control, Ego-Resiliency, and IQ in Delay of Gratification in Adolescence. J. Personal. Soc. Psychol. 57 (6), 1041–1050.
Gale, C.R., Batty, G.D., Deary, I.J., 2008. Locus of Control at Age 10 Years and Health Outcomes and Behaviors at Age 30 Years: The 1970 British Cohort Study. Psychosom. Med. 70 (4), 397–403.
Gallo, W.T., Endrass, J., Bradley, E.H., Hell, D., Kasl, S.V., 2003. The Influence of Internal Control on the Employment Status of German Workers. Schmollers Jahrbuch 123 (1), 71–81.
Galton, F., 1883. Inquiries into Human Faculty and Its Development. Macmillan and Company, London.
Gardner, H., 2004. Frames of Mind: The Theory of Multiple Intelligences. Basic Books, New York, NY.
Garon, N., Bryson, S.E., Smith, I.M., 2008. Executive Function in Preschoolers: A Review Using an Integrative Framework. Psychol. Bull. 134 (1), 31–60.
Geiser, S., Santelices, M.V., 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., Schmeidler, D., 1989. Maxmin Expected Utility with Non-Unique Prior. J. Math. Econ. 18 (2), 141–153.
Gillham, J.E., Reivich, K.J., 1999. Prevention of Depressive Symptoms in School Children: A Research Update. Psychol. Sci. 10 (5), 461–462.
Gneezy, U., List, J.A., 2006. Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments. Econometrica 74 (5), 1365–1384.
Goldberg, L.R., 1993. The Structure of Phenotypic Personality Traits. Am. Psychol. 48 (1), 26–34.
Goldberg, L.R., Sweeney, D., Merenda, P.F., Hughes, Jr. J.E., 1998. Demographic Variables and Personality: The Effects of Gender, Age, Education, and Ethnic/Racial Status on Self-Descriptions of Personality Attributes. Pers. Individ. Dif. 24 (3), 393–403.
Goldberger, A.S., 1968. Topics in Regression Analysis. Macmillan, New York.
Goldsmith, H.H., Buss, A.H., Plomin, R., Rothbart, M.K., Thomas, A., Chess, S., Hinde, R.A., McCall, R.B., 1987. Roundtable: What Is Temperament? Four Approaches. Child Dev. 58 (2), 505–529.
Gorsuch, R., 1983. Factor Analysis. Lawrence Erlbaum Associates, Hillsdale, NJ.
Gottfredson, L.S., 1998. The General Intelligence Factor. Sci. Am. Presents 9 (4), 24–29.
Gottfredson, L.S., 2002. g: Highly General and Highly Practical. In: Sternberg, R.J., Grigorenko, E.L., (Ed.), The General Factor of Intelligence: How General Is It? Lawrence Erlbaum Associates Publishers, Mahwah, NJ, pp. 331–380.
Gottfredson, M.R., Hirschi, T., 1990. A General Theory of Crime. Stanford University Press, Palo Alto, CA.
Gottschalk, P., 2005. Can Work Alter Welfare Recipients' Beliefs? J. Policy Anal. Manage 24 (3), 485–498.
Gough, H.G., 1964. Graduation from High School as Predicted from the California Psychological Inventory. Psychol. Sch. 3 (3), 208–216.
Gough, H.G., Heilbrun, A.B., 1983. The Adjective Check List Manual. Consulting Psychologists Press, Palo Alto, CA.
Gray, J.R., 2004. Integration of Emotion and Cognitive Control. Curr. Dir. Psychol. Sci. 13 (2), 46–48.
Green, L., Fry, A.F., Myerson, J., 1994. Discounting of Delayed Rewards: A Life-Span Comparison. Psychol. Sci. 5 (1), 33–36.
Green, L., Myerson, J., 2004. A Discounting Framework for Choice with Delayed and Probabilistic Rewards. Psychol. Bull. 130 (5), 769.
Greene, A.C., Sapp, G.L., Chissom, B., 1990. Validation of the Stanford-Binet Intelligence Scale: Fourth Edition with Exceptional Black Male Students. Psychol. Sch. 27 (1), 35–41.
Griliches, Z., 1977. Estimating the Returns to Schooling: Some Econometric Problems. Econometrica 45 (1), 1–22.
Groth-Marnat, G., 2009. Handbook of Psychological Assessment. John Wiley and Sons, Hoboken, NJ.
Gul, F., Pesendorfer, W., 2004. Self-Control and the Theory of Consumption. Econometrica 72 (1), 119–158.

Gunderson, J.G., Triebwasser, J., Phillips, K.A., Sullivan, C.N., 1999. Personality and Vulnerability to Affective Disorders. In: Cloninger, C.R., (Ed.), Personality and Psychopathology. American Psychiatric Press, Washington, DC.

Gustafsson, J.-E., 1988. Hierarchical Models of Individual Differences in Cognitive Abilities. In: Sternberg, R.J., (Ed.), Advances in the Psychology of Human Intelligence. Lawrence Erlbaum Associates, Hillsdale, NJ, p. 4.

Ham, R., Junankar, P.N., Wells, R., 2009. Occupational Choice: Personality Matters. IZA Discussion Paper No. 4105.

Hampson, S.E., Friedman, H.S., 2008. Personality and Health: A Lifespan Perspective. In: John, O.P., Robins, R., Pervin, L., (Ed.), The Handbook of Personality: Theory and Research. Guilford, New York, pp. 770–794.

Hampson, S.E., Goldberg, L.R., Vogt, T.M., Dubanoski, J.P., 2007. Mechanisms by Which Childhood Personality Traits Influence Adult Health Status: Educational Attainment and Healthy Behaviors. Health Psychol. 26 (1), 121–125.

Hampson, S.E., Tildesley, E., Andrews, J.A., Luyckx, K., Mroczek, D.K., 2010. The Relation of Change in Hostility and Sociability During Childhood to Substance Use in Mid Adolescence. J. Res. Personal. 44 (1), 103–114.

Hanoch, Y., Johnson, J.G., Wilke, A., 2006. Domain Specificity in Experimental Measures and Participant Recruitment: An Application to Risk-Taking Behavior. Psychol. Sci. 17 (4), 300.

Hansen, L.P., 2005. Comment On "Exotic Preferences for Macroeconomics." By D.K Backus, B.R. Routledge, and, S.E. Zin, In: Gertler, M. Rogoff, K., (Ed.), NBER Macroeconomics Annual. MIT Press, Cambridge, MA, pp. 391–405.

Hansen, K.T., Heckman, J.J., Mullen, K.J., 2004. The Effect of Schooling and Ability on Achievement Test Scores. J. Econom. 121 (1–2), 39–98.

Hansen, L.P., Sargent, T.J., 2008. Robustness. Princeton University Press, Princeton, NJ.

Harman, H.H., 1976. Modern Factor Analysis, third ed. University of Chicago Press, Chicago, IL.

Harrison, G.W., Lau, M.I., Williams, M.B., 2002. Estimating Individual Discount Rates in Denmark: A Field Experiment. Am. Econ. Rev. 92 (5), 1606–1617.

Hartigan, J.A., Wigdor, A.K., 1989. Fairness in Employment Testing: Validity Generalization, Minority Issues, and the General Aptitude Test Battery. National Academy Press, Washington, DC, p. 354.

Hartlage, L.C., Steele, C.T., 1977. WISC and WISC-R Correlates of Academic Achievement. Psychol. Sch. 14 (1), 15–18.

Hartog, J., 1980. Earnings and Capability Requirements. Rev. Econ. Stat. 62 (2), 230–240.

Hartog, J., 2001. On Human Capital and Individual Capabilities. Rev. Income Wealth 47 (4), 515–540.

Hathaway, S.R., Reynolds, P.C., Monachesi, E.D., 1969. Follow-up of the Later Careers and Lives of 1,000 Boys Who Dropped out of High School. J. Consult. Clin. Psychol. 33 (3), 370–380.

Heckman, J.J., 1976. A Life-Cycle Model of Earnings, Learning, and Consumption. J. Polit. Econ. 84 (4, Part 2), S11–S44.

Heckman, J.J., 2007. The Economics, Technology and Neuroscience of Human Capability Formation. Proc. Natl. Acad. Sci. 104 (3), 13250–13255.

Heckman, J.J., 2008a. Econometric Causality. Int. Stat. Rev. 76 (1), 1–27.

Heckman, J.J., 2008b. Schools, Skills and Synapses. Econom. Inquiry 46 (3), 289–324.

Heckman, J.J., Honoré, B.E., 1990. The Empirical Content of the Roy Model. Econometrica 58 (5), 1121–1149.

Heckman, J.J., Humphries, J.E., Mader, N., 2011. The GED. In: Hanushek, E.A., Machin, S., Wößmann, L., (Ed.), Handbook of the Economics of Education. Elsevier, Amsterdam, pp. 423–484.

Heckman, J.J., Humphries, J.E., Urzua, S., Veramendi, G., 2011. The Effects of Schooling on Labor Market and Health Outcomes. Unpublished Manuscript. Department of Economics, University of Chicago.

Heckman, J.J., LaFontaine, P.A., 2010. The American High School Graduation Rate: Trends and Levels. Rev. Econom. Stat. 92 (2), 244–262.

Heckman, J.J., Malofeeva, L., Pinto, R., Savelyev, P.A., 2011. Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes. Unpublished manuscript. Department of Economics, University of Chicago, First draft, 2008.
Heckman, J.J., Masterov, D.V., 2007. The Productivity Argument for Investing in Young Children. Rev. Agric. Econ. 29 (3), 446–493.
Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A., Yavitz, A.Q., 2010a. Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the Highscope Perry Preschool Program. Quant. Econom. 1 (1), 1–46.
Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A., Yavitz, A.Q., 2010b. The Rate of Return to the Highscope Perry Preschool Program. J. Public Econom. 94 (1–2), 114–128.
Heckman, J.J., Navarro, S., 2007. Dynamic Discrete Choice and Dynamic Treatment Effects. J. Econom. 136 (2), 341–396.
Heckman, J.J., Rubinstein, Y., 2001. The Importance of Noncognitive Skills: Lessons from the GED Testing Program. Am. Econom. Rev. 91 (2), 145–149.
Heckman, J.J., Sedlacek, G.L., 1985. Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market. J. Polit. Econ. 93 (6), 1077–1125.
Heckman, J.J., Stixrud, J., Urzua, S., 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. J. Labor Econ. 24 (3), 411–482.
Heckman, J.J., Williams, B., 2011. Estimating Treatment Effects with Proxies for Controls. Unpublished manuscript. Department of Economics, University of Chicago.
Heitz, R.P., Unsworth, N., Engle, R.W., 2005. Working Memory Capacity, Attention Control, and Fluid Intelligence. In: Wilhelm., O., Engle, R.W., (Ed.), Handbook of Understanding and Measuring Intelligence. Sage Publications Inc., Thousand Oaks, CA, pp. 61–77.
Helson, R., Kwan, V.S.Y., John, O.P., Jones, C., 2002. The Growing Evidence for Personality Change in Adulthood: Findings from Research with Personality Inventories. J. Res. Personal. 36 (4), 287–306.
Helson, R., Picano, J., 1990. Is the Traditional Role Bad for Women? J. Personal. Soc. Psychol. 59 (2), 311–320.
Hembree, R., 1988. Correlates, Causes, Effects, and Treatment of Test Anxiety. Rev. Educ. Res. 58 (1), 47–77.
Hennig-Schmidt, H., Rockenbach, B., Sadrieh, A., 2010. In Search of Workers' Real Effort Reciprocity—a Field and a Laboratory Experiment. J. Eur. Econ. Assoc. 8 (4), 817–837.
Herrnstein, R.J., Murray, C.A., 1994. The Bell Curve: Intelligence and Class Structure in American Life. Free Press, New York.
Hicks, J.R., 1946. Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory. Clarendon Press, Oxford.
Hoffman, M., Fessler, D., Gneezy, U., List, J., 2010. Height and Competitiveness. Unpublished manuscript, Department of Economics, University of Chicago.
Hofstee, W.K., de Raad, B., Goldberg, L.R., 1992. Integration of the Big Five and Circumplex Approaches to Trait Structure. J. Personal. Soc. Psychol. 63 (1), 146–163.
Hogan, R., Hogan, J., 2007. Hogan Personality Inventory Manual. Hogan Assessment Systems, Tulsa, OK.
Hogan, R., Hogan, J., Roberts, B.W., 1996. Personality Measurement and Employment Decisions: Questions and Answers. Am. Psychol. 51 (5), 469–477.
Hogan, J., Holland, B., 2003. Using Theory to Evaluate Personality and Job-Performance Relations: A Socioanalytic Perspective. J. Appl. Psychol. 88 (1), 100–112.
Hogan, R., Roberts, B.W., 2001. Personality Psychology in the Workplace. American Psychological Association, Washington, DC.
Holmlund, H., Silva, O., 2009. Targeting Non-Cognitive Skills to Improve Cognitive Outcomes: Evidence from a Remedial Education Intervention. IZA Discussion Paper No. 4476.
Holt, M.M., Hobbs, T.R., 1979. The Effects of Token Reinforcement, Feedback and Response Cost on Standardized Test Performance. Behav. Res. Ther. 17 (1), 81–83.

Hopkins, K.D., Bracht, G.H., 1975. Ten-Year Stability of Verbal and Nonverbal IQ Scores. Am. Educ. Res. J. 12 (4), 469–477.

Horn, J.L., 1970. Organization of Data on Life-Span Development of Human Abilities. In: Goulet, L.R., Baltes, P.B., (Ed.), Life-Span Developmental Psychology: Research and Theory. Academic Press, New York, pp. 423–466.

Horn, J.L., McArdle, J.J., 2007. Understanding Human Intelligence since Spearman. In: Cudek, R., MacCallum, R.C., (Ed.), Factor Analysis at 100: Historical Developments and Future Directions. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 205–249.

Horvath, P., Zuckerman, M., 1993. Sensation Seeking, Risk Appraisal, and Risky Behavior. Pers. Individ. Dif. 14 (1), 41–52.

Hough, L.M., 1992. The "Big Five" Personality Variables–Construct Confusion: Description Versus Prediction. Hum. Perf. 5 (1–2), 139–155.

Hough, L.M., Eaton, N.K., Dunnette, M.D., Kamp, J.D., McCloy, R.A., 1990. Criterion-Related Validities of Personality Constructs and the Effect of Response Distortion on Those Validities. J. Appl. Psychol. 75 (5), 581–595.

Hough, L.M., Ones, D.S., 2002. The Structure, Measurement, Validity, and Use of Personality Variables in Industrial, Work, and Organizational Psychology. In: Anderson, N., Ones, D.S., Sinangil, H.K., Viswesvaran, C., (Ed.), Handbook of Industrial, Work and Organizational Psychology, Volume 1: Personnel Psychology. Sage Publications Inc., Thousand Oaks, CA, pp. 233–277.

Hough, L.M., Oswald, F.L., 2000. Personnel Selection: Looking toward the Future–Remembering the Past. Annu. Rev. Psychol. 51, 631–664.

Hsiao, C., 2003. Analysis of Panel Data. Cambridge University Press, New York.

Hunter, J.E., 1986. Cognitive Ability, Cognitive Aptitudes, Job Knowledge, and Job Performance. J. Vocat. Behav. 29 (3), 340–362.

Jackson, D.N., 1974. Personality Research Form Manual. Research Psychologists Press, Goshen, NY.

Jackson, J.J., Hill, P.L., Payne, B.R., Roberts, B.W., Stine-Morrow, E.A.L., 2010. Can an Old Dog Learn (and Want to Experience) New Tricks? Cognitive Training Increases Openness in Older Adults. Unpublished Manuscript. Department of Psychology, University of Illinois.

James, W., 1890. The Principles of Psychology. H. Holt and Company, New York.

Janosz, M., LeBlanc, M., Boulerice, B., Tremblay, R.E., 1997. Disentangling the Weight of School Dropout Predictors: A Test on Two Longitudinal Samples. J. Youth Adolesc. 26 (6), 733–762.

Jensen, A.R., 1998. The g Factor: The Science of Mental Ability. Praeger, Westport, CT.

John, O.P., 1990. The "Big Five" Factor Taxonomy: Dimensions of Personality in the Natural Language and in Questionnaires. In: Pervin, L.A., John, O.P., (Ed.), Handbook of Personality: Theory and Research. Guilford Press, New York, pp. 66–100.

John, O.P., Caspi, A., Robins, R.W., Moffitt, T.E., 1994. The "Little Five": Exploring the Nomological Network of the Five-Factor Model of Personality in Adolescent Boys. Child Dev. 65 (1), 160–178.

John, O.P., Robins, R.W., Pervin, L.A., 2008. Handbook of Personality: Theory and Research. The Guilford Press, New York.

John, O.P., Srivastava, S., 1999. The Big Five Trait Taxonomy: History, Measurement and Theoretical Perspectives. In: Pervin, L.A., John, O.P., (Ed.), Handbook of Personality: Theory and Research. The Guilford Press, New York, 102–138.

Johnson, W., Bouchard, Jr., T.J., Krueger, R.F., McGue, M., Gottesman, I.I., 2004. Just One g: Consistent Results from Three Test Batteries. Intelligence 32 (1), 95–107.

Johnson, J., Wilke, A., Weber, E.U., 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. Pol. Psychol. Bull. 35 (3), 153–172.

Judge, T.A., Bono, J.E., Erez, A., Locke, E.A., 2005. Core Self-Evaluations and Job and Life Satisfaction: The Role of Self-Concordance and Goal Attainment. J. Appl. Psychol. 90 (2), 257–268.

Judge, T.A., Erez, A., Bono, J.E., Thoresen, C.J., 2002. Are Measures of Self-Esteem, Neuroticism, Locus of Control, and Generalized Self-Efficacy Indicators of a Common Core Construct? J. Personal. Soc. Psychol. 83 (3), 693–710.

Judge, T.A., Hurst, C., 2007. Capitalizing on One's Advantages: Role of Core Self-Evaluations. J. Appl. Psychol. 92 (5), 1212–1227.
Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. Econometrica 47 (2), 263–291.
Kelly, F.J., Veldman, D.J., 1964. Delinquency and School Dropout Behavior as a Function of Impulsivity and Nondominant Values. J. Abnorm. Soc. Psychol. 69 (2), 190–194.
Kern, M.L., Friedman, H.S., 2008. Do Conscientious Individuals Live Longer? Health Psychol. 27 (5), 505–512.
Kern, M.L., Friedman, H.S., 2010a. Personality and Pathways of Influence on Physical Health. Soc. Personal. Psychol. Compass 5 (1), 76–87.
Kern, M.L., Friedman, H.S., 2010b. Why Do Some People Thrive While Other Succumb to Disease and Stagnation? Personality, Social Relations, and Resilience. In: Fry, P.S., Keyes, C.L.M., (Ed.), New Frontiers in Resilient Aging: Life-Strengths and Well-Being in Late Life. Cambridge University Press, Cambridge, pp. 162–184.
Kilburn, M.R., Hanser, L.M., Klerman, J.A., 1998. Information in the NELS and Approaches to Imputing Test Scores. Estimating the AFQT Scores for National Educational Longitudinal Study (NELS) Respondents. National Defense Research Institute, Arroya Center: Rand Corp, Santa Monica, CA.
Kirby, K.N., 1997. Bidding on the Future: Evidence against Normative Discounting of Delayed Rewards. J. Exp. Psychol. Gen. 126 (1), 54.
Kirby, K.N., Herrnstein, R.J., 1995. Preference Reversals Due to Myopic Discounting of Delayed Reward. Psychol. Sci. 6 (2), 83–89.
Knudsen, E.I., Heckman, J.J., Cameron, J.L., Shonkoff, J.P., 2006. Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce. Proc. Natl. Acad. Sci. 103 (27), 10155–10162.
Knutson, B., Wolkowitz, O.M., Cole, S.W., Chan, T., Moore, E.A., Johnson, R.C., Terpstra, J., Turner, R.A., Reus, V.I., 1998. Selective Alteration of Personality and Social Behavior by Serotonergic Intervention. Am. J. Psychiatry 155 (3), 373–379.
Kobrin, J.L., Patterson, B.F., Shaw, E.J., Mattern, K.D., Barbuti, S.M., 2008. Validity of the SAT for Predicting First-Year College Grade Point Average. Coll. Board 5:1–10.
Kogan, N., Wallach, M.A., 1967. Risk Taking as a Function of the Situation, the Person, and the Group. In: Mandler, G., Mussen, P., Kogan, N., Wallach, M., (Ed.), New Directions in Psychology III. Holt, Rinehart & Winston, New York, pp. 111–278.
Kohn, M.L., Schooler, C., 1978. The Reciprocal Effects of the Substantive Complexity of Work and Intellectual Flexibility: A Longitudinal Assessment. Am. J. Soc. 84 (1), 24–52.
Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E., 2005. Oxytocin Increases Trust in Humans. Nature 435 (7042), 673–676.
Köszegi, B., 2006. Ego Utility, Overconfidence, and Task Choice. J. Eur. Econ. Assoc. 4 (4), 673–707.
Kovas, Y., Haworth, C.M.A., Philip, S.D., Plomin, R., Weinberg, R.A., Thomson, J.M., Fischer, K.W., 2007. The Genetic and Environmental Origins of Learning Abilities and Disabilities in the Early School Years. Monogr. Soc. Res. Child Dev. 72 (3), vii–144.
Kristensen, P., Bjerkedal, T., 2007. Explaining the Relation between Birth Order and Intelligence. Science 316 (5832), 1717.
Krueger, R.F., Eaton, N., 2010. Personality Traits and the Classification of Mental Disorders: Toward a More Complete Integration in DSM 5 and an Empirical Model of Psychopathology. Personal. Disord. Theory Res. Treat. 1 (2), 97–118.
Krueger, R.F., Johnson, W., 2008. Behavioral Genetics and Personality: A New Look at the Integration of Nature and Nurture. In: John, O.P., Robins, R.W., Pervin, L.A., (Ed.), Handbook of Personality: Theory and Research. Guilford Press, New York, pp. 287–310.
Krueger, R.F., South, S., Johnson, W., Iacono, W.G., 2008. The Heritability of Personality Is Not Always 50%: Gene-Environment Interactions and Correlations between Personality and Parenting. J. Personal. 76 (6), 1485–1522.
Kube, S., Maréchal, M.A., Puppe, C., 2006. Putting Reciprocity to Work–Positive Versus Negative Responses in the Field. University of St. Gallen Working Paper, 2006–27.

Kubzansky, L.D., Sparrow, D., Vokonas, P., Kawachi, I., 2001. Is the Glass Half Empty or Half Full? A Prospective Study of Optimism and Coronary Heart Disease in the Normative Aging Study. Psychosom. Med. 63 (6), 910.

Kuncel, N.R., Hezlett, S.A., 2007. Standardized Tests Predict Graduate Students' Success. Science 315 (5815), 1080–1081.

Kuncel, N.R., Ones, D.S., Sackett, P.R., 2010. Individual Differences as Predictors of Work, Educational, and Broad Life Outcomes. Pers. Individ. Dif. 49 (4), 331–336.

Kvist, A.V., Gustafsson, J.-E., 2008. The Relation between Fluid Intelligence and the General Factor as a Function of Cultural Background: A Test of Cattell's Investment Theory. Intelligence 36 (5), 422–436.

Kyllonen, P.C., Walters, A.M., Kaufman, J.C., 2005. Noncognitive Constructs and Their Assessment in Graduate Education: A Review. Educ. Assess. 10 (3), 153–184.

Larson, G.E., Saccuzzo, D.P., Brown, J., 1994. Motivation: Cause or Confound in Information Processing/Intelligence Correlations? Acta Psychol 85 (1), 25–37.

Layard, R., 2005. Happiness: Lessons from a New Science. Penguin, London.

Le, A.T., Miller, P.W., Slutske, W.S., Martin, N.G., 2010. Are Attitudes Towards Economic Risk Heritable? Analysis Using the Australian Twin Study of Gambling. IZA Discussion Paper No. 4859.

Lear, G.R., 2004. Happy Lives and the Highest Good: Aristotle's Nicomachean Ethics. Princeton University Press, Princeton.

Lejuez, C.W., Aklin, W.M., Zvolensky, M.J., Pedulla, C.M., 2003. Evaluation of the Balloon Analogue Risk Task (BART) as a Predictor of Adolescent Real-World Risk-Taking Behaviours. J. Adolesc. 26 (4), 475–479.

Lejuez, C.W., Read, J.P., Kahler, C.W., Richards, J.B., Ramsey, S.E., Stuart, G.L., Strong, D.R., Brown, R.A., 2002. Evaluation of a Behavioral Measure of Risk Taking: The Balloon Analogue Risk Task (BART). J. Exp. Psychol. Appl. 8 (2), 75–84.

Levitt, S.D., List, J.A., 2007. What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? J. Econ. Perspect. 21 (2), 153–174.

Lezak, M.D., 1987. Relationships between Personality Disorders, Social Disturbances, and Physical Disability Following Traumatic Brain Injury. J. Head Trauma Rehabil. 2 (1), 57–69.

Lievens, F., Dilchert, S., Ones, D.S., 2009. Personality Scale Validities Increase Throughout Medical School. J. Appl. Psychol. 94 (6), 1514–1535.

Lillard, A., Else-Quest, N., 2006. The Early Years: Evaluating Montessori. Science 313 (5795), 1893–1894.

List, J.A., 2003. Does Market Experience Eliminate Market Anomalies? Q. J. Econom. 118 (1), 41–71.

List, J.A., 2006. The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions. J. Polit. Econ. 114 (1), 1–37.

List, J.A., 2007. On the Interpretation of Giving in Dictator Games. J. Polit. Econ. 115 (3), 482.

List, J.A., 2009. Social Preferences: Some Thoughts from the Field. Ann. Rev. Econom. 1, 563–583.

Lizzeri, A., Siniscalchi, M., 2008. Parental Guidance and Supervised Learning. Q. J. Econom. 123 (3), 1161–1195.

Lleras, C., 2008. Do Skills and Behaviors in High School Matter? The Contribution of Noncognitive Factors in Explaining Differences in Educational Attainment and Earnings. Soc. Sci. Res. 37 (3), 888–902.

Lochner, L.J., Moretti, E., 2004. The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports. Am. Econom. Rev. 94 (1), 155–189.

Lounsbury, J.W., Steel, R.P., Loveland, J.M., Gibson, L.W., 2004. An Investigation of Personality Traits in Relation to Adolescent School Absenteeism. J. Youth Adolesc. 33 (5), 457–466.

Lubbers, M.J., Van der Werf, M.P.C., Kuyper, H., Hendriks, A.A.J., 2010. Does Homework Behavior Mediate the Relation between Personality and Academic Performance? Learn. Individ. Differ. 20 (3), 203–208.

Lubinski, D., 2004. Introduction to the Special Section on Cognitive Abilities: 100 Years after Spearman's (1904) '"General Intelligence," Objectively Determined and Measured'. J. Personal. Soc. Psychol. 86 (1), 96.

Lucas, R.E., Donnellan, M.B., 2009. If the Person-Situation Debate Is Really over, Why Does It Still Generate So Much Negative Affect? J. Res. Personal. 43 (2), 146–149.

Lykken, D.T., 2007. A More Accurate Estimate of Heritability. Twin Res. Hum. Genet. 10 (1), 168–173.
Machin, S., Marie, O., Vujić, S., 2010. The Crime Reducing Effect of Education. CEP Discussion Paper No. 979.
Mandelbrot, B.B., 1962. Paretian Distributions and Income Maximization. Q. J. Econom. 76 (1), 57–85.
Martin, R.P., 1989. Activity Level, Distractibility, and Persistence: Critical Characteristics in Early Schooling. In: Kohnstamm, G.A., Bates, J.E., Rothbart, M.K., (Ed.), Temperament in Childhood. John Wiley and Sons Ltd., Chichester, England, pp. 451–461.
Martins, P.S., 2010. Can Targeted, Non-Cognitive Skills Programs Improve Achievement? IZA Discussion Paper No. 5266.
Martin, L.R., Friedman, H.S., Schwartz, J.E., 2007. Personality and Mortality Risk across the Life Span: The Importance of Conscientiousness as a Biopsychosocial Attribute. Health Psychol. 26 (4), 428–436.
Mataró, M., Jurado, Á., García-Sánchez, C., Barraquer, L., Costa-Jussá, F.R., Junqué, C., 2001. Long-Term Effects of Bilateral Frontal Brain Lesion: 60 Years after Injury with an Iron Bar. Arch. Neurol. 58 (7), 1139–1142.
Matsumoto, K., Tanaka, K., 2004. Conflict and Cognitive Control. Science 303 (5660), 969–970.
Mayer, J.D., Salovey, P., 1997. What Is Emotional Intelligence? In: Salovey, P., Sluyter, D.J., (Ed.), Emotional Development and Emotional Intelligence: Educational Implications. Basic Books, New York, p. 3.
McAdams, D.P., 2006. The Person: A New Introduction to Personality Psychology. J. Wiley and Sons, Hoboken, NJ.
McArdle, J.J., Hamagami, F., Meredith, W., Bradway, K.P., 2000. Modeling the Dynamic Hypotheses of Gf-Gc Theory Using Longitudinal Life-Span Data. Learn. Individ. Differ. 12 (1), 53–79.
McCabe, L.A., Cunnington, M., Brooks-Gunn, J., 2004. The Development of Self-Regulation in Young Children: Individual Characteristics and Environmental Contexts. In: Baumeister, R.F., Vohs, K.D., (Ed.), Handbook of Self-Regulation: Research, Theory, and Applications. Guilford Press, New York, pp. 340–356.
McClelland, M.M., Cameron, C.E., Connor, C.M., Farris, C.L., Jewkes, A.M., Morrison, F.J., 2007. Links between Behavioral Regulation and Preschoolers' Literacy, Vocabulary, and Math Skills. Dev. Psychol. 43 (4), 947–959.
McClelland, D.C., Koestner, R., Weinberger, J., 1989. How Do Self-Attributed and Implicit Motives Differ? Psychol. Rev. 96 (4), 690–702.
McClure, S.M., Laibson, D.I., Loewenstein, G.F., Cohen, J.D., 2004. Separate Neural Systems Value Immediate and Delayed Monetary Rewards. Science 306 (5695), 503–507.
McCrae, R.R., Costa, P.T., 1990. Personality in Adulthood. Guilford Press, New York.
McCrae, R.R., Costa, P.T., 1994. The Stability of Personality: Observation and Evaluations. Curr Dir. Psychol. Sci. 3 (6), 173–175.
McCrae, R.R., Costa, P.T., 1996. Toward a New Generation of Personality Theories: Theoretical Contexts for the Five-Factor Model. In: Wiggins, J.S. (Ed.), The Five-Factor Model of Personality: Theoretical Perspectives. Guilford Press, New York, pp. 51–87.
McCrae, R.R., Costa, P.T., 1997a. Conceptions and Correlates of Openness to Experience. In: Hogan, R., Johnson, J.A., Briggs, S.R., (Ed.), Handbook of Personality Psychology. Academic Press, San Diego, CA, pp. 825–847.
McCrae, R.R., Costa, P.T., 1997b. Personality Trait Structure as a Human Universal. Am. Psychol. 52 (5), 509–516.
McCrae, R.R., Costa, P.T., 2003. Personality in Adulthood: A Five-Factor Theory Perspective. Guilford Press, New York.
McCrae, R.R., Costa, Jr., P.T., 2008. The Five-Factor Theory of Personality. In: John, O.P., Robins, R.W., Pervin, L.A., (Ed.), Handbook of Personality: Theory and Research. The Guilford Press, New York, pp. 159–181.
McGee, A., 2010. How the Perception of Control Influences Unemployed Job Search. Unpublished manuscript. Department of Economics, Ohio State University.
McGue, M., Bacon, S., Lykken, D.T., 1993. Personality Stability and Change in Early Adulthood: A Behavioral Genetic Analysis. Dev. Psychol. 29 (1), 96–109.

McGue, M., Bouchard, T.J., Iacono, W.G., Lykken, D.T., 1993. Behavioral Genetics of Cognitive Ability: A Life-Span Perspective. In: Plomin, R., McClearn, G.E., (Ed.), Nature, Nurture and Psychology. American Psychological Association, Washington, DC, pp. 59–76.
McHenry, J.J., Hough, L.M., Toquam, J.L., Hanson, M.A., Ashworth, S., 1990. Project a Validity Results: The Relationship between Predictor and Criterion Domains. Pers. Psychol. 43 (2), 335–354.
Meier, S., 2007. A Survey of Economic Theories and Field Evidence on Pro-Social Behavior. In: Frey, B.S., Stutzer, A., (Ed.), Economics and Psychology: A Promising New Cross-Disciplinary Field. MIT Press, Cambridge, MA, pp. 51–88.
Mershon, B., Gorsuch, R.L., 1988. Number of Factors in the Personality Sphere: Does Increase in Factors Increase Predictability of Real-Life Criteria? J. Personal. Soc. Psychol. 55 (4), 675–680.
Michalko, K.T., Saklofske, D.H., 1996. A Psychometric Investigation of the Wechsler Individual Achievement Test with a Sample of Saskatchewan Schoolchildren. Can. J. Sch. Psychol. 12 (1), 44–54.
Miller, E.K., 2000. The Prefrontal Cortex: No Simple Matter. NeuroImage 11 (5), 447–450.
Miller, E.K., Cohen, J.D., 2001. An Integrative Theory of Prefrontal Cortex Function. Ann. Rev. Neurosci. 1 (1), 167–202.
Mischel, W., 1968. Personality and Assessment. Wiley, New York.
Mischel, W., 2004. Toward an Integrative Science of the Person. Ann. Rev. Psychol. 55: 1–22.
Mischel, W., 2009. From *Personality and Assessment* (1968) to Personality Science, 2009. J. Res. Personal. 43 (2), 282–290.
Mischel, W., Ayduk, O., Berman, M.G., Casey, B.J., Gotlib, I.H., Jonides, J., Kross, E., Teslovich, T., Wilson, N.L., Zayas, V., Shoda, Y., 2010. 'Willpower' over the Life Span: Decomposing Self-Regulation. Soc. Cogn. Affect. NeuroSci. http://scan.oxfordjournals.org/content/early/2010/09/18/scan.nsq081.full.pdf+html.
Mischel, W., Metzner, R., 1962. Preference for Delayed Reward as a Function of Age, Intelligence, and Length of Delay Interval. J. Abnorm. Soc. Psychol. 64 (6), 425–431.
Mischel, W., Shoda, Y., 1995. A Cognitive-Affective System Theory of Personality: Reconceptualizing Situations, Dispositions, Dynamics, and Invariance in Personality Structure. Psychol. Rev. 102 (2), 246–268.
Mischel, W., Shoda, Y., 2008. Toward a Unifying Theory of Personality: Integrating Dispositions and Processing Dynamics within the Cognitive-Affective Processing System. In: John, O.P., Robins, R.W., Pervin, L.A., (Ed.), Handbook of Personality Psychology. Guilford Press, New York, pp. 209–241.
Mischel, W., Shoda, Y., Rodriguez, M.L., 1989. Delay of Gratification in Children. Science 244 (4907), 933–938.
Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., 2000. The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. Cogn. Psychol. 41 (1), 49–100.
Möbius, M.M., Niederle, M., Niehaus, P., Rosenblat, T.S., 2010. Managing Self-Confidence: Theory and Experimental Evidence. Unpublished manuscript. Department of Economics, Iowa State University.
Moffitt, T.E., Caspi, A., Rutter, M., 2005. Strategy for Investigating Interactions between Measured Genes and Measured Environments. Arch. Gen. Psychiatry 62 (5), 473–481.
Monterosso, J.R., Luo, S., 2010. An Argument against Dual Valuation System Competition: Cognitive Capacities Supporting Future Orientation Mediate Rather Than Compete with Visceral Motivations. J. Neurosci. Psychol. Econom. 3 (1), 1–14.
Morgeson, F.P., Campion, M.A., Dipboye, R.L., Hollenbeck, J.R., Murphy, K., Schmitt, N., 2007. Reconsidering the Use of Personality Tests in Personnel Selection Contexts. Pers. Psychol. 60 (3), 683–729.
Moskowitz, D.S., 1982. Coherence and Cross-Situational Generality in Personality: A New Analysis of Old Problems. J. Personal. Soc. Psychol. 43 (4), 754–768.
Moutafi, J., Furnham, A., Tsaousis, I., 2006. Is the Relationship between Intelligence and Trait Neuroticism Mediated by Test Anxiety? Pers. Individ. Dif. 40 (3), 587.
Mroczek, D.K., Spiro, A., 2007. Personality Change Influences Mortality in Older Men. Psychol. Sci. 18 (5), 371–376.

Mueller, G., Plug, E., 2006. Estimating the Effect of Personality on Male and Female Earnings. Ind. Labor Relat. Rev. 60 (1), 3–22.

Mueser, P.R., 1979. The Effects of Noncognitive Traits. In: Jencks, C., Bartlett, S., Corcoran, M., Crouse, J., Eaglesfield, D., Jackson, G., McClelland, K., Mueser, P., Olneck, M., Schwartz, J., Ward, S., Williams, J., (Ed.), Who Gets Ahead? The Determinants of Economic Success in America. Basic Books, New York, pp. 122–158.

Mundlak, Y., 1978. On the Pooling of Time Series and Cross Section Data. Econometrica 46 (1), 69–85.

Murray, H.A., 1938. Explorations in Personality: A Clinical and Experimental Study of Fifty Men of College Age. Oxford University Press, New York.

Neal, D.A., Johnson, W.R., 1996. The Role of Premarket Factors in Black-White Wage Differences. J. Polit. Econ. 104 (5), 869–895.

Neisser, U., Boodoo, G., Bouchard, Jr., T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J., Urbina, S., 1996. Intelligence: Knowns and Unknowns. Am. Psychol. 51 (2), 77–101.

Newman, D.L., Caspi, A., Moffitt, T.E., Silva, P.A., 1997. Antecedents of Adult Interpersonal Functioning: Effects of Individual Differences in Age 3 Temperament. Dev. Psychol. 33 (2), 206–217.

Noftle, E.E., Robins, R.W., 2007. Personality Predictors of Academic Outcomes: Big Five Correlates of GPA and SAT Scores. J. Personal. Soc. Psychol. 93 (1), 116–130.

Nyhus, E.K., Pons, E., 2005. The Effects of Personality on Earnings. J. Econom. Psychol. 26 (3), 363–384.

O'Gorman, J.G., Baxter, E., 2002. Self-Control as a Personality Measure. Pers. Individ. Dif. 32 (3), 533–539.

O'Leary, U.-M., Rusch, K.M., Guastello, S.J., 1991. Estimating Age-Stratified WAIS-R IQs from Scores on the Raven's Standard Progressive Matrices. J. Clin. Psychol. 47 (2), 277–284.

Oliver, P.H., Guerin, D.W., Gottfried, A.W., 2007. Temperamental Task Orientation: Relation to High School and College Educational Accomplishments. Learn. Individ. Differ. 17 (3), 220–230.

Omizo, M.M., 1980. The Differential Aptitude Tests as Predictors of Success in a High School for Engineering Program. Educ. Psychol. Meas. 40 (1), 197–203.

Ones, D.S., Viswesvaran, C., 1998. The Effects of Social Desirability and Faking on Personality and Integrity Assessment for Personnel Selection. Hum. Perf. 11 (2–3):245–269.

Osiński, J., 2009. Kin Altruism, Reciprocal Altruism and Social Discounting. Pers. Individ. Dif. 47 (4), 374–378.

Pailing, P.E., Segalowitz, S.J., 2004. The Error-Related Negativity as a State and Trait Measure: Motivation, Personality, and Erps in Response to Errors. Psychophysiology 41 (1), 84–95.

Parco, J.E., Rapoport, A., Stein, W.E., 2002. Effects of Financial Incentives on the Breakdown of Mutual Trust. Psychol. Sci. 13 (3), 292–297.

Paulhus, D.L., 1984. Two-Component Models of Socially Desirable Responding. J. Personal. Soc. Psychol. 46 (3), 598–609.

Paunonen, S.V., Ashton, M.C., 2001. Big Five Factors and Facets and the Prediction of Behavior. J. Personal. Soc. Psychol. 81 (3), 524–539.

Perkins, D.N., Tishman, S., 2001. Dispositional Aspects of Intelligence. In: Collis, J.M., Messick, S., (Ed.), Intelligence and Personality: Bridging the Gap in Theory and Measurement. Lawrence Erlbaum Associates, Publishers, Mahwah, NJ, pp. 233–257.

Pesonen, A.-K., Räikkönen, K., Heinonen, K., Andersson, S., Hovi, P., Järvenpää, A.-L., Eriksson, J.G., Kajantie, E., 2008. Personality of Young Adults Born Prematurely: Helsinki Study of Very Low Birth Weight Adults. J. Child Psychol. Psychiatry 49 (6), 609–617.

Peters, E., Vastfjall, D., Slovic, P., Mertz, C.K., Mazzocco, K., Dickert, S., 2006. Numeracy and Decision Making. Psychol. Sci. 17 (5), 407–413.

Phelps, E.A., 2006. Emotion and Cognition: Insights from Studies of the Human Amygdala. Ann. Rev. Psychol. 57, 27–53.

Phelps, E.S., Pollak, R.A., 1968. On Second-Best National Saving and Game-Equilibrium Growth. Rev. Econom. Stud. 35 (2), 185–199.

Piatek, R., Pinger, P., 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, S., 2002. The Blank Slate: The Modern Denial of Human Nature. Penguin Books, New York.
Plomin, R., DeFries, J.C., Craig, I.W., McGuffin, P., 2002. Behavioral Genetics in the Postgenomic Era. American Psychological Association, Washington, DC.
Plomin, R., Nesselroade, J.R., 1990. Behavioral Genetics and Personality Change. J. Personal. 58 (1), 191–220.
Ponitz, C.C., McClelland, M.M., Jewkes, A.M., Connor, C.M., Farris, C.L., Morrison, F.J., 2008. Touch Your Toes! Developing a Direct Measure of Behavioral Regulation in Early Childhood. Early Child. Res. Q. 23 (2), 141–158.
Ponitz, C.C., McClelland, M.M., Matthews, J.S., Morrison, F.J., 2009. A Structured Observation of Behavioral Self-Regulation and Its Contribution to Kindergarten Outcomes. Dev. Psychol. 45 (3), 605–619.
Poropat, A.E., 2009. A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance. Psychol. Bull. 135 (2), 322–338.
Posner, M.I., Rothbart, M.K., 2000. Developing Mechanisms of Self-Regulation. Dev. Psychopathol. 12 (3), 427–441.
Project Head Start, 1969. The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development. Westinghouse Learning Corporation, Bladensburg, MD.
Putnam, S.P., Ellis, L.K., Rothbart, M.K., 2001. The Structure of Temperament from Infancy through Adolescence. In: Eliasz, A., Angleitner, A., (Ed.), Advances in Research on Temperament. Pabst Science Publishers, Lengerich, Germany, pp. 165–182.
Rabin, M., 2000. Diminishing Marginal Utility of Wealth Cannot Explain Risk Aversion. In: Kahneman, D., Tversky, A., (Ed.), Choices, Values, and Frames. Cambridge University Press, New York, pp. 202–208.
Raudenbush, S.W., Bryk, A.S., 2001. Hierarchical Linear Models: Applications and Data Analysis Methods (Advanced Quantitative Techniques in the Social Sciences). Sage Publications, Inc, Thousand Oaks, CA.
Raven, J.C., 1962. Advanced Progressive Matrices: Sets I and II. Lewis, London.
Raven, J., Raven, J.C., Court, J.H., 1988. Manual for Raven's Progressive Matrices and Vocabulary Scales. Harcourt Assessment, San Antonio, TX.
Raver, C.C., Zigler, E.F., 1997. Social Competence: An Untapped Dimension in Evaluating Head Start's Success. Early Child. Res. Q. 12 (4), 363–385.
Read, D., Read, N.L., 2004. Time Discounting over the Lifespan. Organ. Behav. Hum. Decis. Process. 94 (1), 22–32.
Revelle, W., 1997. Extraversion and Impulsivity: The Lost Dimension? In: Nyborg, H., (Ed.), The Scientific Study of Human Nature: Tribute to Hans J. Eysenck at Eighty. Pergamon/Elsevier Science, Amsterdam, p. 189.
Revelle, W., Wilt, J., Condon, D.M., 2011. Individual Differences and Differential Psychology: A Brief History and Prospect. In: Chamorro-Premuzic, T., Furnham, A., von Stumm, S., (Ed.), The Wiley-Blackwell Handbook of Individual Differences, Wiley-Blackwell, Hoboken, NJ.
Richard, F.D., Bond, C.F., Stokes-Zoota, J.J., 2003. One Hundred Years of Social Psychology Quantitatively Described. Rev. Gen. Psychol. 7 (4), 331–363.
Rindermann, H., 2007. The g-Factor of International Cognitive Ability Comparisons: The Homogeneity of Results in PISA, TIMSS, PIRLS and IQ-Tests across Nations. Eur. J. Personal. 21 (5), 667.
Roberts, B.W., 1997. Plaster or Plasticity: Are Adult Work Experiences Associated with Personality Change in Women? J. Personal. 65 (2), 205–232.
Roberts, B.W., 2006. Personality Development and Organizational Behavior. In: Staw, B., 27, (Ed.), Research on Organizational Behavior. Elsevier Science/JAI Press, Oxford, 1–41.
Roberts, B.W., 2007. Contextualizing Personality Psychology. J. Personal. 75 (6), 1071–1082.
Roberts, B.W., 2009. Back to the Future: *Personality and Assessment* and Personality Development. J. Res. Personal. 43 (2), 137–145.

Roberts, B.W., Chapman, C.N., 2000. Change in Dispositional Well-Being and Its Relation to Role Quality: A 30-Year Longitudinal Study. J. Res. Personal. 34 (1), 26–41.
Roberts, B.W., Chernyshenko, O.S., Stark, S., Goldberg, L.R., 2005. The Structure of Conscientiousness: An Empirical Investigation Based on Seven Major Personality Questionnaires. Pers. Psychol. 58 (1), 103–139.
Roberts, B.W., DelVecchio, W.F., 2000. The Rank-Order Consistency of Personality Traits from Childhood to Old Age: A Quantitative Review of Longitudinal Studies. Psychol. Bull. 126 (1), 3–25.
Roberts, R.D., Goff, G.N., Anjoul, F., Kyllonen, P.C., Pallier, G., Stankov, L., 2000. The Armed Services Vocational Aptitude Battery (ASVAB): Little More Than Acculturated Learning (Gc)!? Learn. Individ. Differ. 12 (1), 81–103.
Roberts, B.W., Harms, P., Smith, J.L., Wood, D., Webb, M., 2006. Using Multiple Methods in Personality Psychology. In: Eid, M., Diener, E., (Ed.), Handbook of Multimethod Measurement in Psychology. American Psychological Association, Washington, DC, pp. 321–335.
Roberts, B.W., Helson, R., Klohnen, E.C., 2002. Personality Development and Growth in Women across 30 Years: Three Perspectives. J. Personal. 70 (1), 79–102.
Roberts, B.W., Jackson, J.J., 2008. Sociogenomic Personality Psychology. J. Personal. 76 (6), 1523–1544.
Roberts, B.W., Kuncel, N.R., Shiner, R., Caspi, A., Goldberg, L.R., 2007. The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes. Perspect. Psychol. Sci. 2 (4), 313–345.
Roberts, R.D., Markham, P.M., Matthews, G., Zeidner, M., 2005. Assessing Intelligence: Past, Present, and Future. In: Wilhelm, O., Engle, R.W., (Ed.), Handbook of Understanding and Measuring Intelligence. Sage Publications Inc., Thousand Oaks, CA, pp. 333–360.
Roberts, B.W., Mroczek, D., 2008. Personality Trait Change in Adulthood. Curr. Dir. Psychol. Sci. 17 (1), 31–35.
Roberts, B.W., Walton, K.E., Viechtbauer, W., 2006. Patterns of Mean-Level Change in Personality Traits across the Life Course: A Meta-Analysis of Longitudinal Studies. Psychol. Bull. 132 (1), 1–25.
Roberts, B.W., Wood, D., Caspi, A., 2008. Personality Development. In: John, O.P., Robins, R.W., Pervin, L.A., (Ed.), Handbook of Personality: Theory and Research. Guilford Press, New York, pp. 375–398.
Robson, A.J., 1996. A Biological Basis for Expected and Non-Expected Utility. J. Econom. Theory 68, 397–424.
Robson, A.J., 2001. Why Would Nature Give Individuals Utility Functions? J. Polit. Econ. 109 (4), 900–914.
Robson, A.J., Samuelson, L., 2007. The Evolution of Intertemporal Preferences. Am. Econ. Rev. 97 (2), 496–500.
Robson, A.J., Samuelson, L., 2009. The Evolution of Time Preference with Aggregate Uncertainty. Am. Econ. Rev. 99 (5), 1925–1953.
Rodríguez-Planas, N., 2010. Mentoring, Educational Services, and Economic Incentives: Longer-Term Evidence on Risky Behaviors from a Randomized Trial. IZA Discussion Paper No. 4968.
Rosenberg, M., 1965. Society and the Adolescent Self-Image. Princeton University Press, Princeton, NJ.
Rosenberg, M., 1989. Society and the Adolescent Self-Image (Rev. Ed.). Wesleyan University Press, Middletown, CT.
Ross, L., Nisbett, R.E., 1991. The Person and the Situation: Perspectives of Social Psychology. McGraw-Hill, New York.
Rothbart, M.K., Ahadi, S.A., Evans, D.E., 2000. Temperament and Personality: Origins and Outcomes. J. Personal. Soc. Psychol. 78 (1), 122–135.
Rothlisberg, B., 1987. Comparing the Stanford-Binet, Fourth Edition to the WISC-R: A Concurrent Validity Study. J. Sch. Psychol. 25 (2), 193–196.
Rotter, J.B., 1966. Generalized Expectancies for Internal Versus External Control of Reinforcement. Psychol. Monogr. 80 (1), 1–28.
Roy, A.D., 1951. Some Thoughts on the Distribution of Earnings. Oxf. Econ. Pap. 3 (2), 135–146.

Rueda, M.R., Rothbart, M.K., McCandliss, B.D., Saccomanno, L., Posner, M.I., 2005. Training, Maturation, and Genetic Influences on the Development of Executive Attention. Proc. Natl. Acad. Sci. 102 (41), 14931–14936.
Rust, J., 2008. Dynamic Programming. In: Durlauf, S.N., Blume, L.E. (Ed.), The New Palgrave Dictionary of Economics. Palgrave Macmillan, Ltd, London.
Rutter, M., 2006a. Genes and Behavior: Nature-Nurture Interplay Explained. Blackwell Publishers, Oxford.
Rutter, M., 2006b. Implications of Resilience Concepts for Scientific Understanding. Ann. N.Y. Acad. Sci. 1094 (1), 1–12.
Ryden, A., Sullivan, M., Torgerson, J.S., Karlsson, J., Lindroos, A.K., Taft, C., 2003. Severe Obesity and Personality: A Comparative Controlled Study of Personality Traits. Int. J. Obes. 27 (12), 1534–1540.
Sackett, P.R., Borneman, M.J., Connelly, B.S., 2008. High Stakes Testing in Higher Education and Employment: Appraising the Evidence for Validity and Fairness. Am. Psychol. 63 (4), 215–227.
Sackett, P.R., Hardison, C.M., Cullen, M.J., 2004. On Interpreting Stereotype Threat as Accounting for African American-White Differences on Cognitive Tests. Am. Psychol. 59 (1), 7–13.
Safire, W., 1989. On Language; the Elision Fields. N.Y. Times 13 August.
Sahm, C.R., 2007. Does Risk Tolerance Change? Unpublished manuscript, University of Michigan Institute for Social Research.
Salgado, J.F., 1997. The Five Factor Model of Personality and Job Performance in the European Community. J. Appl. Psychol. 82 (1), 30–43.
Samuel, D.B., Widiger, T.A., 2008. A Meta-Analytic Review of the Relationships between the Five-Factor Model, DSM-IV-Tr Personality Disorders: A Facet Level Analysis. Clin. Psychol. Rev. 28 (8), 1326–1342.
Samuelson, P.A., 1956. Social Indifference Curves. Q. J. Econom. 70 (1), 1–22.
Saucier, G., 2003. An Alternative Multi-Language Structure for Personality Attributes. Eur. J. Pers. 17 (3), 179–205.
Savage, L.J., 1954. The Foundations of Statistics. Wiley, New York.
Savelyev, P.A., 2010. Conscientiousness, Education, and Longevity of High-Ability Individuals. Unpublished manuscript. Department of Economics, University of Chicago.
Scarr, S., Weinberg, R.A., Waldman, I.D., 1993. IQ Correlations in Transracial Adoptive Families. Intelligence 17 (4), 541–555.
Schaie, K.W., 1994. The Course of Adult Intellectual Development. Am. Psychol. 49 (4), 304–313.
Schmeichel, B.J., Vohs, K.D., Baumeister, R.F., 2003. Intellectual Performance and Ego Depletion: Role of the Self in Logical Reasoning and Other Information Processing. J. Personal. Soc. Psychol. 85 (1), 33–46.
Schmidt, F.L., Hunter, J.E., 1983. Individual Differences in Productivity: An Empirical Test of Estimates Derived from Studies of Selection Procedure Utility. J. Appl. Psychol. 68 (3), 407–414.
Schmidt, F.L., Hunter, J.E., 1998. The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. Psychol. Bull. 124 (2), 262–274.
Schmidt, F.L., Hunter, J., 2004. General Mental Ability in the World of Work: Occupational Attainment and Job Performance. J. Personal. Soc. Psychol. 86 (1), 162–173.
Schmidt, F.L., Hunter, J.E., Larson, M., 1988. General Cognitive Ability Versus General and Specific Aptitudes in the Prediction of Training Performance: Some Preliminary Findings. Paper presented at the 96th Annual Convention of the American Psychological Association, Atlanta, GA.
Schooler, C., Mulatu, M.S., Oates, G., 1999. The Continuing Effects of Substantively Complex Work on the Intellectual Functioning of Older Workers. Psychol. Aging 14 (3), 483–506.
Schuerger, J.M., Witt, A.C., 1989. The Temporal Stability of Individually Tested Intelligence. J. Clin. Psychol. 45 (2), 294–302.
Schulz, R., Bookwala, J., Knapp, J.E., Scheier, M., Williamson, G.M., 1996. Pessimism, Age, and Cancer Mortality. Psychol. Aging 11, 304–309.
Schumpeter, J.A., 1954. History of Economic Analysis. Oxford University Press, New York.

Schwartz, S.H., 1992. The Universal Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. In: Zanna, M.P., (Ed.), Advances in Experimental Social Psychology, Academic Press, Inc., San Diego, CA. pp. 1–65.
Schweinhart, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R., Nores, M., 2005. Lifetime Effects: The High/Scope Perry Preschool Study through Age 40. High/Scope Press, Ypsilanti, MI.
Segal, C., 2008. Motivation, Test Scores, and Economic Success. Department of Economics and Business, Universitat Pompeu Fabra, Working Paper No. 1124.
Sell, A., Tooby, J., Cosmides, L., 2009. Formidability and the Logic of Human Anger. Proc. Natl. Acad. Sci. 106 (35), 15073–15078.
Shamosh, N., Gray, J.R., 2007. Delay Discounting and Intelligence: A Meta-Analysis. Department of Psychology, Yale University, Unpublished manuscript.
Shiner, R.L., 1998. How Shall We Speak of Children's Personalities in Middle Childhood? A Preliminary Taxonomy. Psychol. Bull. 124 (3), 308–332.
Shiner, R., Caspi, A., 2003. Personality Differences in Childhood and Adolescence: Measurement, Development, and Consequences. J. Child Psychol. Psychiatry 44 (1), 2–32.
Siegler, R.S., 1992. The Other Alfred Binet. Dev. Psychol. 28 (2), 179.
Sijtsma, K., 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. Econom. Theory 12 (2), 293–312.
Slonim, R., Roth, A.E., 1998. Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic. Econometrica 66 (3), 569–596.
Slovic, P., 1962. Convergent Validation of Risk Taking Measures. J. Abnorm. Soc. Psychol. 65 (1), 68–71.
Slovic, P., 1966. Risk-Taking in Children: Age and Sex Differences. Child Dev. 37 (1), 169–176.
Slovic, P., 1972. Psychological Study of Human Judgment: Implications for Investment Decision Making. J. Finance 27 (4), 779–799.
Smith, M.L., Glass, G.V., Miller, T.I., 1980. The Benefits of Psychotherapy. Johns Hopkins University Press, Baltimore.
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, Washington, DC.
Spear, L.P., 2000a. The Adolescent Brain and Age-Related Behavioral Manifestations. Neurosci. Biobehav. Rev. 24 (4), 417.
Spear, L.P., 2000b. Neurobehavioral Changes in Adolescence. Curr. Dir. Psychol. Sci. 9 (4), 111.
Spearman, C., 1904. "General Intelligence," Objectively Determined and Measured. Am. J. Psychol. 15 (2), 201–293.
Spearman, C., 1927. The Abilities of Man; Their Nature and Measurement. The Macmillan Company, New York.
Stankov, L., 2005. g Factor: Issues of Design and Interpretation. In: Wilhelm, O., Engle, R.W., (Ed.), Handbook of Understanding and Measuring Intelligence. Sage Publications Inc., Thousand Oaks, CA, 279–293.
Stanovich, K.E., 1999. Who Is Rational?: Studies of Individual Differences in Reasoning. Lawrence Erlbaum Associates, Mahwah, NJ.
Starmer, C., 2000. Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk. J. Econom. Lit. 38 (2), 332–382.
Steele, C.M., Aronson, J., 1998. Stereotype Threat and the Test Performance of Academically Successful African Americans. In: Jencks, C., Phillips, M., (Ed.), The Black-White Test Score Gap. Brookings Institution Press, Washington, DC, pp. 401–427.
Steinberg, L., 2008. A Social Neuroscience Perspective on Adolescent Risk-Taking. Dev. Rev. 28 (1), 78–106.
Steinberg, L., Cauffman, E., Woolard, J., Graham, S., Banich, M., 2009. Are Adolescents Less Mature Than Adults? Minors' Access to Abortion, the Juvenile Death Penalty, and the Alleged APA 'Flip-Flop'. Am. Psychol. 64 (7), 583–594.

Sternberg, R.J., 2001. Successful Intelligence: Understanding What Spearman Had Rather Than What He Studied. In: Collis, J.M., Messick, S., Schiefele, U., (Ed.), Intelligence and Personality: Bridging the Gap in Theory and Measurement. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 347–373.
Sternberg, R.J., Forsythe, G.B., Hedlund, J., Horvath, J.A., Wagner, R.K., Williams, W.M., Snook, S.A., Grigorenko, E., 2000. Practical Intelligence in Everyday Life. Cambridge University Press, New York.
Stevens, C., Fanning, J., Coch, D., Sanders, L., Neville, H., 2008. Neural Mechanisms of Selective Auditory Attention Are Enhanced by Computerized Training: Electrophysiological Evidence from Language-Impaired and Typically Developing Children. Brain Res. 1205, 55–69.
Sticht, T.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.
Sticht, T.G., Hooke, L.R., Caylor, J.S., 1982. Literacy, Oracy, and Vocational Aptitude as Predictors of Attrition and Promotion in the Armed Services. Human Resources Research Organization, Washington, DC.
Störmer, S., Fahr, R., 2010. Individual Determinants of Work Attendance: Evidence on the Role of Personality. IZA Discussion Paper No. 4927.
Sutter, M., Feri, F., Kocher, M.G., Martinsson, P., Nordblom, K., Rützler, D., 2010. Social Preferences in Childhood and Adolescence: A Large-Scale Experiment. IZA Discussion Paper No. 5016.
Sylva, K., 1997. The Quest for Quality in Curriculum. In: Schweinhart, L.J., Weikart, D.P., (Ed.), Lasting Differences: The High/Scope Preschool Curriculum Comparison Study through Age 23. High/Scope Press, Ypsilanti, pp. 89–93.
Tang, T.Z., DeRubeis, R.J., Hollon, S.D., Amsterdam, J., Shelton, R., Schalet, B., 2009. Personality Change During Depression Treatment: A Placebo-Controlled Trial. Arch. Gen. Psychiatry 66 (12), 1322–1330.
Tausch, F., Potters, J., Riedl, A., 2010. Preferences for Redistribution and Pensions: What Can We Learn from Experiments? CESifo Working Paper No. 3156.
Tellegen, A., 1985. Structures of Mood and Personality and Their Relevance to Assessing Anxiety, with an Emphasis on Self-Report. In: Tuma, A.H., Maser, J.D., (Ed.), Anxiety and the Anxiety Disorders. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 681–706.
Terman, L.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. Houghton Mifflin Co, Boston.
Terman, L.M., Merrill, M.A., 1960. Stanford-Binet Intelligence Scale: Manual for the Third Revision Form, L-M. Houghton Mifflin, Boston.
Thaler, R.H., 1980. Toward a Positive Theory of Consumer Choice. J. Econ. Behav. Organ. 1 (1), 39–60.
Thaler, R.H., 2008, July 25–27. A Short Course in Behavioral Economics. Edge Master Class 2008 Retrieved November 4, 2010, http://www.edge.org/3rd_culture/thaler_sendhil08/class6.html.
Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., Gottesman, I.I., 2003. Socioeconomic Status Modifies Heritability of IQ in Young Children. Psychol. Sci. 14 (6), 623–628.
Uylings, H.B.M., 2006. Development of the Human Cortex and the Concept Of "Critical" Or "Sensitive" Periods. Lang. Learn. 56 (Suppl 1), 59–90.
Valiente, C., Lemery-Chalfant, K., Castro, K.S., 2007. Children's Effortful Control and Academic Competence. Merrill Palmer Q. 53 (1), 1–25.
Valiente, C., Lemery-Chalfant, K., Swanson, J., 2010. Prediction of Kindergartners' Academic Achievement from Their Effortful Control and Emotionality: Evidence for Direct and Moderated Relations. J. Educ. Psychol. 102 (3), 550–560.
Valiente, C., Lemery-Chalfant, K., Swanson, J., Reiser, M., 2008. Prediction of Children's Academic Competence from Their Effortful Control, Relationships, and Classroom Participation. J. Educ. Psychol. 100 (1), 67–77.
van Eijck, K., de Graaf, P.M., 2004. The Big Five at School: The Impact of Personality on Educational Attainment. Neth. J. Soc. Sci. 40 (1), 24–40.
van Praag, B.M.S., 1985. Linking Economics with Psychology: An Economist's View. J. Econom. Psychol. 6 (3), 289–311.

van Praag, B.M.S., van Weeren, J., 1988. Memory and Anticipation Processes and Their Significance for Social Security and Income Inequality. In: Maital, S., (Ed.), Applied Behavioral Economics Volume 1. New York University Press, New York, pp. 731–751.

Vaughn, B.E., Kopp, C.B., Krakow, J.B., 1984. The Emergence and Consolidation of Self-Control from Eighteen to Thirty Months of Age: Normative Trends and Individual Differences. Child Dev. 55 (3), 990–1004.

Vazsonyi, A.T., Pickering, L.E., Junger, M., Hessing, D., 2001. An Empirical Test of a General Theory of Crime: A Four-Nation Comparative Study of Self-Control and the Prediction of Deviance. J. Res. Crime Delinq. 38, 91–131.

Viswesvaran, C., Ones, D.S., 1999. Meta-Analyses of Fakability Estimates: Implications for Personality Measurement. Educ. Psychol. Meas. 59 (2), 197–210.

Vitaro, F., Brendgen, M., Larose, S., Tremblay, R.E., 2005. Kindergarten Disruptive Behaviors, Protective Factors, and Educational Achievement by Early Adulthood. J. Educ. Psychol. 97 (4), 617–629.

Wang, L., 1993. The Differential Aptitude Test: A Review and Critique. Annual Meeting of the Southwest Educational Research Association, Austin, TX.

Watson, D., Clark, L.A., Chmielewski, M., 2008. Structures of Personality and Their Relevance to Psychopathology: II. Further Articulation of a Comprehensive Unified Trait Structure. Journal of Personality, 76 (6), 1545–1586.

Webb, E., 1915. Character and Intelligence. Br. J. Psychol. Suppl. 1 (3), 1–99.

Weber, E.U., 2001. Personality and Risk Taking. In: Smelser, N.J., Baltes, P.B., (Ed.), International Encyclopedia of the Social and Behavioral Sciences. Pergamon, Oxford, p. 11274.

Weber, E.U., Blais, A.-R., Betz, N.E., 2002. A Domain-Specific Risk-Attitude Scale: Measuring Risk Perceptions and Risk Behaviors. J. Behav. Decis. Mak. 15 (4), 263–290.

Weber, E.U., Hsee, C., 1998. Cross-Cultural Differences in Risk Perception, but Cross-Cultural Similarities in Attitudes Towards Perceived Risk. Manag. Sci. 44 (9), 1205.

Weber, E.U., Milliman, R.A., 1997. Perceived Risk Attitudes: Relating Risk Perception to Risky Choice. Manag. Sci. 43 (2), 123.

Wechsler, D., 1939. The Measurement of Adult Intelligence. Williams & Wilkins Co., Baltimore, MD.

Weikart, D.P., Epstein, A.S., Schweinhart, L., Bond, J.T., 1978. The Ypsilanti Preschool Curriculum Demonstration Project: Preschool Years and Longitudinal Results. In: Monographs of the High/Scope Educational Research Foundation. High/Scope Press, Ypsilanti, MI, p. 4.

Whisenton, J.T., Lorre, M.R., 1970. A Comparison of the Values, Needs, and Aspirations of School Leavers with Those of Non-School Leavers. J. Negro Educ. 39 (4), 325–332.

Widiger, T.A., Costa, Jr., P.T., 2002. Five-Factor Model Personality Disorder Research. In: Costa, Jr., P.T., Widiger, T.A., (Ed.), Personality Disorders and the Five-Factor Model of Personality. American Psychological Association, Washington, DC.

Widiger, T.A., Trull, T.J., Clarkin, J.F., Sanderson, C., Costa, Jr., P.T., 2002. A Description of the DSM-IV Personality Disorders with the Five-Factor Model of Personality. In: Costa, Jr., P.T., Widiger, T.A. (Ed.), Personality Disorders and the Five-Factor Model of Personality. American Psychological Association, Washington, DC.

Wilhelm, O., Engle, R.W., 2005. Intelligence: A Diva and a Workhorse. In: Wilhelm, O., Engle, R.W. (Ed.), Handbook of Understanding and Measuring Intelligence. Sage Publications, Thousand Oaks, CA.

Willingham, W.W., 1985. Success in College: The Role of Personal Qualities and Academic Ability. College Entrance Examination Board, New York.

Willingham, W.W., Pollack, J.M., Lewis, C., 2002. Grades and Test Scores: Accounting for Observed Differences. J. Educ. Meas. 39 (1), 1–37.

Willis, R.J., Rosen, S., 1979. Education and Self-Selection. J. Polit. Econ. 87 (5, Part 2), S7–S36.

Winship, C., Korenman, S., 1997. Does Staying in School Make You Smarter? The Effect of Education on IQ in the Bell Curve. In: Devlin, B., Feinberg, S., Resnick, D., Roeder, K., (Ed.), Intelligence, Genes, and Success: Scientists Respond to the Bell Curve. Springer, Copernicus, New York, pp. 215–234.

Wolfe, R.N., Johnson, S.D., 1995. Personality as a Predictor of College Performance. Educ. Psychol. Meas. 55 (2), 177–185.

Wong, M.M., Csikszentmihalyi, M., 1991. Motivation and Academic Achievement: The Effects of Personality Traits and the Duality of Experience. J. Personal. 59 (3), 539–574.
Wood, D., Roberts, B.W., 2006. Cross-Sectional and Longitudinal Tests of the Personality and Role Identity Structural Model (PRISM). J. Personal. 74 (3), 779–809.
Young, J.W., Kobrin, J.L., 2001. Differential Validity, Differential Prediction, and College Admission Testing: A Comprehensive Review and Analysis. College Entrance Examination Board, New York, NY.
Zentner, M., Bates, J.E., 2008. Child Temperament: An Integrative Review of Concepts, Research Programs, and Measures. Eur. J. Dev. Sci. 2 (1–2), 7–37.
Zhou, Q., Main, A., Wang, Y., 2010. The Relations of Temperamental Effortful Control and Anger/Frustration to Chinese Children's Academic Achievement and Social Adjustment: A Longitudinal Study. J. Educ. Psychol. 102 (1), 180–196.
Zigler, E., Butterfield, E.C., 1968. Motivational Aspects of Changes in IQ Test Performance of Culturally Deprived Nursery School Children. Child Dev. 39 (1), 1–14.
Zuckerman, M., 1994. Behavioral Expressions and Biosocial Bases of Sensation Seeking. Cambridge University Press, New York.
Zuckerman, M., 2007. Sensation Seeking and Risky Behavior. American Psychological Association, Washington, DC.
Zuckerman, M., Kolin, E.A., Price, L., Zoob, I., 1964. Development of a Sensation-Seeking Scale. J. Consult. Psychol. 28 (6), 477–482.