Universiteit
Leiden
The Netherlands

Fostering and measuring skills: improving cognitive and non-cognitive skills to promote lifetime success

Kautz, T.; Heckman, J.J.; Diris, R.; Ter Weel, B.; Borghans, L.

Citation
Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: improving cognitive and non-cognitive skills to promote lifetime success*. *OECD Education Working Papers*. OECD. doi:10.1787/5jxsr7vr78f7-en

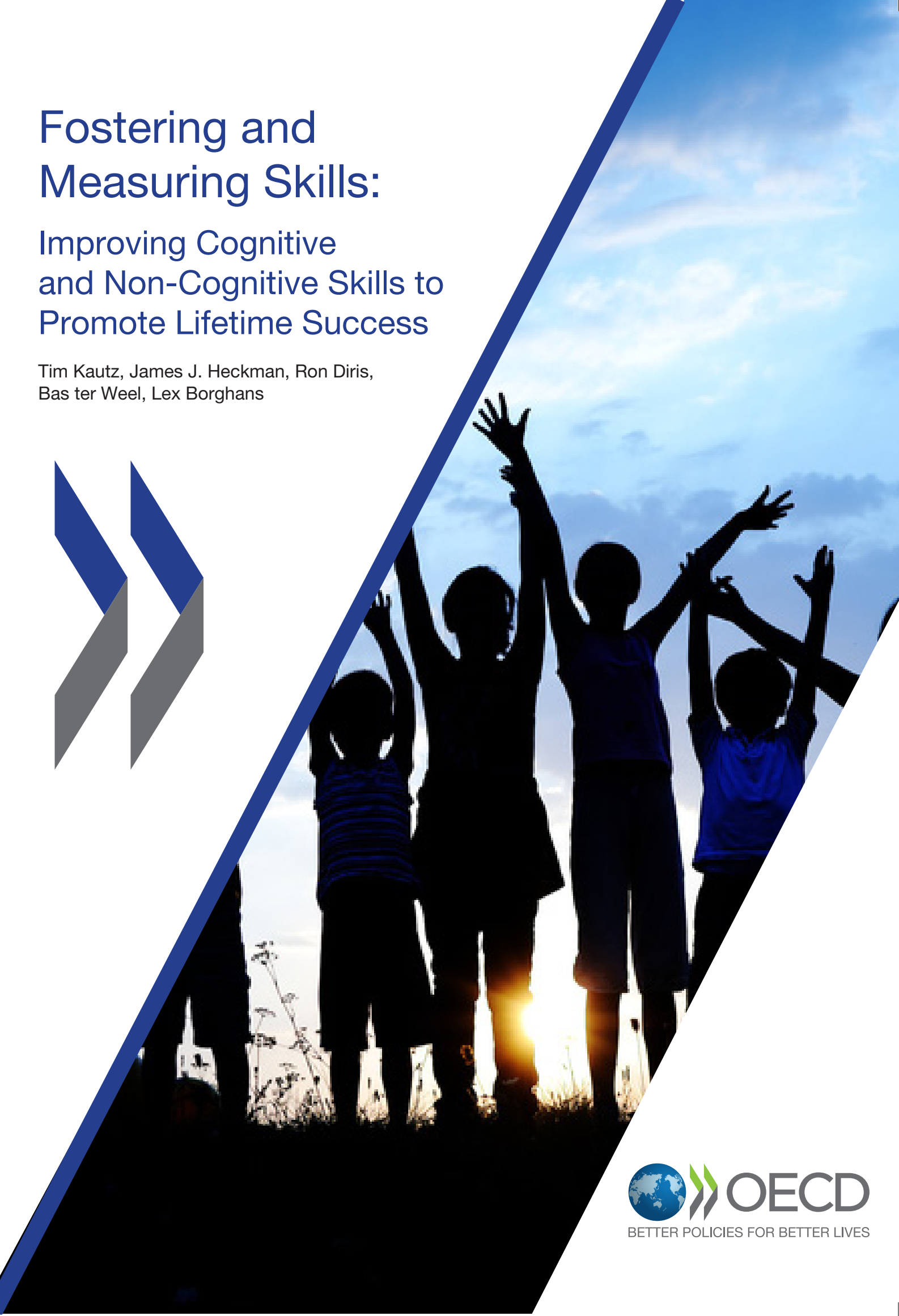
Version: Not Applicable (or Unknown)
License: Leiden University Non-exclusive license
Downloaded from: https://hdl.handle.net/1887/3146399

Note: To cite this publication please use the final published version (if applicable).

Fostering and Measuring Skills:

Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success

Tim Kautz, James J. Heckman, Ron Diris,
Bas ter Weel, Lex Borghans



OECD
BETTER POLICIES FOR BETTER LIVES

FOSTERING AND MEASURING SKILLS: IMPROVING COGNITIVE AND NON-COGNITIVE SKILLS TO PROMOTE LIFETIME SUCCESS

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Photo credits: © Shutterstock

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of the source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to *rights@oecd.org*. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at *info@copyright.com* or the Centre français d'exploitation du droit de copie (CFC) at *contact@cfcopies.com*.

ACKNOWLEDGEMENTS

This report was commissioned by the OECD through its project on Education and Social Progress.

We thank Linor Kiknadze and Edward Sung for valuable research assistance. We received helpful comments from Richard Boyle, Zidi Chen, Maryclare Griffin, Robert Lerman, Seong Hyeok Moon, Dan Moran, Maria Rosales, and Indra Wechsberg.

TABLE OF CONTENTS

EXECUTIVE SUMMARY 7

INTRODUCTION 9

Discarding Obsolete Notions and Embracing the Findings of Recent Research 11
The Organisation of the Rest of the Paper 12

MEASURES OF COGNITIVE AND NON-COGNITIVE SKILLS 13

Cognitive Skills 13
Measuring Non-cognitive Skills 13
A Task-Based Framework for Identifying and Measuring Skills 16
Reference Bias 18
Measuring Skills Using Behaviours 20
Are Non-Cognitive Skills Stable? 21

THE PREDICTIVE POWER OF NON-COGNITIVE SKILLS 23

Correlational Evidence 23
Evidence from the General Educational Development (GED) Programme 25

THE SKILLS NEEDED FOR SUCCESS IN THE LABOUR MARKET 29

A FRAMEWORK FOR UNDERSTANDING INTERVENTIONS 31

SUMMARY OF EMPIRICAL EVIDENCE ON THE EFFICACY OF INTERVENTIONS 33

EARLY-LIFE INTERVENTIONS THAT BEGIN BEFORE FORMAL SCHOOLING 35

Infant Programmes and Model Preschools 38
Nurse-Family Partnership 38
Jamaican Study 40
Perry Preschool Programme 40
Abecedarian Programme 42
Large-Scale Infant and Young Child Programmes 43
Head Start 43
Chicago Child–Parent Center (CPC) 44

EDUCATION AND INTERVENTIONS IN KINDERGARTEN AND ELEMENTARY SCHOOL 45

Targeted Non-Cognitive Skill Interventions 45
The Seattle Social Development Project (SSDP) 45
The Montreal Longitudinal Experimental Study (MLES) 47
Cambridge-Somerville Program 48
Project STAR 48

EDUCATION AND INTERVENTIONS TARGETED TOWARD ADOLESCENTS AND YOUNG ADULTS 49

Adolescent Mentorship Programmes 49

Quantum Opportunity Program 49
Becoming a Man 50
Pathways to Education Programme 50
Big Brothers Big Sisters 51
Empresários Pela Inclusão Social (EPIS) Program 51
H&R Block FAFSA experiment 51
Dartmouth College Coaching Program 52
Residential-Based Programmes 52
Job Corps 52
National Guard ChalleNGe 52
Workplace-Based Adolescent Intervention Programmes 53
Career Academies 54
Year-Up Programme 55
Self-Sufficiency Project 56
Apprenticeship Programmes 56
OTHER CURRICULA THAT HAVE BEEN APPLIED TO MULTIPLE AGE GROUPS 59
Tools of the Mind 59
Studies that Teach the Incremental Theory of Intelligence 59
Prevention vs. Remediation 60
THE EFFECTS OF EDUCATION AND PARENTAL INVESTMENT ON COGNITIVE AND NON-COGNITIVE SKILL 61
SUMMARY 66
REFERENCES 67

Tables

Table 1. The Big Five Domains and Their Facets 15
Table 2. Summary of Effects for Main Interventions 36
Table 3. Summary of Effects of the Nurse-Family Partnership 39
Table 4. Summary of Treatment Effects from the Seattle Social Development Project 47
Table 5. Summary of Treatment Effects from Career Academies within 96-Month Follow-Up after Scheduled High School Graduation 55

Figures

Figure 1. Determinants of Task Performance 16
Figure 2. Decomposing Variance Explained for Achievement Tests and Grades into IQ and Non-Cognitive Skills: Stella Maris Secondary School, Maastricht, Holland 18
Figure 3. National Rank in Big Five Conscientiousness and Average Annual Hours Worked 19
Figure 4. Associations with Job Performance 23
Figure 5. Correlations of Mortality with Non-Cognitive Skills, IQ, and Socioeconomic Status (SES) 24
Figure 6. Distribution of Cognitive Ability by Educational Status (No College Sample, All Ethnic Groups) 26
Figure 7. Hourly Wage Differences of GED Recipients and Traditional Graduates Compared to Uncertified Dropouts - Ages 20-39 27
Figure 8. Distribution of Non-Cognitive Skills by Education Group 28
Figure 9. Framework for Understanding Skill Development 32
Figure 10. Perry Preschool Program: IQ, by Age and Treatment Group 41
Figure 11. Histograms of Indices of Non-Cognitive Skills and CAT Scores 42
Figure 12. Casual Effect of Schooling on ASVAB Measures of Cognition 62
Figure 13. Casual Effect of Schooling on Two Measures of Non-Cognitive Skill 63

EXECUTIVE SUMMARY[1]

This paper reviews the recent literature on measuring and fostering cognitive and non-cognitive skills. IQ tests and achievement tests do not adequately capture non- cognitive skills personality traits, goals, character, motivations, and preferences that are valued in the labour market, in school, and in many other domains. For many outcomes, their predictive power rivals or exceeds that of cognitive skills.

Evidence from the General Educational Development (GED) testing programme in the United States shows the importance of non-cognitive skills. The GED is an achievement test which dropouts can take to certify that they are equivalent to secondary school graduates. The programme is based on the widely held belief that tests capture the important skills learned in school. On the surface, the programme is successful. Based on test scores, GED recipients are just as smart as high school graduates. When it comes to outcomes that matter, such as college completion and labour market success, GED recipients perform much worse in the labour market and in a variety of other life domains than traditional secondary school graduates. Achievement tests like the GED do not adequately capture valuable non-cognitive skills. This evidence should cause policymakers to think twice about relying on achievement tests to evaluate the effectiveness of educational systems.

Reliable measures of non-cognitive skills are available, and they are discussed in this paper. In developing any measure of non-cognitive skills, it is essential to recognise that all measures of skill are based on performance of some task. Traditional personality tests are based on the performance of the task of self-description. Performance on any task depends on multiple skills as well as the effort expended on it. Effort, in turn, depends on the incentives offered to exert the effort to perform the task. Since all measures of cognitive and non-cognitive skill are measures of performance on some task, it is necessary to standardise for incentives, effort, and other skills that determine performance on the task in measuring any skill, yet this is rarely done in conducting skill assessments. Standard measures of cognitive skill have been shown to be sensitive to incentives and levels of other skills. Test scores for young children can be improved by one standard deviation by offering candy for correct answers. The responsiveness to incentives in turn depends on a child's levels of conscientiousness. Using measured behaviours to capture non-cognitive skill is a promising approach that has been shown to be empirically effective. Such measures are available in administrative data that are collected routinely by schools and government agencies.

Skills are stable across situations with different incentives, although manifestations of skills vary with incentives. Though stable at any age, skills are not immutable traits that are set in stone over the life cycle. They have a genetic basis but are also shaped by environments, including families, schools, and peers. Skill development is a dynamic process. The early years are important in shaping all skills and in laying the foundations for successful investment and intervention in the later years. During the early years, both cognitive and non-cognitive skills are highly malleable. During the adolescent years, non-cognitive skills are more malleable than cognitive skills. The differential plasticity of different skills by age has important implications for the design of effective policies.

---

[1] This paper draws on Heckman and Kautz (2014a) and Heckman and Kautz (2014b). Kautz: Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago IL 60637. Tel: 773-702-0634. Email: tkautz@uchicago.edu. Heckman: Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago IL 60637; American Bar Foundation, 750 North Lake Shore Drive, Chicago, IL 60611; Institute for Fiscal Studies, University College London. Tel: 773-702-0634; Fax: 773-702-8490. Email: jjh@uchicago.edu. Diris: Research Centre of Public Economics, KU Leuven, Naamsestraat 69 - Box 3565, 3000 Leuven Belgium. Ter Weel: CPB and Maastricht University, PO Box 80510, 2508 GM Den Haag, The Netherlands, b.ter.weel@cpb.nl. Borghans: Department of Economics and ROA, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands.

This paper reviews a variety of interventions targeted to different stages of the life cycle. We interpret all of the studies we examine within an economic model of skill development. While it is difficult to compare different interventions because they are often multifaceted and target different populations, nonetheless, four conclusions emerge.

First, the evidence base is larger on the long-term effectiveness of interventions that start in early childhood and elementary school compared to their adolescent counter-parts. Many evaluations of early programmes measure a diverse set of outcomes and have follow-ups lasting more than 20 years. Evidence on adolescent interventions is scarcer. Follow-ups for them are typically shorter and fewer outcomes are analysed over shorter horizons. For this reason, we can draw stronger conclusions about the long-run efficacy of early programmes and how they work.

Second, when evaluating skill enhancement programmes it is vital to consider outcomes other than IQ or achievement test scores. Only interventions that start long before kindergarten begins have been shown to have long-term effects on IQ. If IQ were the only measure of success, most intervention programmes would seem futile. Using a diverse set of outcomes presents a more optimistic point of view. Many early programmes improve later-life outcomes, even though they do not improve IQ.

These programmes work because they foster non-cognitive skills. Some have annual rates of return that are comparable to those from investments in the stock market. Parental involvement is an important component of successful early interventions just as successful adolescent mentoring is an age-appropriate version of parental involvement.

Third, the available evidence suggests that the most successful adolescent remediation programmes are not as effective as the most successful early childhood and elementary school programmes, although adolescent mentoring and the provision of information can be very effective. Building an early base of skills that promote later-life learning and engagement in school and society is often a better strategy than waiting for problems to occur. Prevention is more effective than remediation if at-risk populations are sufficiently well targeted.

Fourth, adolescent remediation is possible for children who grow up in disadvantaged environments in their early years. The available evidence suggests that the most promising adolescent interventions are those that target non-cognitive skills as well as programmes that offer mentoring, guidance and information. Many adolescent programmes that focus on academic skills or temporarily change a participant's environment are only successful in the short run although the short-term results can often appear to be spectacular. Workplace-based programmes that teach non-cognitive skills appear to be effective remedial interventions for adolescents. They motivate acquisition of work-relevant skills and provide for disadvantaged youth the discipline and guidance which is often missing in their homes or high schools. Successful interventions at any age emulate the mentoring and attachment that successful families give their children.

Skills enable people. They are capacities to function. Greater levels of skill foster social inclusion and promote economic and social mobility. They generate economic productivity and create social well-being. Skills give agency to people to shape their lives, to create new skills and to flourish.

INTRODUCTION

Modern societies rely on written tests to sift and sort people, to evaluate students and schools, and to assess the performance of entire nations.[2] Achievement tests play a prominent role. The OECD actively promotes PISA tests. In the US, high school dropouts can take a 7-and-a-half hour achievement test – the General Educational Development (GED) exam – to certify that they are equivalent to high school graduates.[3]

Despite their widespread use, achievement tests are not well understood and their creators were well aware of their limitations (see the discussion in Heckman et al., 2014a). Achievement tests were developed in the mid-twentieth century as a way to measure a new concept – "general knowledge" – in an attempt to measure skills that are useful inside and outside of the classroom.[4] Their developers thought that they had designed pencil-and-paper tests that would predict success in the labour market, in education, and in many other aspects of life.

Evaluation of the benefits of these tests is circular. Achievement tests are typically validated using IQ tests and grades, and not by their ability to predict important life outcomes. Fortunately, a recent literature has conducted more meaningful evaluations of these tests and we report its findings here.

Achievement test scores predict only a small fraction of the variance in later-life success. For example, adolescent achievement test scores only explain at most 17% of the variability in later-life earnings. [5] [6] Measurement error accounts for at most 30% of the remaining variability (see Bound et al., 2001).

Achievement tests do not adequately capture non-cognitive skills such as perseverance ("grit"), conscientiousness, self-control, trust, attentiveness, self-esteem and self-efficacy, resilience to adversity, openness to experience, empathy, humility, tolerance of diverse opinions and the ability to engage productively in society, which are valued in the labour market, in school, and in society at large. Until recently these skills have largely been ignored in evaluations of schools and interventions. However, in recent research economists and psychologists have constructed measures of these skills and provided evidence that they are stable across situations and predict meaningful life outcomes.[7]

---

[2] The Programme for International Student Assessment (PISA) evaluates student performance in maths, science, and reading across countries, and its results attract a lot of media attention and influence policy. Scores from the year 2000 PISA test led Germany to reevaluate its educational system and introduce a variety of reforms (Grek, 2009). The creators of the original PISA tests called them literacy tests, not achievement tests, because PISA was designed to capture how knowledge can be applied to other contexts (OECD, 2014). However, this was also the goal of the original achievement tests. We are unable to find any studies establishing that the original PISA measures were fundamentally different skills from those measured by achievement tests. Recently, however, PISA 2012 has added some component tests designed to capture aspects of non-cognitive skills including openness, locus of control, and motivation (OECD, 2013).

[3] See Heckman et al. (2014a) for a detailed discussion of the GED program and an evaluation of its benefits.

[4] For histories of achievement tests see Heckman and Kautz (2014c); Quinn (2014).

[5] See Heckman and Kautz (2012). IQ tests alone explain at most 7% of this variability.

[6] See Bound, Brown, and Mathiowetz (2001).

[7] See the studies by Borghans, Duckworth, Heckman, and ter Weel (2008) and Almlund, Duckworth, Heckman, and Kautz (2011). The modern literature traces back to Bowles and Gintis (1976), and Bowles, Gintis, and Osborne (2001). An important study in sociology is the work of Peter Mueser reported in Jencks (1979).

Skills are not traits set in stone at birth and determined solely by genes. They can be fostered. Cognitive and non-cognitive skills change with age and with instruction. Interventions to improve skills are effective to different degrees for different skills at different ages. Importantly, non-cognitive skills are more malleable at later ages than cognitive skills.

This paper reviews recent evidence from economics and personality psychology on the predictive power of cognitive and non-cognitive skills and how to foster them. A growing body of empirical research shows that non-cognitive skills rival IQ in predicting educational attainment, labour market success, health, and criminality.[8] Both IQ and non-cognitive skills predict scores on achievement tests but non-cognitive skills predict outcomes above and beyond their effects in predicting scores on achievement tests.[9] Non-cognitive skills are universally valued across all cultures, religions and societies.

Because both cognitive and non-cognitive skills can be shaped and changed over the life cycle, they are properly called skills. An older terminology in psychology refers to them as "traits," conveying a sense of immutability or permanence, possibly due to their heritable nature. The distinction between skills and traits is not just a matter of semantics. It suggests new and productive avenues for public policy.

Skills enable people. They are capacities to function. Greater levels of skills foster social inclusion and promote economic and social mobility, economic productivity and well-being. Skills give people the tools with which to shape their lives, to create new skills and to flourish.

There are reliable ways to measure skills, and there are proven ways to enhance them and to evaluate efforts to foster them. Recent research establishes the existence of critical and sensitive periods in the formation of skills over the life cycle. Sensitive periods are those periods where investment is especially productive; critical periods are those periods when investment is essential. Critical and sensitive periods differ across skills. Investments should target those periods.

Effective strategies for promoting human development are based on three guiding principles:

1. The powerful role of families and general social environments in shaping skills.

2. The multiplicity of skills required for successful functioning in society. A core set of skills promotes success in many aspects of life. Different tasks require different skills in different levels and proportions. Successful people pursue the tasks where their skills give them comparative advantage along with personal satisfaction.

3. The technology of skill formation: that skills together with investments beget further skills (Cunha and Heckman, 2007, 2008; Cunha, Heckman, and Schennach, 2010).

Effective policies to promote skills straddle the missions of cabinet agencies and draw on the wisdom of many academic disciplines. They require broad thinking and recognition that both cognitive and non-cognitive skills are important ingredients of successful lives and are malleable to different degrees at different stages of the life cycle. They recognise that different skills cross-fertilise each other. Focusing on

---

Work in psychology going back to Terman, Baldwin, Bronson, DeVoss, Fuller, Lee Kelley, Lima, Marshall, Moore, Raubenheimer, Ruch, Willoughby, Benson Wyman, and Hazeltine Yates (1925) shows that personality skills predict life outcomes (see also Murray, 1938; Terman, Oden, Bayley, Marshall, McNemar, and Sullivan, 1947; and the analysis in Gensowski, 2014).

[8] See Heckman and Kautz (2012, 2014a, 2014c), Almlund et al. (2011), Borghans et al. (2008), and Roberts, Kuncel, Shiner, Caspi, and Goldberg (2007) for reviews.

[9] See Kautz and Zanoni (2014).

one dimension of human skills to the exclusion of other dimensions or on one stage of the life cycle to the exclusion of others misses fundamental aspects of human performance and development. Narrowly focused policies fail to capture synergisms in the expression and development of skills.

Policymakers in the OECD share a common desire to develop human potential. However, current discussions focus on promoting skills by improving schools. Within this very narrow purview, the success of schools is measured by scores on the exams of the students in them. This mindset is a consequence of very limited conceptualisations of human capabilities that assumes that achievement tests capture the important life skills. It misses important dimensions of human flourishing. While schools are important, they are far from being the principal source of the skills that matter.

Discarding Obsolete Notions and Embracing the Findings of Recent Research

In designing effective human development strategies, it is essential to discard obsolete views about the origin and malleability of "traits." What used to be regarded as traits fixed at conception are now understood to be skills that can be augmented through guidance and instruction. Raw intelligence is not fixed solely by parental genes, although heritability plays an important role in shaping it. It is boosted by quality parenting and by caring environments. It becomes solidified around the time of puberty.

Non-cognitive skills can also be shaped. Compared to raw intelligence, they are more malleable until later ages. Neuroscience shows that this malleability is associated with the slow development of the prefrontal cortex.[10]When attempted, adolescent remediation should focus on boosting non-cognitive skills by effective mentoring and guidance.

Skill development is a dynamic process. Boosting non-cognitive skills early in life increases the benefits of education later in life. More persistent students learn more. The levels of cognitive and non-cognitive skills at any age depend in part on levels of those skills at younger ages which depend on earlier investments.[11]

Inequality among families in parenting and lack of support given to children in schools are major contributors to inequality in adult skills. In light of this evidence, it is of great concern that so many children around the world are being born into disadvantage as measured by the quality of parenting and that the trend appears to be accelerating. Everywhere, the traditional family with its secure environment for rearing children is under challenge.[12] A major casualty of this trend is the quality of parenting available to disadvantaged children. It is unreasonable to expect schools to take on the huge burden of supplementing compromised family environments in addition to their traditional job of educating children.

Evidence from the social and biological sciences establishes the importance of the early years in fostering the skills that promote human flourishing. Families do much more than pass on their genes or put food on the table.

Human development is a dynamic process that starts in the womb. Environments and skills interact to foster the development of later life skills and create who we are and what we become. The foundations for adult success are laid down early in life. Many children raised in disadvantaged environments start behind and stay behind. Poverty has lasting effects on brain development, health, and cognitive and non-cognitive skills.[13]Gaps in skills emerge early, before formal schooling begins. Waiting until kindergarten to address these gaps is a poor strategy. It creates achievement gaps for disadvantaged children that are costly to close.

[10] Walsh (2005).

[11] Cunha and Heckman (2007, 2008); Cunha et al. (2010).

[12] See McLanahan (2004) and Heckman (2008) for American data and OECD (2011) for evidence around the world.

[13] See, e.g., the evidence in Shonkoff and Phillips (2000) and Heckman (2008).

Family disadvantage is poorly assessed by conventional measures of poverty that focus on family income, wealth, and parental education. The absence of parental guidance, nourishment, and encouragement is the most damaging condition for child development. Quality parenting – stimulation, attachment, encouragement, and support – is the true measure of child advantage, and not the traditional measures of poverty commonly used in policy discussions.

Children from disadvantaged homes are much more likely to drop out of high school. Single-parent families, compared to two-parent families, often give less cognitive and emotional stimulation to their children (e.g., the evidence in Moon, 2012, 2014a), and as a consequence their children have diminished cognitive and non-cognitive skills.[14] The evidence summarised in this paper suggests that skill deficits can be prevented by improving the early lives of disadvantaged children.

Yet, while important and often neglected in American public policy, the early years do not fully determine adult success. Children are resilient and quality parenting throughout childhood fosters the development and expression of skills. Schooling shapes both cognitive and non-cognitive skills. Certain adolescent remediation programmes appear to be effective.

A variety of adolescent interventions attempt to remediate early-life skill deficits. Alter- natives to the traditional high school curriculum provide potential dropouts with training suitable to their interests and skills. Such programmes can mould non-cognitive skills, even if it is not their primary goal. Other interventions attempt to directly remediate non-cognitive deficits after students drop out of school.

Heckman and Kautz (2014a, 2014c) provide a sober account of what is known. They compare different skill enhancement strategies and consider the features that make some programmes more successful than others. We summarise and update their analysis in this document.

Unfortunately, the field of human development is marred by overzealous advocates who claim miracle fixes from their favoured programmes. The advocates focus on one slice of the life cycle to the exclusion of others. Slogans often replace hard evidence. Most evaluations of interventions have only short-term follow-ups. Many differ in their measures of outcomes and skills and target different demographic groups, so it is difficult to compare alternative programmes. Despite these limitations, several studies suggest that interventions during the preschool years or in kindergarten improve non-cognitive skills in a lasting way, some with annual rates of return that are comparable to those from investments in the stock market in normal times.[15] Several adolescent interventions are promising, particularly those that combine education with work-related experience.[16]There is also some hint that programmes that provide information to adolescents can be effective.

The Organisation of the Rest of the Paper

The remainder of this paper is organised as follows. Section 2 defines cognitive and non-cognitive skills and discusses how they are measured. Section 3 summarises the evidence on the predictive power of non-cognitive skills. Section 4 presents an analysis of the skills needed in the workforce, as garnered from surveys of employers. Section 5 presents a framework for analysing investment in skills over the life cycle. Sections 6 – 11 review evidence about the efficacy of education, parental investment, and interventions in improving non-cognitive skills from preschool, elementary school, and adolescence. Section 12 is a conclusion of all the information presented.

[14] See, e.g., Carneiro and Heckman (2003), Cunha, Heckman, Lochner, and Masterov (2006), and Cunha and Heckman (2009).

[15] See Heckman, Moon, Pinto, Savelyev, and Yavitz (2010a). The rates of return incorporate the benefits to individuals and to society.

[16] See, e.g., Kemple and Willner (2008) and Roder and Elliot (2011).

MEASURES OF COGNITIVE AND NON-COGNITIVE SKILLS

Cognitive Skills

Measures of cognition have been developed and refined over the past century. Cognitive ability has multiple facets.[1] Psychologists distinguish between fluid intelligence (the rate at which people learn) and crystallised intelligence (acquired knowledge).[2] Achievement tests are designed to capture crystallised intelligence[3] whereas IQ tests like Raven's progressive matrices (1962) are designed to capture fluid intelligence.[4] [5]

This new understanding of cognition is not widely appreciated. Many use IQ tests, standardised achievement tests, and even grades as interchangeable measures of "cognitive ability" or intelligence.[6] Scores on IQ tests and standardised achievement tests are strongly correlated with each other and with grades.[7] However, these general indicators of "cognition" measure different skills and capture different facets of cognitive ability.[8] [9]In the following section we show that scores on these tests are also influenced by effort and non-cognitive skills.

Measuring Non-cognitive Skills

Throughout this paper we use the term non-cognitive skills to describe the personal at- tributes not thought to be measured by IQ tests or achievement tests. These attributes go by many names in the literature, including soft skills, personality traits, non-cognitive abilities, character skills, and socio-emotional skills. These different names connote different properties.[10] "Traits" suggests a sense of permanence and possibly also of heritability. "Skills" suggests that these attributes can be learned. In reality, the extent to which these personal attributes can change lies on a spectrum. Both cognitive and non-cognitive skills can change and be changed over the life cycle, but through different mechanisms and with

[1] See Carroll (1993) and Ackerman and Heggestad (1997) for discussions.

[2] See, e.g., Nisbett, Aronson, Blair, Dickens, Flynn, Halpern, and Turkheimer (2012).

[3] Roberts, Goff, Anjoul, Kyllonen, Pallier, and Stankov (2000).

[4] Raven, Raven, and Court (1988). The high correlation between scores on intelligence tests and scores onachievement tests is in part due to the fact that both require intelligence and knowledge. Fluid intelligence promotes the acquisition of crystallised intelligence. Common developmental factors affect the formation of both skills.

[5] Carroll (1993) and Ackerman and Heggestad (1997) discuss more disaggregated facets of cognitive ability.

[6] This practice is true even among leading professional psychologists. For example, all of these measures are assumed to capture intelligence in Flynn (2007), Nisbett (2009), and Nisbett et al. (2012).

[7] See Heckman and Kautz (2012).

[8] See Borghans, Golsteyn, Heckman, and Humphries (2011a).

[9] It is an irony of the testing literature that high school grades are more predictive of first-year college performance than SAT scores (Bowen, Chingos, and McPherson, 2009). The SAT and related tests were once thought to be a more objective measure of student quality than high school grades (Lemann, 1999).

[10] See Almlund et al. (2011) and Borghans et al. (2008) for comparisons of some of these different taxonomies.

different ease at different ages. We use the term skill throughout this paper because all attributes can be shaped.

Although non-cognitive skills are overlooked in most contemporary policy discussions and in economic models of choice behaviour, personality psychologists have studied these skills for the past century. Psychologists primarily measure non-cognitive skills by using self-reported surveys or observer reports. They have arrived at a relatively well-accepted taxonomy of non-cognitive skills called the Big Five, with the acronym OCEAN, which stands for: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Table 1 defines these skills and their multiple facets.[11]Some argue that the Big Five are the "longitude and latitude" of non-cognitive skills, by which all more narrowly defined skills may be categorised.[12]

While the Big Five measures are now widely used in psychology, there are several other taxonomies, including the Big Three, the MPQ, and the Big Nine. They are conceptually and empirically related to the Big Five.[13] Other taxonomies, including psychopathology as measured by the DSM-IV and measures of temperament, have also been related to the Big Five.[14]Almlund et al. (2011) and Becker, Deckers, Dohmen, Falk, and Kosse (2012) summarise evidence showing that economic preference parameters are not closely related to the Big Five measures and apparently represent different attributes. Preference parameters along with the non-cognitive skills measured by psychologists govern behaviour.[15]

[11] See, e.g., Borghans et al. (2008).

[12] Costa and McCrae (1992a).

[13] See Borghans et al. (2008) and Almlund et al. (2011) for comparisons of these taxonomies.

[14] See, e.g., Cloninger, Svrakic, Bayon, and Przybeck (1999).

[15] A deeper question, as yet not systematically investigated in the literature in economics or psychology, is whether the "traits" captured by the alternative measurement systems are the expression of a more basic set of preferences or goals. McAdams (2006) adds goals to the list of possible traits. Almlund et al. (2011) and Heckman and Kautz (2012) develop a model in which preferences and endowments of skills determine the effort applied to tasks. As shown in the next section, performance on tasks is the source of any measurement of a trait. Hence, in their framework, measures of skills are determined, in part, by preferences.

Table 1. The Big Five Domains and Their Facets

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

Note: *These temperament attributes may be related to two Big Five factors. Facets specified by the NEO-PI-R personality inventory (Costa and McCrae, 1992b). Adjectives in parentheses from the Adjective Check List (Gough and Heilbrun, 1983).

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

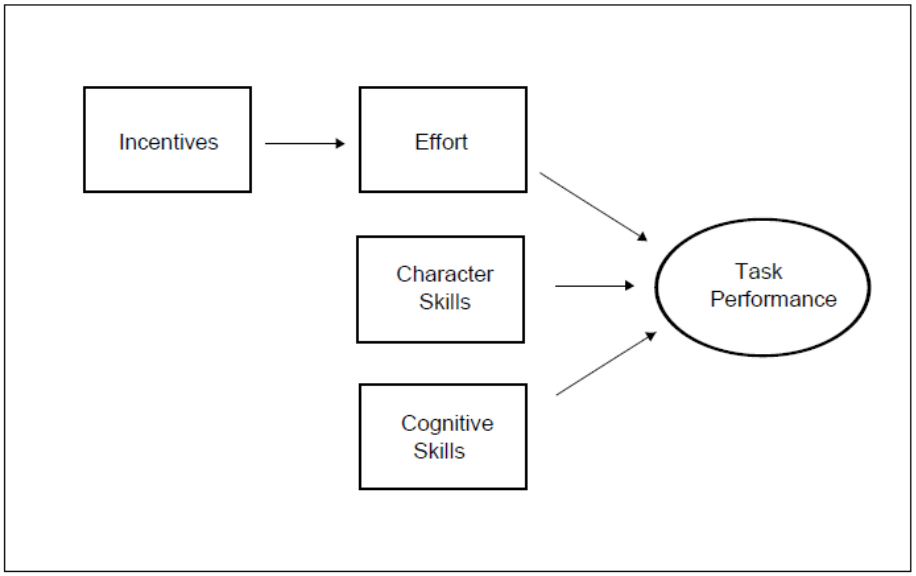
A Task-Based Framework for Identifying and Measuring Skills

A leading personality psychologist defines personality (non-cognitive) traits (skills) as follows:

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

Roberts' definition of personality ("non-cognitive" skills) and the one favoured by Almlund et al. (2011) suggests that all psychological measurements are calibrated on measured behaviour or "tasks" broadly defined. A task could be taking an IQ test, answering a personality questionnaire, performing a job, attending school, completing secondary school, participating in crime, or performing in an experiment run by a social scientist. Figure 1 depicts how performance on a task can depend on incentives, effort, and cognitive and non-cognitive skills. Performance on different tasks depends on these components to different degrees. People can compensate for their shortfalls in one dimension by having strengths in other dimensions.

Figure 1. Determinants of Task Performance



Many believe that personality skills can only be assessed by self-reported questionnaires that elicit skills like the Big Five (see John, 2000; John and Srivastava, 1999 for one widely used questionnaire). However, performance on any task or any observed behaviour can be used to measure personality and other skills.[16] For example, completing high school requires many other skills besides those measured by achievement tests, including showing up in school, paying attention, and behaving in class.[17]

[16] See Almlund et al. (2011).

[17] The idea of using behaviours to measure non-cognitive is old. Ralph Tyler suggested using measures of behaviour to capture non-cognitive skills in his first proposal for the National Assessment of Educational Progress tests. See Tyler (1973) and Rothstein, Jacobsen, and Wilder (2008). This idea is being been pursued in the recent literature (Heckman, Humphries, Urzu´a, and Veramendi, 2011; Jackson, 2013). See Kautz and Zanoni (2014) for a recent application of this idea. We discuss this approach more extensively in Section 2.5.

Inferring skills from performance on tasks requires standardising all of the other contributing factors that produce the observed behaviours. The inability to parse and localise behaviours that depend on a single skill or ability gives rise to a fundamental problem of assessing the contribution of any particular skill to the successful performance on any task (or measure). This problem is commonly ignored in empirical research that studies how cognitive and non-cognitive skills affect outcomes.[18]

There are two distinct issues that need to be addressed in designing measures of skills based on performance of any task. First, behaviour depends on incentives created by situations. Different incentives elicit different amounts of effort on the tasks used to measure skills. Accurately measuring non-cognitive skills requires standardising for the effort applied in any task. Second, performance on most tasks depends on multiple skills. Not standardising for incentives and other relevant skills that determine performance on a particular task used to measure a particular skill can produce misleading estimates of that particular skill.

These issues are empirically relevant. For example, incentives partly determine scores on IQ tests. Studies conducted over the past 40 years show that incentives, like money or candy, can increase IQ scores, particularly among low-IQ individuals. The black – white gap in IQ can be completely eliminated by giving M&M candies for correct answers.[19] However, there is no evidence that this incentive-induced performance persists. It has yet to be shown that creating incentives for performance on one test improves performance on subsequent tests, or in any other life task. Indeed, there is some evidence that such incentives, in fact, may worsen subsequent performance (Deci and Ryan, 1985; Ryan and Deci, 2000).

Not all persons respond with equal strength to incentives. Research by Borghans et al. (2008) and Segal (2012) shows that the responsiveness of persons to incentives on IQ and achievement tests depends on their non-cognitive skills. Duckworth, Quinn, Lynam, Loeber, and Stouthamer-Loeber (2011) show motivation of test takers predicts IQ scores.

The recent literature shows that non-cognitive skills predict standardised achievement test scores, which some psychologists assume are good measures of intelligence.[20] Figure 2 (based on Dutch data) shows how the variability across persons in the scores on one achievement test, the Differential Aptitudes Test (DAT),[21] are determined by IQ and non-cognitive measures. Non-cognitive skills explain a substantial portion of the variability across per- sons in DAT scores. Non-cognitive skills explain the variance in achievement scores above and beyond the variance that IQ explains when both measures of non-cognitive skill and IQ are included in a regression. These findings caution the interpretation that standardised achievement tests only measure cognitive ability. They also capture non-cognitive skills.[22]

Ironically, The Bell Curve by Herrnstein and Murray, which uses an achievement test (the Armed Forces Qualification Test, AFQT) as a measure of intelligence, implicitly shows the power of both cognitive and non-cognitive in shaping life outcomes in the United States.

---

[18] See Borghans et al. (2011a), Almlund et al. (2011), and Heckman and Kautz (2012) for discussions of this problem.

[19] See Ayllon and Kelly (1972); Borghans, Meijers, and ter Weel (2008); Breuning and Zella (1978); Clingman and Fowler (1976); Edlund (1972); Holt and Hobbs (1979); Larson, Saccuzzo, and Brown (1994); Segal (2008). This evidence is summarised in Borghans et al. (2008) and Almlund et al. (2011).

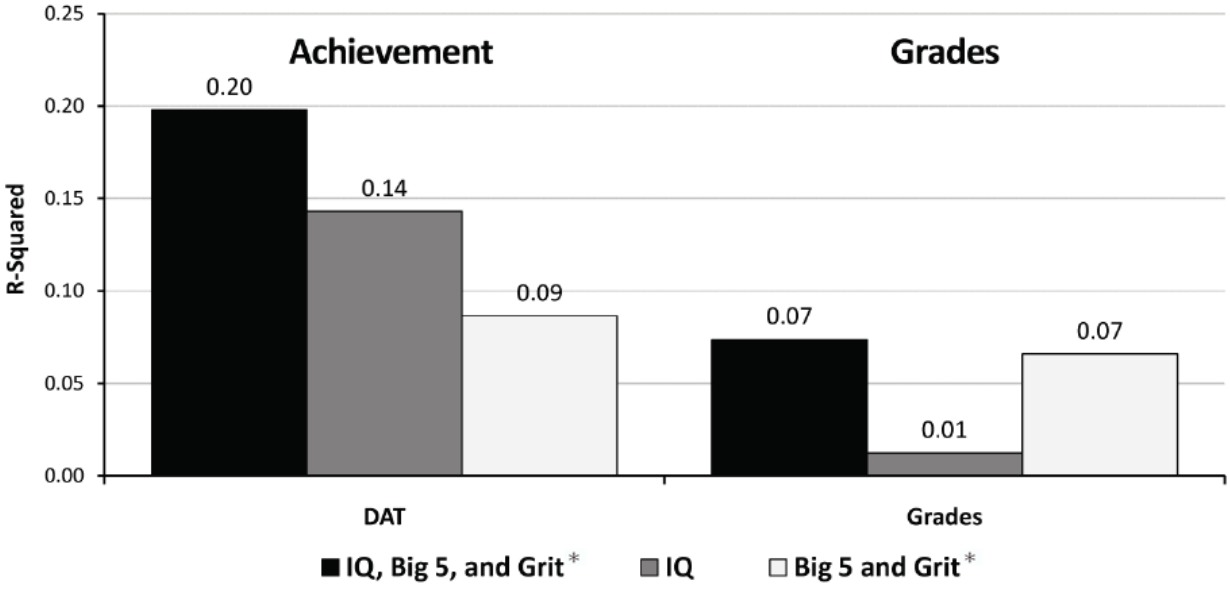
[20] See, e.g., Nisbett (2009).

[21] The correlation between DAT and the widely used Armed Forces Qualification Test (AFQT) scores in the National Longitudinal Study of Youth 1979 (NLSY79) is 0.75 (Borghans, Golsteyn, Heckman, and Humphries, 2011b). Friedman and Streicher (1985) estimate correlations between 0.65 and 0.82 in a sample of high school sophomores and juniors. Kettner (1976) estimates correlations between DAT and the AFQT subtests of 0.76 to 0.89 in a sample of juniors and seniors.

[22] In the Stella Maris data, Openness to Experience is strongly correlated with IQ. See Borghans et al. (2011b).

However, measures of non-cognitive skills beyond achievement tests are predictive in explaining outcomes beyond the predictions based on achievement tests (see Heckman, Humphries, and Veramendi, 2014; Kautz and Zanoni, 2014).

Figure 2. Decomposing Variance Explained for Achievement Tests and Grades into IQ and Non-Cognitive Skills: Stella Maris Secondary School, Maastricht, Holland



Note: Grit is a measure of persistence on tasks (Duckworth, Peterson, Matthews, and Kelly, 2007).

Source: Borghans et al. (2011a).

Reference Bias

Answers from self-reports can be misleading when comparing levels of personality skills across different groups of people. Most personality assessments do not anchor their measurements in any objective outcome.[23] For example, the German Socio-Economic Panel (GSOEP) asks respondents to rate themselves on the following statement: "I see myself as someone who tends to be lazy" (Lang, John, Lüdtke, Schupp, and Wagner, 2011). The scale ranges from 1 = "strongly disagree" to 7 = "strongly agree." In answering this question, people must interpret the definition of "lazy," which likely involves comparing themselves to other people. If different groups have different standards or reference points, comparing traits across groups can be highly misleading. Laziness may mean different things to different groups of people.

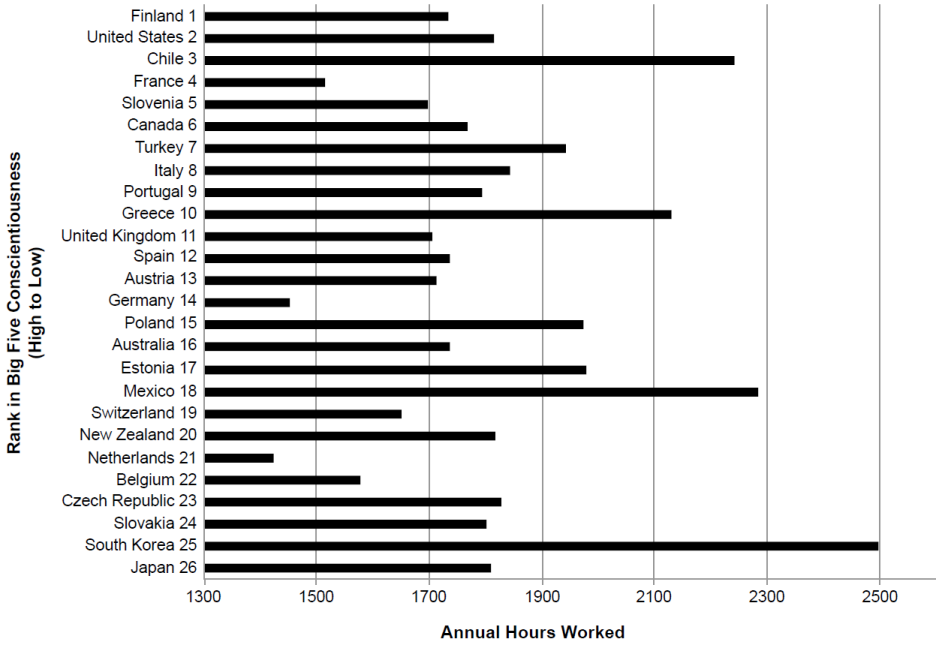
This measurement problem – sometimes called reference bias – is empirically important.[24] Schmitt, Allik, McCrae, and Benet-Martínez (2007) administer a Big Five personality questionnaire to groups of people in a variety of different countries. Using their estimates, Figure 3 shows how Organisation of Economic Cooperation of Development (OECD) countries rank (from high to low) in Conscientiousness – the tendency to be hard-working and persistent. The bars display the average number of hours that people work in the country.

[23] These are called Likert scales (Likert, 1932).

[24] Reference bias is also problematic in health surveys that use self-reported, subjective health assessments. See Groot (2000).

At first glance the results are surprising. South Korea ranks second to last in terms of Conscientiousness but also ranks first in the number of hours worked. South Korea is not an anomaly. Country-level reports of Big Five Conscientiousness are unrelated to the number of hours worked. The rank correlation between hours worked and conscientiousness across countries is negative, though statistically insignificant.[25] This finding contrasts with studies showing that non-cognitive skills tend to be positively related with labour supply within individual countries.[26]

Figure 3. National Rank in Big Five Conscientiousness and Average Annual Hours Worked



Note: Several countries are omitted due to lack of data.

Sources: The Conscientiousness ranks come from Schmitt et al. (2007). These measures were taken in 2001 (Schmitt, 2002). The hours worked estimates come from the OECD (2001).

These findings are relevant for constructing measures of non-cognitive skills. Granting that achievement tests miss important skills, would the self-reported Big Five be a useful supplement to school evaluations? The possibility of reference bias suggests that it might not. Self-reports of Conscientiousness might measure different things for different schools.

Psychologists have attempted to address this problem.[27] Some surveys include vignette-based questions that attempt to standardise for aspects of the culture or situation. They attempt to frame questions so that the people in the survey answer within a common situation. However, this approach might not work well for evaluating schools, especially if teachers have incentives to coach children on taking these tests so that they score better and give answers perceived to be positive. Direct use of standard psychological

[25] $r = -0.07$ ($p = 0.73$).

[26] See the studies summarised in Almlund et al. (2011).

[27] For further discussion of reference bias, see Duckworth (2012); Goldammer (2010); Heine, Buchtel, and Norenzayan (2008); Heine, Kitayama, Lehman, Takata, Ide, Leung, and Matsumoto (2001); Heine, Lehman, Peng, and Greenholtz (2002); Naumann and John (2011); Peng, Nisbett, and Wong (1997); and Schmitt et al. (2007).

measures can be problematic.[28] We consider next a different strategy based on observations of behaviour and not on responses to questionnaires administered by psychologists.

Measuring Skills Using Behaviours

Ralph Tyler, one of the two scholars credited with developing achievement tests, recognised their limitations.[29] He suggested using measures of behaviour such as performance, participation in student activities, and other observations by teachers and school administrators to complement achievement tests when evaluating students and schools. Several recent papers demonstrate that this is a promising approach. Heckman, Pinto, and Savelyev (2013) show that teacher ratings of elementary schoolchildren's behaviours are strong predictors of adult outcomes and that early childhood interventions promote the non-cognitive skills measured by these ratings. Heckman et al. (2014) estimate the causal effect of cognitive and socio-emotional (non-cognitive) skills on a variety of outcomes. They measure socio-emotional (non-cognitive) skills using risky and reckless behaviours measured in the adolescent years.[30]

They develop and apply methods to use high school grades to measure both cognitive and non-cognitive skills. They show that non-cognitive skills promote educational attainment, beneficial labour market outcomes, and health.

Jackson (2013) studies the effect of teachers on student cognitive and non-cognitive skills. In a fashion similar to Heckman et al. (2011, revised 2014), Jackson measures cognitive skills using achievement test scores, while measuring non-cognitive skills using absences, suspensions, grades, and grade progression. These measures of non-cognitive skills predict adult outcomes with a strength similar to measures of cognitive ability. His measures of non-cognitive skills are commonly available from the administrative records of schools. Kautz and Zanoni (2014) use early measures of behaviour in school to predict graduation and college attendance.

Similar to Ralph Tyler's suggestion of using participation in extracurricular activities to measure non-cognitive skills, Lleras (2008) measures non-cognitive skills in part by using tenth grade participation in sports, academic clubs, and fine arts activities. Participation in these activities predicts educational attainment 10 years later, even after controlling for cognitive ability as measured by achievement tests.

Criminologists have debated about whether it is better to measure self-control with self- reported psychological scales or with objective behaviours. Gottfredson and Hirschi (1990) launched a series of studies on the link between self-control and crime. They claim that a single factor, self-control, predicts much of the variance in criminal outcomes.

There is a sharp divide in this literature. Subsequent studies have measured self-control using psychological scales, while others have used behavioural measures. A meta-analysis by Pratt and Cullen (2000) finds that behavioural measures are at least as good at predicting crime as are measures based on self-reported taxonomies. In a similar vein, Benda (2005) uses both types of measures in the same study and finds that behavioural measures predict crime better than psychological scales.

[28] In an attempt to address reference bias, some psychologists measure skills using behaviours. Heine et al. (2008) examine cross-country differences in Conscientiousness using objective measures, including walking speed, postal workers' speed, and the accuracy of clocks in public banks. To measure walking speed, researchers timed how long it took for a random sample of people to walk 60 feet in public areas. Postal workers' speed was assessed by measuring how long it took for postal workers to sell stamps.

[29] See Heckman and Kautz (2014c).

[30] The measure of risky and reckless behaviour is based on whether adolescents engaged in any of the following behaviours: stealing from a store, purposefully damaging property, taking something worth less than $50, or conning someone.

Hirschi and Gottfredson (1993) suggest that objective behavioural measures might be preferred to self-reports, partly because the act of filling out a survey requires some level of self-control. Answering survey questions is in itself a task that relies on skills beyond the ones targeted by the survey.

Some criticise this approach and argue that it is tautological to use measures of behaviour to predict other behaviours even though the measures are taken early in life to predict later life behaviours.[31] As suggested by Figure 1, all tasks or behaviours can be used to infer a skill as long as the measurement accounts for other skills and aspects of the situation. In addition, many of the recent studies in economics use early measures of behaviours to predict behaviours in adulthood. (See, e.g., Heckman et al., 2014, 2013). Self-reported scales should not be assumed to be more reliable than behaviours, although personality psychologists often assume so. The question is which measurements are most predictive and which can be implemented in practice. The literature suggests that there are objective measurements of non-cognitive skills that are not plagued by reference bias.

Are Non-Cognitive Skills Stable?

Many have questioned whether there are stable non-cognitive skills, i.e., whether people exhibit the same non-cognitive skills across different situations at a fixed point in time. The publication of Walter Mischel's 1968 book, Personality and Assessment, gave rise to a heated "personality–situation" debate within psychology, which pitted social psychologists who favoured situational factors as primary determinants of behaviour against personality psychologists who considered stable personality (non-cognitive) traits (skills) as more con- sequential. Mischel argued that aspects of situations overshadow any effect of personality (non-cognitive) traits (skills) on behaviour. Ironically, Mischel himself later demonstrated the stability of non-cognitive skills (as measured by the performance of subjects in demonstrating self-control in early childhood) over the life cycle in his celebrated "marshmallow experiment."[32] Oblivious to this evidence, behavioural economists continue to echo Mischel's 1968 claim. (See, e.g., Thaler, Mullainathan, and Kahneman, 2008).

A large body of evidence reviewed in Almlund et al. (2011) shows that stable non- cognitive skills exist and are predictive of many behaviours.[33] An early paper by Epstein (1979) presents compellingevidence that, averaging over tasks and situations, people act in a predictable fashion with a high level of reliability of average behaviour ("measured non- cognitive skills") across situations.[34]

In addition, measures of non-cognitive skills tend to be about 30%–60% heritable, suggesting that something tied to the person, rather than the situation, influences behaviour (Bouchard and Loehlin, 2001).[35] Evidence from neuroscience suggests that expression of different skills is linked to different regions of the brain (see Canli, 2006, and DeYoung, Hirsh, Shane, Papademetris, Rajeevan, and Gray, 2010). Non-cognitive skills are not epiphenomena.

---

[31] See the discussion in Pratt and Cullen (2000) and Benda (2005).

[32] A participant (usually a child) was given a marshmallow. The experimenter left the room and told the participant that he or she would receive a second marshmallow if he or she resisted consuming the marshmallow until the experimenter returns. The length of time that the participant waits is a measure of short-term discounting. The children who could wait had much better lifetime outcomes. (For a recent discussion of this study, see Mischel, Ayduk, Berman, Casey, Gotlib, Jonides, Kross, Teslovich, Wilson, Zayas, and Shoda, 2011.)

[33] See the special issue of Journal of Research in Personality (2009) entitled "Personality and Assessment at Age 40" for a recent discussion.

[34] R2 of 0.6–0.8, where R2 is a measure of variance explained.

[35] 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 (2006) and an emerging literature on epigenetics (Jablonka and Raz, 2009).

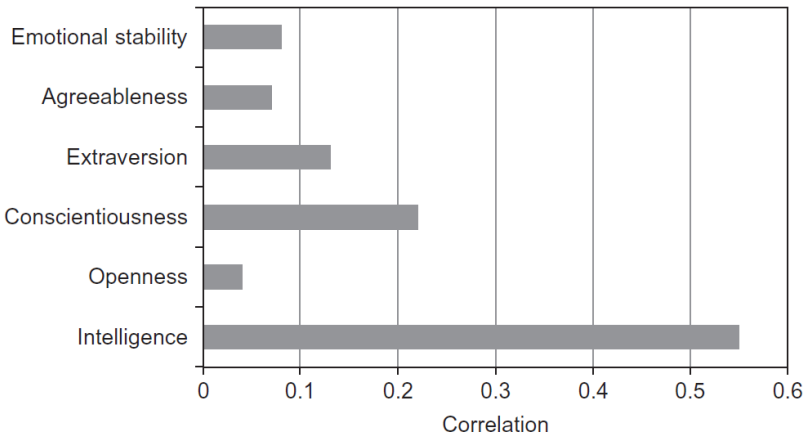
THE PREDICTIVE POWER OF NON-COGNITIVE SKILLS

Correlational Evidence

A substantial body of evidence shows that non-cognitive skills predict a wide range of life outcomes, including educational achievement, labour market outcomes, health, and criminality. For many outcomes, the predictive power of non-cognitive skills rivals that of measures of cognitive ability. Of the Big Five, Conscientiousness – the tendency to be organised, responsible, and hardworking—is the most widely predictive across a variety of outcomes. (See Almlund et al., 2011; Borghans et al., 2008; Heckman and Kautz, 2012; Roberts et al., 2007). Conscientiousness predicts years of schooling with the same strength as measures of intelligence (Almlund et al., 2011).

Aspects of job performance are also related to academic performance. Both require completing work on a schedule and involve intelligence to varying degrees. As with academic performance, numerous studies and meta-analyses have found that Conscientiousness is associated with job performance and wages (Barrick and Mount, 1991; Hogan and Holland, 2003; Nyhus and Pons, 2005; Salgado, 1997). Figure 4 presents correlations of the Big Five and IQ with job performance. Of the Big Five factors, Conscientiousness is the most strongly associated with job performance but is about half as predictive as IQ. Conscientiousness, however, may play a more ubiquitous role than IQ. The importance of IQ increases with job complexity (the information processing requirements of the job). Cognitive skills are more important for professors, scientists, and senior managers than for semiskilled or unskilled labourers (Schmidt and Hunter, 2004). The importance of Conscientiousness does not vary much with job complexity (Barrick and Mount, 1991); this suggests that it applies to a wider spectrum of jobs.

Figure 4. Associations with Job Performance

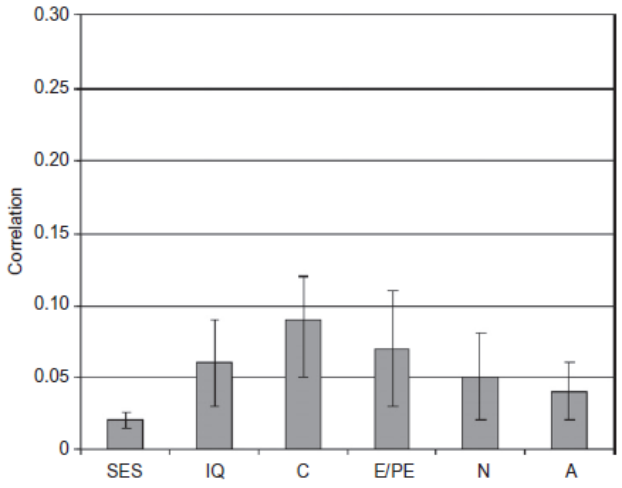


Note: The values for personality are correlations that were corrected for sampling error, censoring, and measurement error. Job performance was based on performance ratings, productivity data, and training proficiency. The authors do report the timing of the measurements of personality relative to job 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 intelligence is a raw correlation.

Sources: The correlations reported for personality traits come from a meta-analysis conducted by Barrick and Mount (1991). The correlation reported for intelligence comes from Schmidt and Hunter (2004).

Measures of non-cognitive skills rival IQ and measures of socioeconomic status in predicting longevity.[1] Roberts et al. (2007) review evidence from 34 different studies on the predictive validity of the Big Five personality measures, relative to that of cognitive ability and socioeconomic status. Most studies they survey 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. Figure 5 presents results from their analyses. Conscientiousness is a stronger predictor of longevity than any other Big Five trait and a stronger predictor than either IQ or socioeconomic status.[2] In general, skills related to Conscientiousness, Openness to Experience, and Agreeableness are associated with longer lives.[3] The magnitudes of the relationships vary across studies.

Figure 5. Correlations of Mortality with Non-Cognitive Skills, IQ, and Socioeconomic Status (SES)



Note: 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 represents standard errors. The length of the studies represented vary from 1 year to 71 years.

Source: Roberts et al. (2007).

Of the Big Five, Conscientiousness and Agreeableness are most predictive of criminality. In a sample of at-risk youth, boys who committed severe delinquent behaviours were more than three quarters of a standard deviation lower in Agreeableness and Conscientiousness, as measured by mothers' reports at age 12 or 13, than boys who had minor or no delinquent behaviours up to that age (John, Caspi, Robins, and Moffitt, 1994).

As with most studies in personality psychology, the evidence presented in Figure 4-5 and most of the literature do not address the question of causality; that is do measured skills *cause* (rather than just predict outcomes)? Empirical associations are not a reliable basis for policy analysis. As previously noted (see Figure 1), multiple skills and effort all generate performance in a given task. Many studies in personality psychology do not control for all of the factors that produce performance on measured tasks. They equate

[1] For a recent study, see Friedman and Martin (2011).

[2] The timing of the measurements of non-cognitive skill relative to the outcomes varies by study.

[3] See Boyle, Williams, Mark, Brummett, Siegler, and Barefoot (2005); Friedman and Martin (2011); Kern and Friedman (2008); Kubzansky, Sparrow, Vokonas, and Kawachi (2001); Martin, Friedman, and Schwartz (2007); Mroczek and Spiro (2007); Schulz, Bookwala, Knapp, Scheier, and Williamson (1996) .

measures of outcomes with the skill being measured.[4] This practice can lead to a substantial bias in inference about the importance of any particular skill. Our discussion of the GED programme and survey of the intervention literature in the following sections presents evidence on the causal relationship between skills and outcomes.

Evidence from the General Educational Development (GED) Programme

Evaluations of the General Educational Development (GED) programme in the United States provide causal evidence on the importance of non-cognitive skills. The GED is an extreme application of the philosophy of achievement testing. It is a seven-and-a-half hour achievement test that secondary school dropouts can take to certify to employers and schools that they have the knowledge and skills of secondary school graduates.

On the surface, the GED programme appears to be successful. Figure 6 shows the distribution of cognitive skill for secondary school graduates, GED recipients, and other secondary school dropouts. As measured by scores on other achievement tests, GED recipients are as smart as secondary school graduates who do not go on to college.

However, test scores are not intrinsically valuable. On outcomes that matter, GED recipients are not equivalent to traditional graduates. Figure 7 shows the wages of GED recipients and secondary school graduates relative to those of dropouts both in the raw data and when adjusted for cognitive ability. At each age, the first set of bars shows the effect when adjusting for basic demographics (age, race, and region of residence). The second set shows the effect after additionally adjusting for cognitive skill as measured by scores on the Armed Forces Qualification Test (AFQT). After adjusting for cognitive skill, GED recipients are indistinguishable from dropouts, whereas secondary school graduates outperform both. Similar patterns arise across a wide range of outcomes. Compared to traditional secondary school graduates, GED recipients have shorter spells of employment, get divorced at higher rates, have worse health, and are more likely to become incarcerated.[5]

If GED recipients are as smart as secondary school graduates, why do they perform worse in the labour market after adjusting for cognitive skill? The GED test misses non-cognitive skills. Figure 8 shows the distribution of non-cognitive skills as measured by risky behaviours during adolescence. By this measure, GED recipients are much more like other dropouts than secondary school graduates. Deficits in non-cognitive skills cause GED recipients to dropout of high school and underperform later in life.[6] Success in high school and in the labour market requires showing up on time, persistence, and organisation.

Evidence from the GED programme teaches broader lessons about the pitfalls of relying on achievement tests. Faith in tests can conceal major social problems. Heckman and LaFontaine (2010) show that if GED recipients are counted as high school dropouts, then the black-white gap in high school graduation for males has declined substantially over the last 30 years. If GED recipients are not counted as high school graduates, there has been virtually no change in the gap.

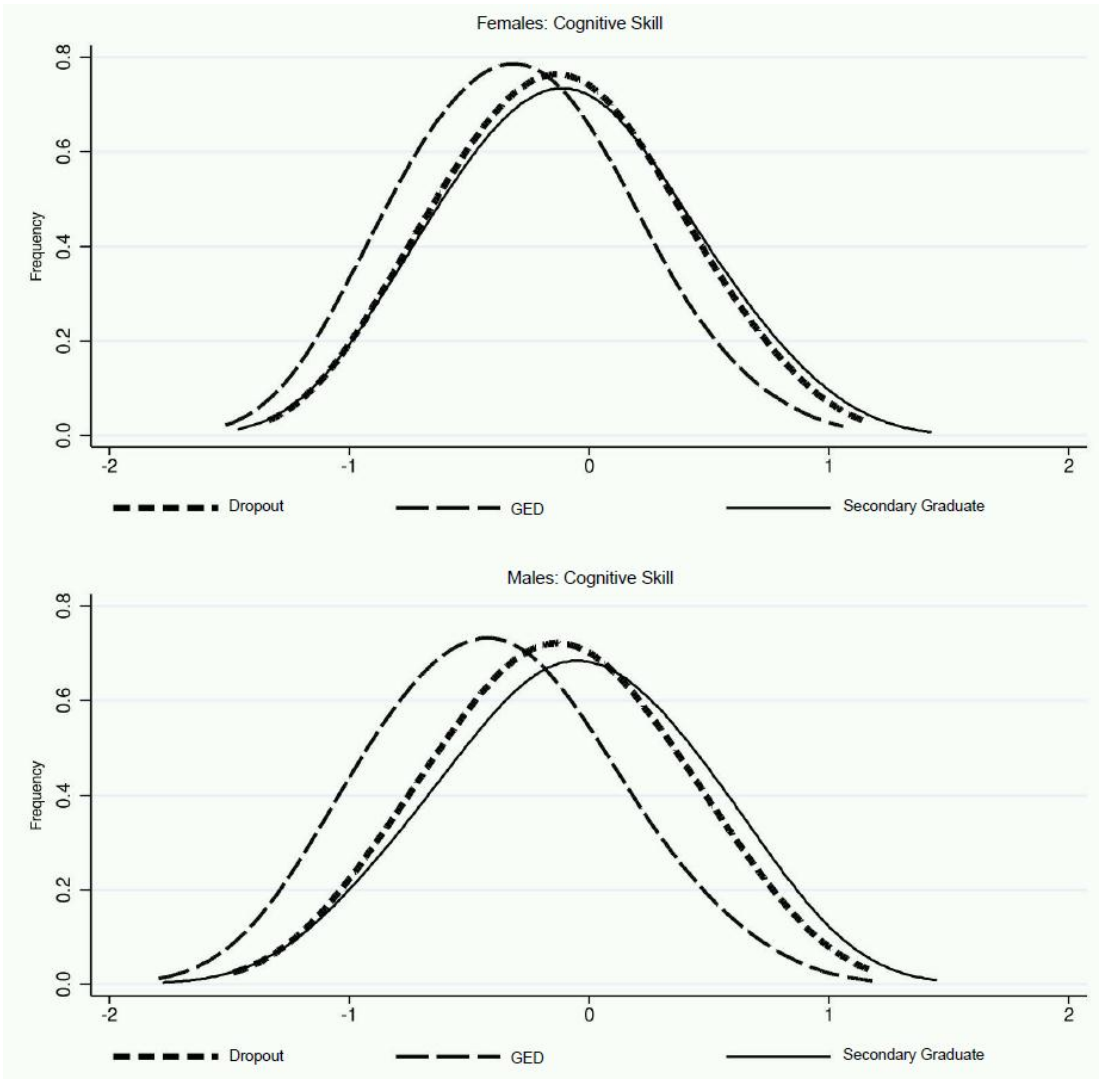
[4] Selecting measures and verifying them is part of the mysterious and inherently subjective process of "construct validity" in psychology. For a discussion, see Borghans et al. (2008).

[5] See the evidence surveyed and presented in Heckman, Humphries, and Kautz (2014b).

[6] See Heckman, Humphries, and Kautz (2014c) for more evidence on the differences between dropouts, GED recipients, and secondary school graduates.

Evaluating social programmes based on test scores alone can mislead policy-makers. Counting GEDs as high school graduates has artificially inflated the perceived performance of many social programmes.[7] Job Corps is a telling example. It is a national job training programme in the United States that costs over 1 billion dollars per year. Early evaluations found that Job Corps was successful, because it produced GEDs and it was assumed that GED recipients would earn as much as traditional graduates.[8] As discussed further in Section 9, later evaluations showed that there were no long-term effects of Job Corps on earnings or other outcomes. Relying on achievement testing to evaluate social programmes can divert funding to ineffective programmes.

Figure 6. Distribution of Cognitive Ability by Educational Status (No College Sample, All Ethnic Groups)



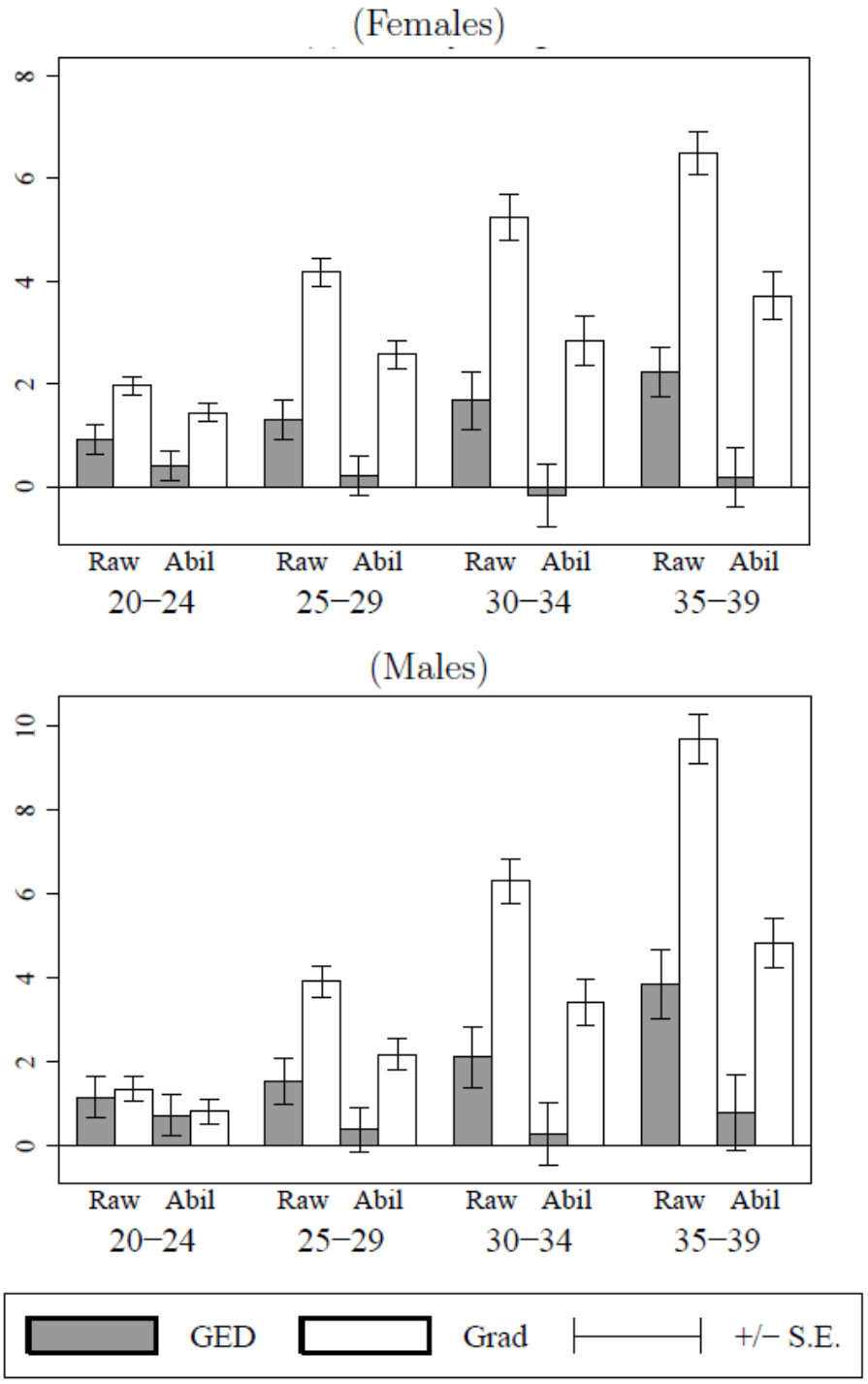
Note: The distributions represent cognitive factors, estimated using a subset of the Armed Services Vocational Aptitude Battery (ASVAB). The factors are adjusted for educational attainment, as laid out in Hansen, Heckman, and Mullen (2004). The sample is restricted to the cross-sectional subsample for both males and females. Distributions show only those with no postsecondary educational attainment. The cognitive factors are normalised by gender to be mean zero, standard deviation one.

Sources: Reproduced from Heckman et al. (2011), based on data from the National Longitudinal Survey of Youth, 1979 (NLSY79).

[7] See Humphries (2014) for a discussion of social programs that are evaluated based on the number of GEDs produced.

[8] See Schochet, Burghardt, and Glazerman (2001).

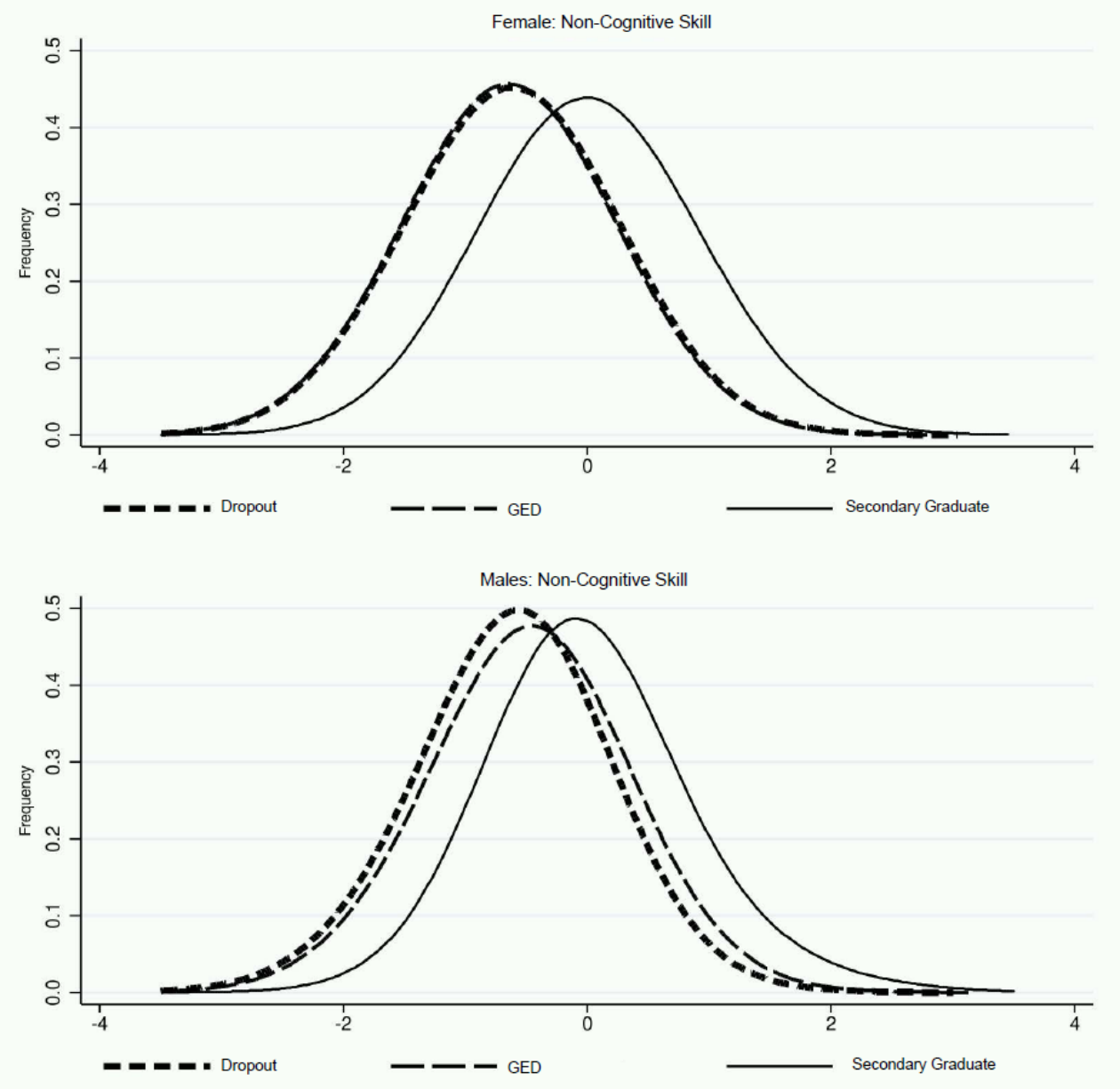
Figure 7. Hourly Wage Differences of GED Recipients and Traditional Graduates Compared to Uncertified Dropouts - Ages 20-39



Note: Controls: "Raw" – age and region or state of residence; "Abil" – AFQT adjusted for schooling at time of test. The intervals around each bar are standard errors centred around the mean – a measure of sampling variability.

Sources: Reproduced from Heckman et al. (2014b), which uses data from the National Longitudinal Survey of Youth, 1979 (NLSY79).

Figure 8. Distribution of Non-Cognitive Skills by Education Group



Note: The distributions represent non-cognitive factors, estimated using measures of early violent crime, minor crime, marijuana use, regular smoking, drinking, and early sexual intercourse. Sample restricted to the cross-sectional subsample for both males and females. Distributions show only those with no postsecondary educational attainment. The non-cognitive factors are normalised to be mean zero, standard deviation one.

Sources: Reproduced from Heckman et al. (2011), which uses data from the National Longitudinal Survey of Youth, 1979 (NLSY79).

THE SKILLS NEEDED FOR SUCCESS IN THE LABOUR MARKET[1]

Another perspective on the importance of non-cognitive skills comes from surveys of employers and workers. In a 1991 American report, the Secretary's Commission on Achieving Necessary Skills (SCANS) conducted an extensive analysis of which skills workers needed in the American workforce.[2] The Commission researched the literature, consulted with experts, and conducted detailed interviews with workers and/or supervisors in 50 occupations. The interviews rated the importance of various skills in the context of illustrative tasks and tools on the job. Using these sources, the Commission categorised necessary skills into basic skills, thinking skills, personal qualities, and a set of workplace competencies. In addition to reading, writing, and maths skills, basic skills include listening and speaking. The thinking skills cover creative thinking, decision making, problem solving, reasoning, and the ability to learn. SCANS specifies that personal qualities include responsibility, self-esteem, sociability, self-management, integrity, and honesty. SCANS identifies five groups of workplace competencies: the ability to allocate resources (time, money, facilities), interpersonal skills (such as teamwork, teaching others, leadership), the ability to acquire and to use information, the ability to understand systems, and the ability to work well with technology.

Employer surveys reinforce the importance of skills that go well beyond academic skills. In a survey of 3 200 employers in four large metropolitan areas in the Unites States, employers reported that such personal qualities as responsibility, integrity, and self-management are as important as or more important than basic skills (Holzer, 1997). In another employer survey undertaken in the mid-1990s of 3 300 businesses (the National Employer Survey), employers ranked attitude, communication skills, previous work experience, employer recommendations, and industry-based credentials above years of schooling, grades, and test scores as part of the skills needed for success in the workplace (Zemsky, 1997).

Non-cognitive skills are especially critical for entry level and hourly workers. Of employers drawn from a national sample in the United States in 1996, 69% reported rejecting hourly applicants because they lacked basic employability skills, such as showing up every day, coming to work on time, and having a strong work ethic. This percentage is more than double the percentage of rejecting applicants due to inadequate reading and writing skills. Rejections for not passing a drug test were almost as common as rejections for lack of literacy skills.[3] In a 2007 survey of employers in Washington State, about 60% reported difficulty in hiring. They experienced less difficulty finding workers with adequate reading, writing, and maths skills than with appropriate occupational, problem solving, teamwork, communication, and adaptability skills as well as positive work habits and a willingness to accept supervision.[4]

Evidence from the United Kingdom supports these findings. A 1998 survey of 4,000 employers found that the four skills found most lacking in 16 to 24-year-olds were technical and practical skills, general communication skills, customer handling skills, and teamwork skills.[5] At the bottom of the list were

[1] We have benefitted from the commentary of Robert Lerman in preparing this section. See Lerman (2013).

[2] Secretary's Commission on Achieving Necessary Skills (1992).

[3] Barton (2006).

[4] Washington Workforce Training Board (2008).

[5] Westwood (2004).

numeracy and literacy skills. In a 2002 survey of 4 000 employers in the UK, 23% of employers reported a significant number of their staff were less than fully proficient at their jobs. Skill shortfalls were most common in communication, teamwork, other technical and practical skills, customer handling, and problem solving and least common in numeracy and literacy.[6]

Consistent with these findings, the Confederation of British Industry defines employability as: *(1)* values and attitudes compatible with the work, including a desire to learn, to apply that learning, to improve, and to take advantage of change; *(2)* basic skills (literacy and numeracy); *(3)* key skills (communication, application of numbers, information technology, improving one's own learning and performance, working with others, problem solving) sufficient for the needs of the work; *(4)* other generic skills such as modern language and customer service skills and *(5)* job-specific skills and the ability to manage one's own career.

An ethnographic approach provides some revealing examples of how skills are used in context and how non-academic skills are often developed and used as part of a "community of practice".[7] In addition to formal knowledge, Nelsen (1997) points out that workplaces require facts, principles, theories, and maths and writing skills, but also informal knowledge embodied in heuristics, work styles, and contextualised understanding of tools and techniques. In her revealing case study of auto repair workers, Nelsen argues that social skills of new workers are very important for learning the informal knowledge of experienced workers, as captured in stories, advice, and guided practice.

[6] Hillage, Regan, Dickson, and McLoughlin (2002).

[7] Stasz (2001).

A FRAMEWORK FOR UNDERSTANDING INTERVENTIONS

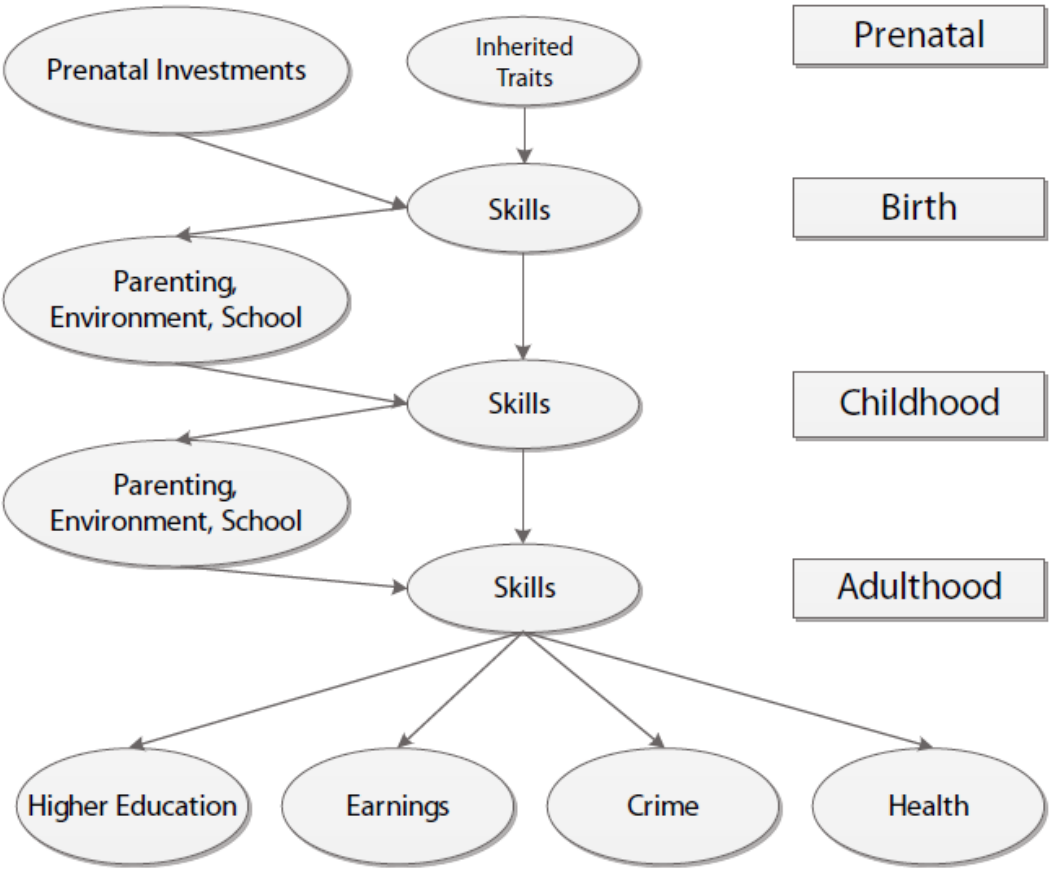
A variety of strategies have been advocated for promoting the development of non-cognitive and cognitive skills. Different advocates press their favourite strategies. Policymakers have limited resources, so it is important to identify which programmes are most effective. The rate of return – the per-period yield of an investment – provides a convenient and interpretable summary of the efficacy of competing programmes. An alternative to funding an early education programme might be to invest the resources in the stock market and to use these resources plus the accrued interest to fund a programme for adolescents or to invest in social infrastructure.

In this section, we present a conceptual scheme for understanding the dynamics of skill development.[1] Figure 9 illustrates the basic framework. In it cognitive and non-cognitive skills are produced by investments at different stages of the life cycle. Skills at birth depend on prenatal investments and inherited traits. Skill formation at later ages depends on the stock of skills acquired earlier as well as prior investments. For example, in a given classroom, a child who has a better attention span more easily acquires more cognitive skills. This concept is called self-productivity and is summarised by the motto "skills beget skills."

Investments (parenting, environment, and schools) also affect skills. The efficacy of investment is moderated by the stock of skills at any age. The benefits of an investment depend on the current level of skills – a phenomenon that economists call static complementarity. More motivated children benefit the most from additional investments. In addition, investment today increases the stock of future skills, which in turn increases the return to future investments. Economists call this phenomenon dynamic complementarity. This channel increases the returns to early investments because it makes future investments more productive. For this reason, Cunha et al. (2010) show that it is economically efficient to invest in the most disadvantaged young children because it raises their payoffs from future investments. Heckman and Mosso (2014) present a more complete discussion of static and dynamic complementarity and a formal proof of when early investment is more effective compared to later investment.

[1] These ideas are formalised in Cunha and Heckman (2007, 2008); Cunha et al. (2010).

Figure 9. Framework for Understanding Skill Development



This framework recognises that different skills might be relatively easy to shape at different stages of the life cycle. Sensitive periods for a given skill are periods when investments are relatively more productive. Critical periods for a particular skill are periods when investment during any other period is not productive.

Figure 9 illustrates why understanding the effects attributable to specific interventions is such a challenging task. Most of the empirical studies that we review only investigate the interventions aimed at one slice of the life cycle. They do not connect the links in the figure or correct for the effects of later investment in producing the outcomes attributed to early investments. One important area for future research on skill formation is to better document how early interventions influence the efficacy of later interventions.

SUMMARY OF EMPIRICAL EVIDENCE ON THE EFFICACY OF INTERVENTIONS

This section summarises the empirical evidence from a variety of interventions ranging from targeting prenatal infants to targeting young adults. Heckman and Kautz (2014a, 2014c) discuss these programmes in greater detail. In our analysis, we focus on programmes that have been well studied, have long-term follow-ups, have been widely adopted, or offer unique insights. Although there are some important exceptions, most of the programmes that they review are based in the United States, primarily because more high-quality evaluations have been conducted there. Descriptions and discussions of other programmes can be found in Heckman and Kautz (2014a) and the adjoining Web Appendix.[1]

For four reasons, evaluating and comparing the evidence from intervention programmes is challenging. First, many interventions are only evaluated with short-term follow-ups, which in practice can lead to upward-biased estimates of returns if the benefits eventually dissipate or to downward-biased estimates of the returns if the effects of the programmes appear later in life. Second, not all studies measure the same outcomes. Ideally, all studies would report outcomes in terms of the rate of return of the programme. Reported outcomes often differ across studies. Many studies only consider the effect of an intervention on a few outcomes. Without knowing the range of outcomes affected, it is difficult to calculate a rate of return. Third, many programmes target specific demographic groups and most of them target disadvantaged populations. Applying the findings from one group to another might be problematic if groups differentially benefit from programmes. Fourth, different programmes use different, often incompatible, measurement schemes.

Table 2 summarises the effects of many of the studies discussed in this report.[2] The table displays information about the nature of the intervention, the quality of the evaluation, the effects on later life outcomes, and estimates of the rate of return and cost-benefit ratio when available. The squares in the "Components" columns indicate the extent to which the programme and the evaluation of it have the features defined in the table. The dots in the "Effects on Outcomes" columns indicate the extent to which the programme influenced skills and outcomes. (The notes at the bottom of the table define the symbols and abbreviations used.)

Three striking patterns emerge about the nature of the programmes and the quality of the available evaluations of them. First, as a group, programmes targeting early childhood and elementary school have longer follow-ups. All of the early childhood or elementary school programmes in Table 2 have evaluations that follow participants for at least 10 years and many follow them more than 20 years, whereas only two evaluations of adolescent programmes follow participants for at least 10 years (the longest is 12).

Second, early childhood programmes tend to measure cognitive and non-cognitive skills in addition to a variety of later-life outcomes, whereas many of the adolescent evaluations focus solely on labour market outcomes. Because of these features of data availability, we can better understand the sources of the effects on adult outcomes of early childhood programmes by considering how these interventions produce skills.

[1] See https://cehd.uchicago.edu/page/web-appendix_fostering_skills_oecd.

[2] Note that there are several programs that appear in Table 2 but do not appear later in the text of this paper. The Web Appendix appendix to Heckman and Kautz (2014a) details these other programs and can be found here: https://cehd.uchicago.edu/page/web-appendix_fostering_skills_oecd.

Due to the absence of measures of skills for many adolescent interventions, understanding these programmes requires examining the curricula of the programmes themselves, for example, whether the programme seeks to foster cognitive or non-cognitive skills.

Third, selection into programmes differs by the age of intervention. In most early childhood evaluations, the programmes first contact parents to participate and then parents opt into the program. In contrast, in most adolescent evaluations, participants themselves chose to enter the programme.

Table 2 also suggests certain features of effective programmes. Only very early interventions (before age three) improve IQ in a lasting way, consistent with the evidence that early childhood is a critical period for cognitive development (see Knudsen, Heckman, Cameron, and Shonkoff, 2006). The most successful interventions target pre-schoolers (after age three) and primary school children. They improve later-life outcomes by developing non-cognitive skills.

Programmes that target adolescents have not been established to be as effective as programmes that target children at earlier ages, in part because there have been fewer long-term evaluations of them. Several of the successful adolescent mentoring or residential programmes improve labour market and social outcomes, but have relatively short follow-ups. The two programmes with the longest follow-ups improve outcomes in the short run, but the benefits fade after a few years. These programmes alter participants' environments and incentives during the intervention, which could influence their behaviour in the short term without having a lasting effect.

The most promising adolescent programmes integrate aspects of work into traditional education. Such programmes break down the rigid separation between school and work that characterises the American high school. In addition, a nascent literature reports substantial benefits from giving information to adolescents.

High schools create an adolescent society with values distinct from those of the larger society and removed from the workplace.[3] Even in affluent communities, the adolescent society has an anti-academic, anti-achievement bias. It was not until 1940 that more than half of each birth cohort graduated from high school in the United States.[4] In earlier times, adolescents took apprenticeships and jobs where they were supervised and mentored by adults. Mentoring involved teaching valuable non-cognitive skills—showing up for work, cooperating with others, and persevering on tasks. These skills could be fostered in high schools, but with the relaxation of discipline in the schools, it is more difficult to do so.[5]

The apparent success of apprenticeship programmes might arise in part from their cultivation of non-cognitive skills. The attachment of a supervisor to an apprentice helps create non- cognitive skills in a version of the attachment bond between parent and child.[6]

[3] See Coleman (1961).

[4] See Goldin and Katz (2008).

[5] See Arum (2005).

[6] See Bowlby (1951); Sroufe (1997); Sroufe, Egeland, Carlson, and Collins (2005).

EARLY-LIFE INTERVENTIONS THAT BEGIN BEFORE FORMAL SCHOOLING

This section summarises the interventions listed in Table 2 that start before children enrol in kindergarten. We divide the early interventions into two categories: *(1)* infant programmes and model preschools that target infants by directly providing cognitive and socio-emotional stimulation to children, by instructing young mothers or by providing centre-based care to specific populations and *(2)* large-scale programmes that have been implemented in schools and are relatively less expensive than centre-based care. Many early interventions have successfully boosted cognitive skills, non-cognitive skills, or both.

Given that there are more differences among evaluations than there are evaluations, it is difficult to understand exactly why some programmes are more successful than others. Nevertheless, some patterns emerge. Only programmes that start before the age of 3 have had lasting impacts on IQ. Some of the most successful programmes have had no long-term impact on IQ but have improved later-life outcomes by improving non-cognitive skills.

As a general rule, large-scale programmes have been less successful than smaller-scale pro- grams, but the large-scale programmes differ from small-scale programmes in important ways. For example, small-scale programmes tend to be more intensive and also involve parents more actively.

The evidence suggests that involving parents is an important component of successful interventions. While most interventions are temporary, parents are a lasting fixture in the lives of children. The parents with children in these interventions have responded by investing more in their children in later years.[1]

[1] See Heckman and Mosso (2014) and Moon (2014b) for summaries of the evidence on this issue.

Table 2. Summary of Effects for Main Intervention

	Participant/Evaluation Characteristics						Components					Effects on Outcomes							Return/Benefits	
Program	Age	Duration	Target	Selection	Follow-Up	Sample Size	Home	Health	Parental	On Site	Group	IQ	School	Non-Cog	Education	Health	Crime	Earnings	Return	Benefit/Cost
Early																				
NFP	< 0	2Y	SES	Prgrm	19Y	640	☒	☒	☒	□	□	●	◐	●	○	◍	◐	.		2.9
ABC	0	5Y	SES	Refer	30Y	90	□	☒	☒	☒	☒	●	●	◐	◐	◍	◐	◐		3.8
IHDP	0	3Y	Health	Prgrm	18Y	640	☒	□	☒	☒	☒	◐	◐	◐	○	○	○	.		
FDRP	0	5Y	SES	Prgrm	15Y	110	☒	☒	☒	☒	☒	○	◐	◍	.	.	●	.		
PCDC	1	2Y	SES	Prgrm	15Y	170	☒	□	☒	☒	☒	◍	◍	●	.	.	.	.		
JSS	1–2	2Y	Health	Prgrm	22Y	160	☒	☒	☒	□	□	●	●	●	◍	●	.	●		
Perry	3	2Y	SES, IQ	Prgrm	37Y	120	☒	□	☒	☒	☒	◍	●	●	◐	○	●	◐	7–10	7.1–12.2
Head Start	3	2Y	SES	Prnt	23Y	4,170	☒	☒	☒	☒	☒	◍	◍	○	●	●	◐	●		
CPC	3–4	2Y	SES	Prnt	25Y	1,470	□	☒	☒	☒	☒	.	●	●	●	●	●	●	18	10.8
TEEP	3,5	2Y	SES	Prgrm	22Y	260	☒	☒	☒	☒	□	◍	●	●	●	.	.	.		
STAR	5–6	4Y	SES	Prgrm	22Y	11,000	□	□	□	☒	☒	.	◍	●	●	.	.	●	6.2	
Elementary																				
LA's Best	5–6	6Y	SES	Schl	12Y	19,320	□	☒	☒	☒	☒	.	◍	.	●	.	○	.		0.9
CSP	5–13	5Y	Behav	Refer	35Y	510	☒	☒	☒	☒	☒	.	.	○	.	⊗	○	.		
SSDP	6–7	6Y	Crime	Prgrm	21Y	610	□	☒	☒	☒	☒	.	●	●	●	●	◍	○		3.1
Adolescence																				
BBBS	10-16	1Y	SES	Self	1Y	960	□	□	□	□	□	.	◐	○	.	◐	○	.		1.0
IHAD	11–12	7Y	SES	Prgrm	8Y	180	□	☒	☒	☒	☒	.	.	.	●	.	.	.		
EPIS	13–15	3Y	Schl	Schl	2Y	45,070	□	□	□	☒	☒	.	●	.	.	.	.	.		0.9–3.0
xl club	14	2Y	Schl	Schl	2Y	261,420	□	□	□	☒	☒	.	○	.	.	.	.	.		
SAS	14–15	5Y	Schl, SES	Schl	6Y	430	□	□	□	□	□	.	◍	○	●	.	.	.		
STEP	14–15	2Y	Schl, SES	Self	4Y	4,800	□	☒	□	☒	☒	.	○	.	○	.	.	○		
QOP	14–15	5Y	Schl	Prgrm	10Y	1,070	□	☒	□	☒	☒	.	○	.	◍	○	⊗	○		0.42
Academies	13–16	4Y	Schl, SES	Self	12Y	1,460	□	□	□	☒	☒	.	◍	○	○	○	◍	◐		
ChalleNGe	16–18	1Y	Dropout	Self	3Y	1,200	□	☒	□	☒	☒	.	.	◍	●	◐	◍	●	6.4	2.66
Job Corps	16–24	1Y	SES	Self	9Y	15,300	□	☒	□	☒	☒	.	.	.	○	◍	◍	◍		0.22
Dominican	16–29	1Y	SES	Self	2Y	9,410	□	☒	□	☒	☒	.	.	●	.	◍	.	◐		
Year-Up	18–24	1Y	SES	Self	4Y	200	□	□	□	☒	☒	.	.	.	○	.	.	●		
SSP	19+	3Y	SES	Prgrm	4Y	4,960	□	□	□	□	□	.	.	●	.	.	.	◐		2.67

Note: ☐ – Does not include intervention component. ☒ – Includes intervention component. ○ – No effects. ● – Positive effects. ◍ – Weakly positive effects. ◐ – Mixed effects (either different studies find different results or only particular sub-populations benefited). ⊗ – Negative effects. "." – Not measured. "Age" – The age at which participants entered the programme. For programmes that targeted grades, rather than ages, it was assumed that children entered kindergarten at ages 5-6 and the age range advanced one year for each subsequent grade. "Duration" – Length of the treatment. In cases where the treatment length varied for participants, the longest duration was presented. "Target" – Population that was targeted by the programme. SES – Socioeconomic status or disadvantage. Behav- Behaviour. Schl – School performance. Crime – Local crime rates. IQ – Low IQ. "Selection" – The party that acted first in joining the sample. Prgrm – Evaluation programme contacted participants. Refer – Other party referred participants to programme. Prnt – Parent applied to programme. Self – Participant applied to programme. Schl – School selected participants. "Follow-up" – Duration of longest follow-up evaluation in years. "Sample" – Largest sample size from the studies examined (rounded to nearest 10). "Home" – Included home visits. "Health"– Included a nutritional component. "Parental" – involved parents. "On Site" – Took place at an on site location. "Group"– Whether the intervention combined participants in groups. "ÏQ" – IQ score. "School" – school performance. "Non-cog" – measured non-cognitive skills. "Education" – educational attainment. "Health" – health (including drug use). "Crime"– crime. "Earnings" – earnings or related outcomes. "Return"– Annual rate of return. "Benefit/Cost"- Estimated benefits divided by costs.

Sources: **Nurse Family Partnership (NFP)** – Eckenrode, Campa, Luckey, Henderson, Cole, Kitzman, Anson, Sidora-Arcoleo, and Olds (2010); Kitzman, Olds, Cole, Hanks, Anson, Arcoleo, Luckey, Knudtson, Henderson, and Holmberg (2010); Olds (2006); Olds, Kitzman, Cole, Hanks, Arcoleo, Anson, Luckey, Knudtson, Henderson, Bondy, and Stevenson (2010); Olds, Kitzman, Cole, Robinson, Sidora, Luckey, Henderson, Hanks, Bondy, and Holmberg (2004); Olds, Kitzman, Hanks, Cole, Anson, Sidora-Arcoleo, Luckey, Henderson, Holmberg, Tutt, Stevenson, and Bondy (2007); Olds, Robinson, Pettitt, Luckey, Holmberg, Ng, Isacks, Sheff, and Henderson (2004). **Abecedarian Project (ABC)** – Breitmayer and Ramey (1986); Heckman, Moon, and Pinto (2014); Ramey, Collier, Sparling, Loda, Campbell, Ingram, and Finkelstein (1976); Temple and Reynolds (2007). **Infant Health and Development Program (IHDP)** – McCormick, Brooks-Gunn, Buka, Goldman, Yu, Salganik, Scott, Bennett, Kay, Bernbaum, Bauer, Martin, Woods, Martin, and Casey (2006). **Syracuse Family Development Research Program (FDRP)** – Lally, Mangione, and Honig (1987). **Houston Parent-Child Development Center (PCDC)** – Besharov, Germanis, Higney, and Call (2011); Bridgeman, Blumenthal, and Andres (1981); Johnson and Walker (1991); Walker and Johnson (1988). **Jamaican Supplementation Study (JSS)** – Gertler, Heckman, Pinto, Zanolini, Vermeersch, Walker, Chang, and Grantham-McGregor (2013); Grantham-McGregor, Powell, Walker, and Himes (1991); Walker, Chang, Powell, and Grantham-McGregor (2005); Walker, Chang, Powell, Simonoff, and Grantham-McGregor (2007). **Perry Preschool Program (Perry)** – Heckman et al. (2010a); Heckman, Moon, Pinto, Savelyev, and Yavitz (2010b); Heckman et al. (2013); Schweinhart, Barnes, and Weikart (1993). **Head Start** – Carneiro and Ginja (2012); Currie and Thomas (1995); Deming (2009); Garces, Thomas, and Currie (2002); Ludwig and Miller (2007); Westat (2010). **Chicago Child-Parent Center Program (CPC)** – Niles, Reynolds, and Nagasawa (2006); Reynolds (1994); Reynolds and Temple (1998); Reynolds, Temple, Ou, Arteaga, and White (2011); Reynolds, Temple, Robertson, and Mann (2002); Reynolds, Temple, White, Ou, and Robertson (2011). **Turkish Early Enrichment Project (TEEP**) – Kagitcibasi, Sunar, and Bekman (2001); Kagitcibasi, Sunar, Bekman, Baydar, and Cemalcilar (2009). **Project STAR (STAR)** – Chetty, Friedman, Hilger, Saez, Diane, and Yagan (2011); Krueger (2003). **LA's BEST** – Goldschmidt and Huang (2007); Huang, Gribbons, Kim, Lee, and Baker (2000); Huang, Kim, Marshall, and Pérez (2005). **Cambridge Somerville Program (CSP)** – McCord (1978). **Seattle Social Development Project (SSDP)** – Aos, Lieb, Mayeld, Miller, and Pennucci (2004); Hawkins, Catalano, Kosterman, Abbott, and Hill (1999); Hawkins, Kosterman, Catalano, Hill, and Abbott (2005, 2008**). Big Brothers Big Sisters (BBBS)** – Aos et al. (2004); Tierney, Baldwin-Grossman, and Resch (1995). **I Have a Dream (IHAD**) – Kahne and Bailey (1999). **Empresários Pela Inclusão Social (EPIS)** – Martins (2010). **xl Club** – Holmlund and Silva (2009). **Sponsor a Scholar (SAS)** – Johnson (1999). **Summer Training and Education Program (STEP)** – Walker and Vilella-Velez (1992). **Quantum Opportunity Program (QOP)** – Aos et al. (2004); Rodríguez-Planas (2010, 2012). **Career Academies (Academies)** – Kemple and Snipes (2000); Kemple and Willner (2008). **National Guard ChalleNGe Program (ChalleNGe)** – Bloom, Gardenhire-Crooks, and Mandsager (2009); Millenky, Bloom, and Dillon (2010); Millenky, Bloom, Muller-Ravett, and Broadus (2011); Perez-Arce, Constant, Loughran, and Karoly (2012). **Job Corps** – Schochet et al. (2001); Schochet, Burghardt, and McConnell (2006, 2008). **Dominican Youth Employment Program (Dominican)** – Gertler and Martinez (2012). **Year Up** – Roder and Elliot (2011); Roder and Elliott (2014). **Self-Suffciency Project (SSP)** – Card and Hyslop (2009); Gottschalk (2005); Michalopoulos, Tattrie, Miller, Robins, Morris, Gyarmati, Redcross, Foley, and Ford (2002); Zabel, Schwartz, and Donald (2010, 2013).

Infant Programmes and Model Preschools

Nurse-Family Partnership

Evidence from the Nurse-Family Partnership (NFP) programme shows how improving prenatal conditions and early parenting can improve child skills and outcomes. NFP aims at first- time, low-income, unmarried, and/or adolescent mothers without previous births. In the programme, nurses visit young mothers from the first or second trimester of the mother's pregnancy until the second birthday of her first child. The programme encourages mothers to reduce smoking, teaches mothers how to care for their children, and helps mothers to pursue education and find jobs. It was introduced, as a pilot, at Elmira, New York. Later, the programme was introduced at Memphis, Tennessee, and Denver, Colorado.[2] At Elmira and Memphis, mothers receive around 30 to 35 visits of around 75–90 minutes each. It is evaluated by the method of randomised assignment (Howard and Brooks-Gunn, 2009; Olds, 2006).

NFP reduces risky behaviours of mothers and their children (see Table 3). Children exhibit persistently higher IQ scores through age 6 (Olds, Kitzman, Cole, Robinson, Sidora, Luckey, Henderson, Hanks, Bondy, and Holmberg, 2004).[3] At age 12, the children in the treatment group have lower rates of substance abuse and exhibit lower levels of internalising behaviour (e.g. anxiety, depression, and withdrawal).[4] When the children are 12, mothers are much less dependent on welfare.[5] By age 19, children are less likely to engage in crime, even though there are only weak effects on achievement test scores or grades in the long run.[6] [7]

These results suggest that the programme works by improving non-cognitive skills. NFP might have had long term effects on these skills because parents responded to the intervention by improving the home environments of their children in a lasting way.[8]

[2] Our discussion will focus on the effects of the program up to age 12 for the Memphis sample and the effects at age 19 for the Elmira sample.

[3] Measured by the mental processing domain of the Kaufman Assessment Battery for Children (KABC).

[4] Kitzman et al. (2010). The measure of internalizing behaviour is based on self-reports from the Achenbach Child Behaviour Checklist (Achenbach and Rescorla, 2001). See Almlund et al. (2011) for a discussion of the relationship between internalizing behaviour and the Big Five.

[5] Olds et al. (2010).

[6] Eckenrode et al. (2010).

[7] Illustrated by the strong discrepancy in Table 3 between estimates for "Ever Arrested" versus "Arrested in Last Year."

[8] See Heckman and Mosso (2014) and Moon (2014b) for summaries.

Table 3. Summary of Effects of the Nurse-Family Partnership

Outcome	Treatment	Control	Treatment Effect
Age 6[1]			
Vocabulary Skills(ES)			0.17**
Internalizing Disorders(%)	12.6	14.7	−2.1
Externalizing Disorders(%)	17.4	20.2	−2.8
Age 9[2]			
GPA(ES)			0.09
Antisocial Behavior(%)			−0.03
Grade Retention(%)	16.0	12.4	3.6
Age 12[3]			
GPA			0.08
Achievement Tests[a]			1.09
Grade Retention(%)	24.9	20.8	4.1
Internalizing Disorders(%)[b]	22.1	30.9	−8.8**
Externalizing Disorders(%)[c]	19.7	17.8	1.9
Used Substance Last 30 Days(%)	1.7	5.1	−3.4**
Ever Arrested(%)	3.1	3.1	0.0
Welfare Benefits Mother[d]	8772	9797	−1025**
Age 19[4]			
Ever Arrested(%)	21.3	37.4	−16.1**
Arrested in Last Year(%)	8.2	5.5	2.7
Illicit Drug Use(%)	48.7	51.9	−3.2
Has HS Diploma(%)	70.6	74.5	−3.9
Economically Productive(%)[e]	71.4	68.3	3.1

Note: The estimates are coefficients from regressions that control for sample member characteristics. % refers to treatment effects in terms of changes in prevalence in outcome variable in percentage points. ES indicates Effect Size, reporting the treatment effect in standard deviations of the outcome variable. The Age 6-12 estimates come from the Memphis site. The Age 19 estimates come from the Elmira site. (a) Is based on group reading and math achievement test scores and is in percentile units. (b) Uses student self-reports on domains such as anxiety, depression, somatisation, and withdrawal to assess if students pass a clinical threshold, based on the Achenbach Child Behaviour Checklist (Achenbach and Rescorla, 2001). (c) Uses student, teacher, and parent reports on domains such as conduct problems, aggression, and total problems to assess if students pass a clinical threshold, based on the Achenbach Child Behaviour Checklist. (d) Measures the average yearly receipt of welfare during the child's first 12 years of life, in US$. (e) Measures if someone is involved in education, a job, the military, or job training. ** Attains 5% significance level.

Sources: (1) The estimates at Age 6 come from Olds et al. (2004). (2) The estimates at Age 9 come from Olds et al. (2007). (3) The estimates at Age 12 come from Kitzman et al. (2010). (4) The estimates at Age 19 come from Eckenrode et al. (2010).

Jamaican Study

Many childhood programmes have been introduced in less developed countries. The Jamaican Supplementation Study (JSS) was one of the few with a long-term follow-up.[9] Like NFP, JSS targeted health and parenting skills. This randomised programme consisted of a two-year nutritional and stimulation programme for stunted children (low height for age), aged 9-24 months at the start of the programme. Participants either received supplementation (milk formula), stimulation (encouraged mother to play with kids in an effective manner), or both (Grantham-McGregor et al., 1991).

In general, the stimulation intervention outperformed the supplementation intervention. Both interventions improved early cognitive development in the short-term, but only the stimulation had long-term effects on skills. The stimulation intervention improved cognitive and non-cognitive skills at age 17-18 by about half a standard deviation (Gertler et al., 2013; Grantham-McGregor et al., 1991). As with the NFP programme, the stimulation programme improved internalising behaviour (Walker et al., 2005). The stimulation intervention also improved earnings by 33% at age 22 (Gertler et al., 2013; Gertler, Heckman, Pinto, Zanolini, Vermeersch, Walker, Chang, and Grantham-McGregor, 2014).[10] This intervention supports the findings of NFP that relatively small interventions that teach parenting skills can have a big impact later in life.

Perry Preschool Programme

Evaluations of the Perry Preschool programme provide some of the most compelling evidence that non-cognitive skills can be boosted in ways that produce adult success. The Perry Preschool programme enriched the lives of 3- and 4-year-old low-income black children with initial IQs below 85 at age three.[11] They attended two-and-a-half-hours of centre-based preschool five days a week for two years. In addition, home visits promoted parent-child relationships. The programme ended after two years of enrolment, and both treatment and control groups entered the same school.

Participants were taught social skills in a daily "plan-do-review" sequence in which children first planned a task, executed it, and then reviewed it with teachers and fellow students. They learned to work with others when problems arose.[12] The programme was evaluated by the method of random assignment.

---

[9] Most programs lack long-term follow-ups (even beyond age 10), or any follow-up at all, and few employ methods of random assignment. Short-term effects are generally positive; intervention leads to better cog- nitive skills and, in most cases, lower grade retention in the early years of primary school. For an overview, see Engle, Fernald, Alderman, Behrman, O'Gara, Yousafzai, Cabral de Mello, Hidrobo, Ulkuer, Ertem, and Iltus (2011); Grantham-McGregor, Cheung, Cueto, Glewwe, Richter, and Strupp (2007); Grantham-McGregor et al. (1991); Jolly (2007); Lake (2011); Myers (1992); Walker, Wachs, Grantham-McGregor, Black, Nelson, Huffman, Baker-Henningham, Chang, Hamadani, Lozoff, Meeks Gardner, Powell, Rahman, and Richter (2011); Walker et al. (2005); Walker, Chang, Vera-Hern´andez, and Grantham-McGregor (2011); Walker, Wachs, Gardner, Lozoff, Wasserman, Pollitt, Carter, and The International Child Development Steering Group (2007); Young (1996) Engle, Black, Behrman, Cabral de Mello, Gertler, Kapiriri, Martorell, Eming Young, and The International Child Development Steering Group (2007).

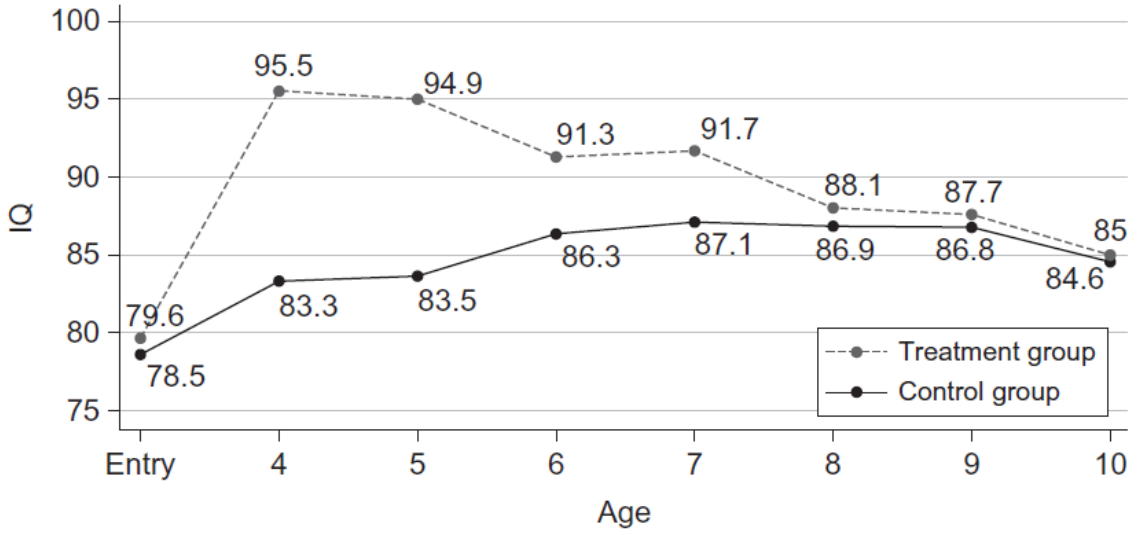
[10] The supplementation program might have been ineffective in the long run because the supplement was relatively weak, contained few micronutrients, and had no fortified formula. Some of the interventions referred to in Myers (1992) and Young (1996) also reported stronger effects of stimulation over supplementation.

[11] We draw on the analysis of Heckman et al. (2010a, 2010b, 2013).

[12] Sylva (1997) describes the Perry program as a program that fosters non-cognitive skills. It has features in common with the Tools of the Mind intervention discussed below. See Bodrova and Leong (2001) for the Tools of

The programme did not improve IQ scores in a lasting way. Figure 10 shows that, by age ten, treatment and control groups had the same average IQ scores. Many critics of early childhood programmes seize on this finding and related evidence to dismiss the value of early intervention studies. Arthur Jensen's (1969) discussion of IQ fadeout in Head Start and other compensatory programmes promoted the widespread embrace of the notion that intervention efforts are ineffective and that intelligence is genetically determined. His uncritical reliance on intelligence test scores illustrates the fallacy of relying on monodimensional measurements of human skills.

Figure 10. Perry Preschool Programme: IQ, by Age and Treatment Group



Note: 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.

Sources: Cunha et al. (2006) and Heckman and Masterov (2007), based on data provided by the High Scope Foundation.

The Perry intervention provides an effective rebuttal to these arguments. The programme greatly improved outcomes for both participating boys and girls, resulting in a statistically significant rate of return around 7%-10% per annum for both genders (see Heckman et al., 2010a). These returns are in the range of the post-World War II, pre-2008 meltdown stock market returns to equity in the U.S. labour market, estimated to be 6.9% per annum.[13]

Heckman, Pinto, and Savelyev (2013) show that the Perry treatment effects arise primarily from lasting chances in non-cognitive skills and not from changes in IQ. Figure 11 shows histograms of measures of non-cognitive skills for the treatment and control groups. The treatment groups of both genders improved their teacher-reported externalising behaviour, a skill related to Agreeableness and Conscientiousness. For girls the programme also improved Openness to Experience (proxied by academic motivation). Figure 11 also shows that the programme improved scores on the California Achievement Test (CAT), even though it did not have a lasting effect on IQ. This evidence is consistent with the evidence of Figure 2 showing that achievement test scores depend strongly on non-cognitive skills (see Borghans et al., 2008; Borghans, Golsteyn, Heckman, and Meijers, 2009). Achievement tests measure

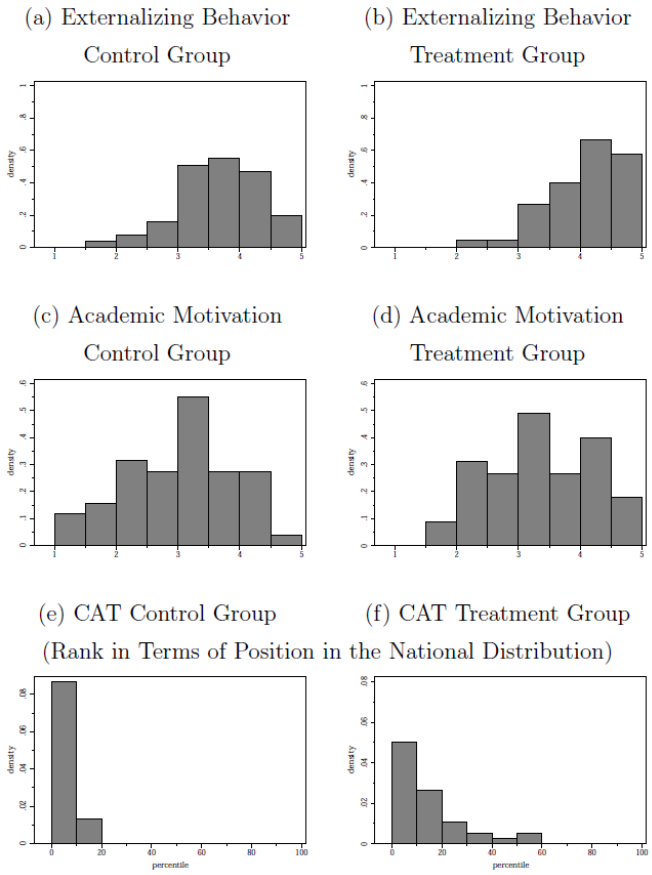
---

the Mind curriculum. See Heckman and Pinto (2015) for a discussion of the Perry curriculum and the Tools of the Mind curriculum.

[13] See DeLong and Magin (2009).

general knowledge. The acquisition of general knowledge (crystallised intelligence) depends on persistence, curiosity, and focus. As in the NFP programme, some of the benefits of Perry are due to enhanced parenting. Heckman and Mosso (2014) and Moon (2014b) show that the intervention also changed parents' belief in the importance of parenting, especially for the parents of boys.[14]

Figure 11. Histograms of Indices of Non-Cognitive Skills and CAT Scores



Note: Indices for Externalising Behaviour and Academic Motivation are based on items of the Pupil Behaviour Inventory (PBI), teacher ratings of student behaviour. The units are expressed in terms of standard deviations. The scale on "Externalising Behaviour" is normalised so that a higher score corresponds to better behaviour. The PBI includes whether the student disrupts classroom. "CAT" is the California Achievement Test score expressed in percentiles of the general population distribution of the scores. The one-sided *p*-values for difference in means are 0.001, 0.043, and 0.000 for Externalising Behaviour, Academic Motivation, and CAT scores, respectively. Histograms are based on the pooled sample of males and females.

Source: Heckman et al. (2013)

Abecedarian Programme

The Abecedarian (ABC) programme was another programme aimed at disadvantaged black children. It started earlier, lasted much longer, and was much more intensive than the Perry programme, combining a preschool intervention that started when children were as young as 6 weeks old with school-age treatment

[14] In line with these results in a late life book Arthur Jensen (1998) acknowledged that Conscientiousness was an important determinant of life success although he did not consider how it (or IQ) might be fostered by interventions.

through grade three.[15] The preschool component was full-day child care five days per week, 50 weeks a year. The curriculum focused on a series of educational games but also had a medical and nutritional component.[16] During grades one through three, teachers and parents interacted on a bi-weekly basis.

In contrast to Perry, the preschool component of ABC led to lasting improvements in IQ.[17] For girls, the programme improved IQ through age 21. The effect for boys was positive but was less precisely estimated. Girls and boys also scored better on achievement tests. ABC likely improved IQ because it started at an earlier age than Perry. Very early childhood appears to be a critical period for shaping IQ.

As with Perry, the benefits of the ABC programme differ across genders. For girls, the programme improved educational attainment, reduced participation in criminal activity, decreased substance abuse, and improved internalising and externalising behaviour. Like the Perry programme, ABC improved employment and health for males and produced substantial improvements in non-cognitive skills.[18]

The ABC programme also sheds some light on how later interventions might complement early interventions. Heckman et al. (2014) also study the effects of the school-age follow- up treatment. Half of the original treatment group and half of the original control group were assigned to a three-year, school-age treatment after the preschool programme. For boys, the school-age treatment had little effect regardless of whether they were in the preschool treatment group or control group. For girls, the school-age treatment had a positive effect for those in the preschool treatment group but not for those who were in the control group. The evidence for girls can be explained in part by dynamic complementarity – the early programme enhanced the benefits of the later programme.[19]

Large-Scale Infant and Young Child Programmes

Head Start

The Head Start programme receives considerable attention both in the evaluation literature and in public discussion. Head Start children are eligible for enrolment from ages three to five, although they generally enter at age four and receive one year of treatment. In addition to a centre-based (preschool) intervention, Head Start includes medical services and parental assistance (Ludwig and Miller, 2007). Its implementation differs greatly across sites, making it difficult to evaluate its overall effectiveness (Deming, 2009). Parental participation was not mandatory, and there is some evidence that many do not participate.[20]

There have been no long-term evaluations of Head Start based on randomised assignment. Instead, most evaluations of Head Start compare siblings with different enrolment statuses[21] or use regression discontinuity designs.[22] Empirical evidence on Head Start is mixed. Many studies find that Head Start

[15] It is easy to exaggerate the difference in intensity between the Perry program and the ABC program. While the ABC program starts earlier and spends more time each day with the child, the amount of time per day spent on essentially the same type of learning activities is very similar at ages 3 and 4. See Griffin, Heckman, and Moon (2013).

[16] See Campbell, Pungello, Miller-Johnson, Burchinal, and Ramey (2001); Campbell, Ramey, Pungello, Sparling, and Miller-Johnson (2002).

[17] We report on the treatment effects of the preschool component from Heckman et al. (2014).

[18] Conti, Heckman, Moon, and Pinto (2014) report substantial beneficial effects of the program on adult health for both genders.

[19] Campbell, Conti, Heckman, Moon, Pinto, Pungello, and Pan (2014) show that the ABC program had substantial effects on health.

[20] Schumacher (2003).

[21] For example, Deming (2009); Garces et al. (2002).

[22] For example, Carneiro and Ginja (2011); Ludwig and Miller (2007).

improves IQ scores and achievement test scores, but that the improvements fade by age ten (Deming, 2009; Ludwig and Miller, 2007). For some subgroups, there is evidence of persistent gains in test scores and reductions in retention rates.[23] Garces et al. (2002) find that Head Start improves outcomes for educational attainment for whites and reduces criminality for blacks, whereas Deming (2009) finds that Head Start improves educational attainment for blacks.

A short term experimental study by Westat (2010) based on randomised assignment offers a more pessimistic view. The study finds that most positive impacts on achievement tests[24] and socio-emotional development[25] fade by first grade. This evaluation likely underestimates the true programme effects. Many members of the control group joined the programme at sites other than where they initially applied or were enrolled in other childhood programmes that were more intensive than Head Start.

Chicago Child–Parent Center (CPC)

The Chicago Child–Parent Center (CPC) is an apparently successful large-scale preschool programme. It is directed toward disadvantaged, predominantly black, inner-city children in Chicago. The programme provides a half- or full-day preschool programme for three- and four-year-olds. The programme was designed to develop basic reading, writing, and maths skills. In contrast to other large-scale programmes, parents are required to visit the centres and receive advice on good parenting behaviour (Reynolds, 2000).

CPC has not been evaluated by the random assignment method. All evaluations are based on samples of children attending CPC kindergarten matched to "comparable" children attending non-CPC kindergarten. Based on these methods, Reynolds et al. (2011) and Reynolds et al. (2011) report that the programme increased high school graduation by about seven percentage points, reduced substance abuse by four-six percentage points, and reduced arrest rates by six percentage points. Because IQ was not measured, it is difficult to tell what causes the improvement in adult outcomes. Some evidence suggests that non-cognitive skills played a role. At age 13, the participants have higher levels of social and emotional competence. The programme also improves achievement test scores by about 1/3 of standard deviation (Iowa Tests of Basic Skills),[26] but improvements in non-cognitive skills can lead to improvements in achievement tests without affecting IQ, as was the case with the Perry programme (Reynolds, 1995).

CPC might have been more successful than other similar programmes, like Head Start, because it encouraged parental involvement. Parents are induced to visit the centres, and the staff teaches them various forms of stimulative play that they can do at home and assists parents interested in pursuing further education and seeking jobs (Reynolds, 2000).

[23] Deming (2009) reports persistent gains for males and those with higher ability mothers as measured by summary test score that includes the Peabody Picture Vocabulary Test (PPVT), Peabody Individual Achievement Math (PIATMT), and Reading Recognition (PIATRR). Retention rates were lower for boys, blacks, and students with low maternal AFQT.

[24] The tests included the Peabody Picture Vocabulary Test and several batteries of the Woodcock-Johnson III test.

[25] The measures included parent- and teacher-reported measures of aggression, hyperactivity, withdrawn behaviour, and total problem behaviour. They use several test batteries, including the Achenbach Classroom Behaviour Checklist (CBCL; Achenbach and Rescorla, 2001) and the Adjustment Scales for Preschool Intervention (ASPI; Lutz, Fantuzzo, and McDermott, 2000).

[26] Hieronymus, Lindquist, and Hoover (1980).

EDUCATION AND INTERVENTIONS IN KINDERGARTEN AND ELEMENTARY SCHOOL

One hundred and fifty years ago, moral education, rooted in Protestant Christianity, was taught in American public schools.[1] Moral and character education disappeared from the curriculum of public schools. However, moral education is still featured in Catholic education. Some private schools teach a secular version of moral education. For example, the Knowledge Is Power Program (KIPP) Charter Schools enforce strict codes of discipline and give character (non- cognitive skill) training to their participants (The KIPP Foundation, 2011). In contemporary society, discussions of moral and character education provoke controversy even among its advocates. Scholars disagree about the origins of character and morality and how they can best be fostered.[2]

In this section, we survey the evidence on secular versions of non-cognitive skill development programmes. This does not reflect any antireligious bias on our part. Instead, it reflects the absence of convincing evaluations of the effectiveness of non-cognitive skill development programmes in religious schools.

Targeted Non-Cognitive Skill Interventions

Few well-evaluated interventions have targeted children in elementary school. Durlak, Weissberg, Dymnicki, Taylor, and Schellinger (2011) present a meta-analysis of 213 school-based social and emotional learning programmes. While their study suggests that these programmes have been successful, many of the evaluations they review suffer from substantial methodological problems. For example, only 15% of the studies have follow-ups that last beyond 6 months. For the studies with longer follow-up, the mean impact is positive and statistically significant. However, these follow-ups are still short compared to those for the Perry Preschool programme and ABC.[3] Although these programmes appear promising, evidence on how effective they are in the long-run is lacking.

The Seattle Social Development Project (SSDP)

Of all of the programmes reported in Durlak et al. (2011), the Seattle Social Development Project (SSDP) has been evaluated the most thoroughly. The programme targeted public elementary schools in high-crime areas of Seattle. The full programme lasted from first grade through sixth grade.

[1] See the discussions in Heckman and Kautz (2014c) and Kaestle (1984).

[2] The conceptualisations of morality and character differ in the literature. Non-cognitive skill is usually viewed as a skill acquired through habituation and practice. In its perfected state it becomes a gut reaction. Morality is viewed as a consequence of conscious choices. Many dismiss discussions of morality and character out of hand because of their religious connotations or because they suggest a prudish vision of society that does not tolerate cultural pluralism. Lapsley and Yeager (2012) survey the controversies in the field, the curricula available and evidence on their effectiveness.

[3] The largest federal study to date on character education programs failed to find evidence for improvements in behaviour or academic performance (see Social and Character Development Research Consortium, 2010).

The programme emphasised attachment and interaction between children and their parents and teachers. Throughout elementary school, the participants' teachers received five days of training per year that included proactive classroom management, interactive teaching, and cooperative learning. The first-grade teachers received additional lessons based on the Cognitive Problem-Solving curriculum, which teaches children to resolve conflicts with peers (Shure and Spivack, 1988). In first and second grade, parents were offered a seven- session course on behavioural management training.[4] In second and third grade, parents were offered a four-session course to support their children's academic achievement.[5] Finally, in fifth and sixth grade, parents were offered a five-session course designed to reduce drug use.[6] Parents of 43% of the children participated in parenting classes.[7]

The programme improved a variety of long-run outcomes. Table 4 shows that SSDP improved grades and behaviour during adolescence. At age 21, participants were 10 percentage points more likely to have graduated from high school (or have earned an equivalency degree). By age 24, this effect faded, but this likely occurred because members of the control group earn GED certificates.[8] GED certificates do not confer the same benefits as a traditional high school diploma.[9] By ages 24 and 27, the participants were 12 percentage points more likely to earn an associate's degree. The treatment group had higher earnings at ages 24 and 27, although treatment effects on income are not precisely determined. The programme improved self-efficacy at age 21 and 24 but did not have a statistically significant effect on achievement test scores. The programme improved mental health in the long run. Like many other programmes, if one judged it solely on the basis of achievement tests, SSDP would be unsuccessful. Considering meaningful life outcomes and non-cognitive outcomes presents a far more positive picture.

[4] The course followed the "Catch 'Em Being Good" curriculum (Hawkins, Von Cleve, and Catalano Jr, 1991).

[5] The course consisted of four sessions that followed the "How to Help Your Child Succeed in School" curriculum (Abbott, O'Donnell, Hawkins, Hill, Kosterman, and Catalano, 1998).

[6] The course followed the "Preparing for the Drug (Free) Years" curriculum (Haggerty, Kosterman, Catalano, and Hawkins, 1999).

[7] Hawkins et al. (1999).

[8] Few people earn a traditional high school degree between age 21 and 24.

[9] See Heckman et al. (2014b) and the literature they review. Unfortunately, Hawkins et al. (1999, 2005, 2008) fail to distinguish high school graduates from GED recipients, a common failing in the literature.

Table 4. Summary of Treatment Effects from the Seattle Social Development Project

	Age			
Outcome	Age 18[h]	Age 21[i]	Age 24[j]	Age 27[j]
GPA	0.24*			
CAT(ES)[a]	0.05			
Grade Repetition(%)	−8.7**			
Dropout(%)	−7.3			
School Misbehavior[b]	−1.41**			
Violent Crime(%)	−11.4**			
Ever Arrested(%)	−6.0			
Arrested past year(%)		−2.0	1.0	1.0
Smoking(%)	−0.7			
Pregnancy(%)	−9.3*		−9.0	−8.0
Anxiety(%)[c]		−2.0		
Depression(%)[c]		−8.0*		
High School Graduate/GED(%)		10.0***	6.0	6.0
More than 2 Years of College(%)		8.0***		
Self-efficacy[d]		0.17***	0.13*	−0.01
Associate's Degree(%)			12.0*	12.0*
Bachelor's Degree(%)			7.0	6.0
Substance Abuse Index[e]			3.0	−3.0
Mental Health Disorder Index[f]			−9.0*	−11.0**
Income(in thousands)[g]			3.51	3.12

Note: The estimates are coefficients from regressions that control for sample member characteristics. % refers to treatment effects in terms of changes in prevalence in outcome variable in percentage points. ES indicates Effect Size, reporting the treatment effect in standard deviations of the outcome variable. *(a)* CAT stands for California Achievement Test score and combines reading, language, and mathematics subtests. It has been standardised based on the sample of ninth-grade Seattle students. *(b)* Measures frequency of occurrence of skipping, cheating, and being sent from class. *(c)* Anxiety, social phobia, and depression were measured using the *DSM-IV* (American Psychiatric Association, 1994). The mental health disorder index groups this for anxiety, social phobia, posttraumatic stress disorders, and major depressive episodes. *(d)* Measured as mean score on six items concerning perceived future opportunities, on a scale of 1-4. *(e)* The Substance Abuse Index measures dependence on substances (tobacco, alcohol, and illicit drugs) using *DSM-IV* criteria. *(f)* The Mental Health Disorder Index summarises problems of anxiety, social phobia, posttraumatic stress and depression, using *DSM-IV* criteria. *(g)* Refers to income from all sources, before taxes. Includes zero-earners; income is top-coded at $200,000. *(h)* Hawkins et al. (1999). *(i)* Hawkins et al. (2005). *(j)* Hawkins et al. (2008). * 10% significance; ** 5% significance; *** 1% significance.

The Montreal Longitudinal Experimental Study (MLES)

The Montreal Longitudinal Experimental Study (MLES) provides further evidence that non-cognitive skills can be shaped during elementary school.[10] MLES shared a number of similarities with SSDP. They both targeted similar ages, focused on behavioural skills training, involved parents, and had similar effects on outcomes. MLES was a two-year programme that lasted between ages seven and nine and targeted boys who had behavioural issues in kindergarten. The intervention consisted of 19 sessions that included a variety of social behaviour/skills training administered by professionals. It also included a parental training component that focused on fostering similar skills.

Algan et al. (2014) evaluate the intervention by the method of random assignment. They find that the programme improved a range of adult outcomes. It increased full time employment or school enrolment at ages 17-26 by 11 percentage points, and secondary school graduation by ages 23-24 by 19 percentage points. It reduced the probability of having a criminal record by ages 23-24 by 11 percentage points. They find that the programme also boosted measures of non-cognitive skills and grades during adolescence. They do not collect data on test scores and psychological measures, so it is not clear whether the

[10] See Algan, Beasley, Vitaro, and Tremblay (2014) for a detailed description of the program and study.

improvements in grades are due to changes in non-cognitive skills alone or whether it also improved cognitive skills (as measured by test scores).

Cambridge-Somerville Program

The Cambridge-Somerville Program is an example of a primary school programme that harmed its participants. This five-year intervention targeted five to thirteen year-old boys with behavioural problems. The programme was evaluated by the method of random assignment. The treatment differed among participants, but included medical assistance, tutoring, access to summer camps, and meetings with parents (McCord and McCord, 1959). A 30-year follow-up showed that participants exhibited worse drinking habits, more serious mental diseases, more heart problems, higher blood pressure, (modestly) higher crime rates, and died younger (McCord, 1978). Despite the negative effects, participants reported that they thought the programme helped them. McCord (1978) speculates that the programme failed because it did not create a sense of autonomy among its participants. When the intensive assistance was removed, the treatment group might have fallen back into unfavourable behaviours. Unlike SSDP and MLES, it did not actively engage parents and it did not target specific behavioural skills. These factors might have reduced its efficacy.

Project STAR

Project STAR provides evidence that higher quality kindergarten classes improve later-life outcomes by improving non-cognitive skills. From kindergarten through third grade, children and teachers were randomly assigned to classrooms of differing class sizes. In an analysis that builds on the earlier work of Heckman et al. (2013), Chetty et al. (2011) examine the Project STAR programme and find that students placed in higher quality kindergarten classes tend to have higher test scores at the end of kindergarten. They measure quality by peer average performance on the SAT. By eighth grade, the effect on test scores faded over time.

As with the Perry programme, benefits emerge later in life. Children placed in better kindergarten classrooms had significantly higher earnings in early adulthood. Furthermore, kindergarten classroom quality also predicts better fourth-and eighth-grade behaviour as measured by teacher-assessed effort, initiative, interest in the class, and disruptive behaviour. In turn, measured behaviour predicts earnings in adulthood, suggesting that improvement in non-cognitive skills is the main channel through which better kindergarten classrooms improve earnings. The studies of Perry and STAR demonstrate the importance of long-run follow-ups, and the importance of accounting for non-cognitive skills in evaluating intervention programmes.

EDUCATION AND INTERVENTIONS TARGETED TOWARD ADOLESCENTS AND YOUNG ADULTS

Compared to early-childhood programmes, the evidence on adolescent programmes is less abundant. There are, however, some well-documented and promising interventions. We divide adolescent programmes into four main categories: *(1)* mentoring programmes for at-risk students; *(2)* residential-based education programmes for high school dropouts; *(3)* in-school, professional training and *(4)* incentives for student performance. Several mentorship programmes and residential-based programmes appear promising but have relatively short follow-ups. Some- what similar programmes with long follow-ups have had no long-term effects. These programmes appear to achieve their short-term effects by closely monitoring participants and removing them from their usual environments.[1] Programmes that combine work and education are more promising and have been shown to have lasting effects. A few programmes that combine long-term mentoring with financial support to attend college have improved educational attainment but do not measure other outcomes.

Few programmes that target adolescents have long-term follow-ups. As Table 2 shows, only two evaluations of adolescent programmes have follow-ups of at least ten years, compared to 14 early-childhood and elementary school evaluations. Most adolescent intervention programmes measure fewer outcomes and focus mainly on schooling and employment.

Our framework on the life-cycle development of skills provides an interpretation of the relative failure of adolescent programmes. Early childhood is a sensitive period for the development of both cognitive and non-cognitive skills—prevention is more effective than remediation. Early investments are productive because early skills promote the development of later skills and, through dynamic complementarity, make later investments more productive. The finding that short-term benefits fade away quickly in most adolescent interventions might indicate the importance of incentives and situations in affecting behaviour. As previously noted, several adolescent programmes temporarily supervise and control the environments of adolescents, thereby changing their incentives which beneficially, but temporarily, affects their behaviour.

Adolescent Mentorship Programmes

Quantum Opportunity Program

The Quantum Opportunity Program (QOP) is one of the few adolescent mentorship pro- grams with a long-term follow-up. Rodríguez-Planas (2012) estimates the effect of the QOP on labour market outcomes, educational attainment, and risky behaviours. The programme was an intensive after-school programme that provided mentoring, educational services, and financial incentives during the four years of high school (plus one additional year in case students fell behind).

QOP was not successful in improving risky behaviours, and, if anything, caused harm for male participants. Ten years after the start of the programme, males in the treatment group were eight percentage

[1] These are called “incapacitation” effects in criminology.

points more likely to have been recently arrested. Males were more likely to have completed two years of college or training, but the increased educational attainment did not translate into higher earnings in the follow-up period. Female participants had better educational outcomes during the first follow-up near the end of the programme, but the control group caught up by the end of the three-year follow-up.

Rodríguez-Planas (2010) suggests several reasons why the programme might have failed. First, increasing awareness about the consequences of risky behaviour could have caused teens to become fatalistic; for example, some might feel that drug addiction was inevitable. Second, the mentors might have served as advocates when students were in trouble at school or with the law, lowering the cost of engaging in problem behaviours. Third, students might have used their stipends to purchase drugs or alcohol. Fourth, the programme organised group activities among the participants, which could have induced negative peer-effects.

Becoming a Man

A recent evaluation of Chicago's "Becoming a Man" (B.A.M.) programme provides evidence that mentorship programmes that target non-cognitive skills and are coupled with academic tutoring might improve cognitive and academic outcomes.[2] The programme starts in 9th and 10th grade and lasts for three-quarters of an academic year. It consists of two components. First, it targets "social-cognitive" (non-cognitive) skills through B.A.M. – a one-hour, weekly Cognitive Behavioural Training (CBT) implemented in groups. Second, it targets academic skills through daily, hour-long tutoring sessions in maths.

Cook et al. (2014) evaluate the programme by the method of random assignment. They find that during the year of the intervention, assignment to the programme was associated with a 0.5 standard deviation increase in maths achievement test scores, a 0.4 standard deviation increase in maths grades, and reduction of absences by 10 days.[3] It did not have an effect on reading scores or reading grades. While the programme is promising, the follow-up period is far too short to determine whether it has lasting effects on meaningful outcomes.

Its implementation differed from QOP in important ways that might explain why B.A.M. appears to be more successful. In contrast to QOP, it did not provide monetary incentives, the tutoring was in very small groups, and the larger group sessions were partly designed to reduce conflict and might have reduced negative peer effects.

Pathways to Education Programme

The Pathways to Education Program (Pathways) is an adolescent intervention that operates in housing projects of Toronto.[4] The programme is available throughout high school for residents living in particular neighbourhoods. It consists of four main components. First, students meet twice-a-week with a "Student-Parent Support Worker" (mentor) who works directly with students and parents and is the main interface between students and the pro- gram. Second, participants are required to attend tutoring sessions twice a week and can attend up to four times a week.[5] Third, in 9th and 10th grade, students are required to participate in group mentorship every two weeks. The group activities range from cooking to attending

[2] See Cook, Dodge, Farkas, Fryer, Guryan, Ludwig, Mayer, Pollack, and Steinberg (2014) for a detailed description of the program.

[3] The effect sizes are relative to the control group.

[4] See Oreopoulos, Brown, and Lavecchia (2014) for a detailed description of the program.

[5] Students with sufficiently high grades are exempt from the mandatory tutoring.

sports events. Fourth, students receive bus passes to help them attend school and financial assistance for college.

Oreopoulos et al. (2014) evaluate the programme non-experimentally by adopting a "difference-in-difference" strategy that compares students from neighbourhoods that offered Pathways to students from other similar neighbourhoods. They estimate that the availability of Pathways increased secondary school completion and post-secondary enrolment by about 15 percentage points. The programme also improves achievement test scores. Non-cognitive skills are not directly measured.

As with B.A.M., Pathways differed from QOP and other unsuccessful interventions in some key ways. First, the financial incentives were only usable for bus fare and college tuition, mitigating the possibility that they could be used to purchase drugs or alcohol as was a concern in QOP. Second, Pathways was introduced at the neighbourhood-level and the take-up rates were over 85%, so it did not group only the highest-risk students together which might have limited the role of negative peer effects. Third, it more actively engaged parents.

Big Brothers Big Sisters

The Big Brothers Big Sisters (BBBS) mentoring programme effectively improves educational outcomes (Tierney et al., 1995). The programme is aimed at children (aged 10-16) living in unstable family environments, which generally are single-parent households. Volunteer mentors have regular and lengthy meetings with the enrolees for an average treatment length of one year.

Tierney et al. (1995) evaluate the programme by comparing treated children to a control group that was put on a waiting list for the programme. BBBS has positive impacts on academic outcomes but only for girls. The effects on direct measures of non-cognitive skills related to Big Five Neuroticism (Self-Worth, Social Acceptance, and Self-Confidence) are positive but not statistically significant. Treatment group children are less likely to hit other children or lie to their parents. The follow-up lasted only 18 months after random assignment. Given the results from QOP, this follow-up is too short to determine whether the effects are persistent. However, BBBS differed in important ways from QOP. BBBS features one-on-one mentorship, which allows mentees and mentors to form stronger attachments and might help avoid the negative peer effects of grouping at-risk youth together.

Empresários Pela Inclusão Social (EPIS) Program

Martins (2010) analyses data from Empresários Pela Inclusão Social (EPIS), a programme developed to improve student achievement of 13- to 15-year-olds in Portugal. Unlike many other interventions, EPIS emphasises non-cognitive skills, aiming to increase motivation, self-esteem, and study skills. The programme consists of one-on-one meetings with a trained staff member or meetings in small groups. The intervention is tailored to each participant's individual skill deficit. EPIS decreases grade retention by 10 percentage points, in a cost- effective way. Other outcomes are not reported.

H&R Block FAFSA experiment

In this paper, we have primarily analysed interventions through the lens of skill development. Other interventions provide adolescents and young adults with specific information or assistance useful in attaining concrete goals or outcomes given a fixed skill base. Bettinger, Long, Oreopoulos, and Sanbonmatsu (2012) analyse an intervention that helped families complete financial aid applications and provided them with information about eligibility for financial aid.

In collaboration with a tax preparation service, Bettinger et al. (2012) identified a group of low-income families with adolescents or young adults who did not yet have a bachelor's degree. They assigned

a member of these families into a group that received help completing a financial aid form and an estimate of their college aid eligibility, a group that received only an estimate of their college aid eligibility, or a control group. They divide the sample into participants who are dependent on parents and independent adults. They find that dependents whose parents received help in completing financial aid forms and eligibility information were eight percentage points more likely to be enrolled in college within the year compared to the control group. No other groups benefited. This evidence suggests that relatively small interventions that provide very specific types of information or assistance to targeted groups can make a big difference. The follow-up period is too short to determine whether members of the control group will eventually catch up.

Dartmouth College Coaching Program

In a similar vein as the H&R Block FAFSA experiment, the "Dartmouth College Coaching Program" was a relatively short intervention that primarily consisted of providing information to students and helping them complete forms.[6] The programme targeted students in their last year of secondary school who were on the cusp of applying to college but had taken few steps in doing so. It provided these students with in-person mentorship, a waiver of their college entrance exam fees, and a $100 bonus for completing the programme. The mentorship component was designed around a series of tasks required for college admission including writing college essays, requesting transcripts, and completing the financial aid form.

Carrell and Sacerdote (2013) evaluate the intervention by the method of random assignment. They find that assignment to the treated group increased college enrolment of female students by about 15 percentage points but had no effect on male students.[7] In one cohort they also offered the control group $100 for completing college applications and estimated similar treatment effects, suggesting that the cash incentives played a small role.

Residential-Based Programmes

Job Corps

The Job Corps is the largest residential training programme in the United States for at-risk youth. Youth spend one year in training and remedial education, and receive counselling and training in social skills. Early evaluations found that the programme improved many outcomes including wages, welfare dependence, college attendance, health, and crime.[8] However, more recent evaluations report that gains in earnings fade and that the reduced crime effects came primarily during the programme and only for the residential participants (Schochet et al., 2001, 2008). The net earnings effects are basically zero.[9] The programme's emphasis on job search might explain the temporary increase in earnings after participants leave it.

National Guard ChalleNGe

The National Guard Youth ChalleNGe programme is a 17-month intervention for youth who dropped out of high school. Like the Job Corps, most training occurs in a residential facility removed from the usual

[6] See Carrell and Sacerdote (2013) for a detailed description of the program.

[7] The enrolment variable is based on whether the student ever enrolled in college during the sample period. They implemented the experiment for different cohorts so that the follow-up duration depends on the cohort.

[8] See, e.g., Mallar, Kerachsky, Thornton, and Long (1982).

[9] For males ages 20–24 there appear to be modest Job Corps effects. However, these findings do not survive adjustments for multiple hypothesis testing (Schochet et al. (2008)).

environments of participants. The programme features a two-week residential orientation and assessment period, a 20-week residential programme often conducted at a military base, and a one-year, non-residential mentoring programme (Bloom et al., 2009; Millenky et al., 2010, 2011).

Nine months after the programme, participants are more likely to have a GED or high school diploma and be employed. They are less likely to have been arrested or convicted. Three years after the intervention, the effects for criminal behaviour and high school graduation decline and become statistically insignificant. The initial reduction in crime likely occurs because participants are housed in a residential programme (an incapacitation effect). In-depth interviews of the participants suggest that it was difficult for participants to maintain the discipline that they had learned once they returned to their original environments (Millenky et al. (2011)). However, the employment and earnings effects persist. Given the similarities between the Job Corps and the ChalleNGe programme, the three-year follow-up might not be long enough to determine whether the labour market effects are persistent.

Workplace-Based Adolescent Intervention Programmes

Workplace-based internships and apprenticeships can boost skills.[10] One hundred and fifty years ago, apprenticeships and workplace-based education were standard for most adolescents. Apprenticeships offer in-depth, work-based learning combined with related academic course work. Few rigorous studies have examined how entering and completing apprenticeships in the United States affects the educational attainment, job skills, non-academic skills, and job market outcomes of young people.[11] A study of high school students who participated in a Wisconsin youth apprenticeship in printing documented participant earnings levels substantially above expected earnings for similar youth (Orr, Bloom, Bell, Lin, Cave, and Doolittle, 1994). Anecdotal evidence suggests that youth apprenticeships motivate participants to do better in school and to pursue difficult courses broadly related to their occupational interests.

Well-structured career-focused programmes can potentially enhance non-cognitive as well as occupational skills, in ways most relevant to the occupations and tasks that engage young people. Halpern (2009) undertook in-depth, observational studies of high school youth apprenticeships. Halpern's qualitative analysis of these programmes is rich and difficult to summarise. Nonetheless, several perspectives are worth noting. He gives many examples of how apprentices gradually develop expertise in an occupation as well as in problem-solving skills. They acquire self-confidence, teamwork, the ability to take direction and take initiatives, and other non-cognitive skills. He notes that participating youth see themselves as judged by their ability to meet or exceed the established standards of an occupation, including meeting deadlines and facing genuine constraints and unexpected difficulties that arise in the profession.[12] Adult mentors scaffold the discipline protocol for the apprentice, sequencing and controlling task demands to keep them on the constructive side of difficulty. They direct apprentices' attention, demonstrate new points, and sometimes collaborate with them.

[10] We have benefitted from the commentary of Robert Lerman in preparing this subsection. See Lerman (2013) for a discussion of related issues.

[11] Students involved in Wisconsin's Youth Apprenticeship Program believed their worksite learning expe- riences were excellent, with 84%–86% reporting that they enabled them to improve their problem-solving and teamwork skills (Scribner and Wakelyn, 1998). Students reported gaps between their work-based and school- based learning. But interviews with instructors and employers indicated that apprentices improved their so- cial and interpersonal skills, develop independent decision-making skills, and increased their self-confidence and self-esteem. However, one should be wary of self-reports. Students in the Cambridge-Somerville pro- gram reported satisfaction with the program, even though objective evaluations of it were negative (McCord, 1978).

[12] To quote Halpern, "Young people learn through observation, imitation, trial and error, and reiteration; in other words through force of experience."

Utilising workplaces as learning locations is motivated by evidence of the importance of occupational and non-cognitive skills. Workplace-oriented training can help youth attain development goals such as personal autonomy, efficacy, motivation, realism, optimism, and knowledge of vocations. It facilitates matching between workers and firms, and motivates adolescents to acquire relevant academic and non-cognitive skills. Evidence from other countries shows that workplace-based learning helps students develop an occupational identity, a professional ethic, and self-esteem based on accomplishment (Rauner, 2007). We review some of the better documented programmes in this section.

Career Academies

Career Academies provide evidence that integrating career development into a standard high school can have long-term labour market benefits. These academies operate within regular high schools and prepare students for postsecondary education and employment. In addition to regular high school courses, Career Academies offer courses that are career- or occupation-focused (Kemple and Willner, 2008). Individual academies focus on preparing students for participation in particular industries such as finance, health science, hospitality, and tourism. They include activities that might enhance non-cognitive skills in the work- place (Kemple and Snipes, 2000). Career Academies have many features in common with European apprenticeship programmes.

Compared to other high schools, Career Academies expose students to career-oriented activities, including job shadowing, career fairs, job search workshops, resume preparation, and interview preparation. Students work together in teams when relating their experiences. A recent evaluation based on random assignment estimates the effects of Career Academies on labour market, educational, and social outcomes (Kemple and Willner, 2008). Table 5 shows the differences in outcomes between the control and treatment groups. Career Academies have positive and sustained impacts on earnings and employment among young men but have no impact on high school completion rates and postsecondary educational enrolment and attainment rates.[13] Treatment group participants are more likely to live independently with children and a spouse/partner. Young males in the treatment group are more likely to get married and become custodial parents. For females, most effects are positive but not statistically significant.

Career Academies likely benefit their students by improving non-cognitive skills. The programme improves earnings for males in the long run, even though it does not improve educational attainment or scores on achievement tests. The Career Academies include activities that are designed to improve non-cognitive skills. Internships may teach students the importance of perseverance and Conscientiousness, along with other occupation-related skills that improve their labour market readiness. Programmes that integrate school and work not only motivate students to learn relevant academic material but also integrate adolescents into the larger society and teach children the skills valued in the workplace and in society at large.

[13] The percentage gains in earnings were highest for the students facing the highest risk of dropping out of school. Young women did not experience any statistically significant gains in earnings. Given that only 55% of the treatment group actually finished the full curriculum of the Career Academy, the earnings gains likely understate the actual impact of full participation.

Table 5. Summary of Treatment Effects from Career Academies within 96-Month Follow-Up after Scheduled High School Graduation

Outcome	Males	Females
Labor Market (49–96 Months)		
Monthly Earnings ($)[(a)]	361**	118
Months Employed (#)	2.8***	−0.3
Average Hours Worked per Week (#)	4.1***	0.2
Average Hourly Wages ($)	0.6	0.7
Educational Attainment (After 96 Months)		
High School Diploma	−0.4	0.2
GED	3.6	1.3
Certificate/License	2.0	0.1
AA Degree	−1.0	1.8
BA Degree	−2.2	−1.6
Family Formation (After 96 Months)		
Married and Living Together	9.0**	1.5
Custodial Parent	11.5***	3.7
Non-Custodial Parent	−6.4**	0.2

Note: Impact estimates are regression-adjusted to control for background characteristics of the sample and for the clustering of students within schools and random assignment years. (a) Non-workers were assigned a value of "0" for monthly earnings. * 10% significance; ** 5% significance; *** 1% significance.

Source: Kemple and Willner (2008).

Year-Up Programme

The Year-Up programme provides low-income, mostly black and Hispanic 18- to 24-year-olds with full-time classes and six-month internships. To enter the programme, applicants must submit a written essay, letters of reference, a high school transcript, and attendance records (Grobe, Rosenblum, and Weissman, 2010). Given the requirements, Year-Up applicants may be more motivated than their peers.

During the first six months of the programme, students take classes involving business writing, time management, teamwork, problem solving, and conflict resolution (Grobe et al. (2010)). To remain in the programme, students must maintain high attendance rates, be on time, and complete assignments (Roder and Elliot, 2011). Students who complete classes are placed in internships with companies (Roder and Elliot, 2011). Participants continue to attend a weekly class during their internships where they share experiences about their internships, work on their resumes, do mock job interviews, and look into postsecondary options. Students also learn technical skills

Roder and Elliott (2014) study the effect of Year-Up on labour market outcomes three years after the end of the programme. Qualified applicants on the waiting list were assigned to treatment or control groups. Three years after the programme ended, the treatment group has higher annual earnings, primarily due to their increased hourly wages. The employment rates are similar between the treatment and control groups. The earnings and wage benefits arise from improvements in job quality. Compared to the control group, more members in the treatment group find higher paying jobs in information technology and investment operations.

The three-year follow-up is too short to determine whether the programme has lasting effects, but some evidence suggests that the benefits might be more permanent than other similar programmes. The Job

Corps primarily improves earnings in the short run by increasing employment (Schochet et al. (2008)). In contrast, Year-Up increases hourly wages (Roder and Elliot, 2011). If these wage increases reflect increased human capital formation, the benefits likely last longer. About a quarter of the treatment group is hired by the employer sponsoring their internship.

The evaluation reports no direct measures of non-cognitive skills, but it does provide clues that the programme affected non-cognitive in ways that lead to better outcomes in the job market. Year-Up students were taught that appropriate behaviour, dress, and communication are important for success.

Self-Sufficiency Project

The Canadian Self-Sufficiency Project (SSP) provides direct evidence that working improves non-cognitive skills. New welfare enrolees were told that if they remained on welfare a year later, they would have the chance to participate in the SSP study. After a year, they were either assigned to a treatment group that would receive an earnings subsidy for three years if they worked full time or to a control group that would not.[14]

Assessing the effect of employment on skills is complicated by the possibility of reverse causality. Just as employment might affect skills, skills might affect employment. The SSP programme provides a way to study the effects of exogenous variation of employment status on skills. Gottschalk (2005) uses assignment status in the experiment as an instrumental variable for employment to estimate the effect of employment on Locus of Control (the extent to which people feel they have control over their lives).[15] After 36 months, those who received the subsidy were more likely to have an improved Locus of Control, suggesting that work can improve this non-cognitive skill. This finding supports the earlier evidence that work-based education can improve non-cognitive skills. Other studies evaluate the labour market benefits of the programme, but the findings are mixed.[16]

Apprenticeship Programmes[17]

Apprentices are employees who have formal agreements with employers to carry out a recognised programme of work-based and classroom learning. Apprenticeships typically begin in late adolescence or early adulthood. The training is highly structured, with a well-designed curriculum usually lasting three to four years. It is common for apprentices to spend three to four days per week at workplaces, acquiring knowledge and undertaking productive work involving gradually increasing levels of complexity. Classroom instruction in theoretical and general studies takes place at vocational schools one to two days per week. Apprentices participate in the production process, work with a trainer who is often a mentor, and ultimately gain sufficient occupational mastery to become certified by an external body (Wolter and Ryan, 2011). The emphasis is on occupational skills, but apprentices are likely to improve non-cognitive skills and other generic employment skills as well, including those most relevant for their chosen occupation. The training is employer-led and thus sensitive to market demands and less costly to the government than full-time schooling. Career Academies and apprenticeship programmes share many features in common. They train participants in the skills of specific occupations, supplement the training with general education, and provide guidance and mentoring for participants.

[14] See Michalopoulos et al. (2002) for a detailed description of the program.

[15] Locus of Control is related to Big Five Neuroticism. See Almlund et al. (2011) for a discussion.

[16] See, e.g., Card and Hyslop (2009); Michalopoulos et al. (2002); Zabel et al. (2010, 2013).

[17] The commentary and text of Robert Lerman was helpful in preparing this section. See Lerman (2013).

Apprenticeship training is common in many countries, reaching 55%–70% of youth in Austria, Germany, and Switzerland. The scale of apprenticeship varies widely across countries. Apprenticeships as a share of the total labour force are about 3.7%–3.9% in Australia and Germany, about 2.3%–2.5% in Canada and England, and 1.7% in France but only about 0.3% in the United States.[18]

Given the concern that apprenticeships train workers too narrowly and make apprentices less adaptable to a changing labour market, in evaluating apprenticeship programmes it is important to follow those completing apprenticeships over long time horizons. At a mini- mum, an evaluation has to find an appropriate comparison group, estimate costs to firms, workers, and the government, and take account of the heterogeneity of occupations used for apprenticeships.

Several studies of apprenticeship training in European countries find high rates of returns to participants, often in the range of 10%–15%.[19] Bougheas and Georgellis (2004) find that about 70% of apprentices were in full-time employment immediately after training. Cooke (2003) shows that apprentices do not initially earn more than other workers, but experience greater wage growth.

One recent study of apprenticeship training analyses small Austrian firms that eventually failed. It compares the wages of apprentices who started well before the firm failed (and thus who had a long period in apprenticeship training) with wages of apprentices who spent less time in training because the firm failed earlier in their apprenticeship period (Fersterer, Pischke, and Winter-Ebmer, 2008). Estimates suggest that for a three- to four-year apprenticeship, post-apprenticeship wages are 12%–16% higher than they would otherwise be. Since the worker's costs of participating in an apprenticeship are often minimal, the returns are high.

What about job and occupational mobility? Buechtemann, Schupp, and Soloff (1993) find that about 80% of apprentices join the firm where they apprenticed after completing the programme. In order to examine the adaptability of apprenticeship training, researchers (Clark and Fahr, 2002; Geel and Backes-Gellner, 2009) identified groups of skills that are required for each occupation. Changing occupations may not imply a significant change in the types of expertise required because of the overlap of skill requirements across occupations. While only 42% of apprentices stay in their initial occupation, nearly two-thirds remain with the apprentice occupation or another occupation requiring a similar mix of skills (Geel and Backes-Gellner, 2009). When apprentices move to another occupation that requires similar skills, their wages increase. Those trained in occupations with more specific skills are most likely to remain in their initial occupation or move to occupations within the same cluster.[20]

Germany has formally integrated apprenticeship programmes into its educational system with apparent success. In Germany, there are three types of secondary schools: low (Hauptschule), medium (Realschule), and high (Gymnasium). Students who graduate from any of these secondary schools qualify

[18] Apprenticeships have expanded rapidly in several countries, including Australia (rising from 157 000 in 1996 to 515,000 in 2012) and in England (rising from 53 000 in 1990 to 614 000 in 2012) (National Apprenticeship Service, 2012; National Centre for Vocational Education Research, 2012). Along with the increasing roles of apprenticeship in Australia and England have come initiatives to upgrade the quality of their programs (McDowell, Oliver, Persson, Fairbrother, Wetzlar, Buchanan, and Shipstone, 2011).

[19] See, e.g., Adda, Dustmann, Meghir, and Robin (2013); Bougheas and Georgellis (2004); Clark and Fahr (2002); Winkelmann (1996).

[20] Clark and Fahr (2002) also find that the wage penalty of departing from the occupation for which apprentices trained varies with the distance away from the original occupation and that there is no penalty at all from displacement into a somewhat related occupation. Among the former apprentices changing occupations, about two of five report using many or very many of the skills from their apprenticeship and another 20% used some of the skills.

to participate in the dual apprenticeship programme. Those who graduate from higher quality schools have more prestigious vocational options.[21] The apprenticeships typically last about three years during which participants spend one to two days per week in a part-time vocational training school. They work the remaining time (Franz and Soskice, 1995).

Harhoff and Kane (1997) argue that apprentices occupy a similar place in the German economy as high school graduates occupy in the American economy. They find that apprentices earn about 19% more than those completing ten or fewer years of school but 47% less than college graduates, whereas in the United States high school graduates earn about 23% more than those who complete ten or fewer years of school but 42% less than those who complete college. The estimates are similar over the life cycle.

Adda et al. (2013) estimate that participating in the German apprenticeship programme yields a 9%–10% annual rate of return. This estimate includes the return to non-cognitive skills which the authors do not isolate from the other benefits of the programme. It also includes the return to the education and job training received, which likely produces skills valued in the labour market.

Some American apprenticeships are also promising. Hollenbeck (2008) estimates rates of return to various types of job training in Washington State. He compares outcomes of apprentices to those of other workers with similar backgrounds and earnings histories. His results show that apprenticeship training yields higher returns than other training or postsecondary (mostly community college) occupational programmes. For apprenticeships, the social and governmental rates of return are large (over 20%) within the first two-and-a-half years after apprentices exit the programme.

Another, more recent study of apprenticeship in ten American states documents large and statistically significant earnings gains from participating in an apprenticeship (Reed, Yung- Hsu Liu, Kleinman, Mastri, Reed, Sattar, and Ziegler (2012)). The study estimates how the length of participation in an apprenticeship affects earnings, holding constant the pre-enrolment earnings of apprenticeship participants. Six years after starting a programme, the earnings of the average apprenticeship participant (average duration in an apprenticeship) are 40% higher than the earnings of nonparticipants with the same pre-apprenticeship history.

Workplace-based education is promising. It breaks down the artificial barrier between the culture of high schools and that of the rest of society documented in Coleman (1961). It motivates adolescents to learn relevant academic skills and gives mentoring and guidance to participants, many of whom come from disadvantaged families where such guidance is missing.

[21] See Franz and Soskice (1995) for a detailed description of the program.

OTHER CURRICULA THAT HAVE BEEN APPLIED TO MULTIPLE AGE GROUPS

Tools of the Mind

A random assignment evaluation of Tools of the Mind provides evidence that non-cognitive skills can be fostered. Tools of the Mind is a programme that attempts to teach preschool and early primary schoolchildren to regulate their social and cognitive behaviours. The curriculum encourages children to role-play and learn in groups with other children. In short-term follow-ups, several studies show that it improves classroom behaviour as well as executive functioning, defined as higher-level cognitive skills, including inhibitory control, working memory, and cognitive flexibility.[1] Similar findings are reported for the Montessori preschool curriculum (Lillard and Else-Quest, 2006).

However, the evidence on the programme is not all favourable.[2] There are few long-term evaluations of it. As noted by Heckman et al. (2013), the Perry curriculum and the Tools of the Mind curriculum have common intellectual influences, so the positive long-term evidence on Perry arguably transfers to forecasting the likely long-run effects of Tools of the Mind.

Studies that Teach the Incremental Theory of Intelligence

Several studies examine the effect of teaching students that capabilities are malleable and that learning can change the structure of the brain (the "incremental theory of intelligence"). Dweck (2007) presents an extensive discussion of this approach. The focus in this intervention is on promoting cognition, but the measures used to evaluate the interventions include achievement tests and grades that are also determined by non-cognitive skills. Some of this research also measures perseverance. The theory underlying the intervention is that children who believe that cognitive ability is a fixed trait might have little incentive to improve it.

The "mindset" interventions implement the logic of Figure 1. Effort and non-cognitive skills determine performance on tasks, including performance on grades and achievement tests. Motivating greater effort enhances performance, and the "mindset" intervention is one such motivational exercise. At issue is whether the enhanced effort, however motivated, has lasting consequences. Unfortunately, all of the follow-ups are short-term in nature, so the important question – do these motivational programmes have lasting consequences? – remains unanswered.[3]

Aronson, Fried, and Good (2002) evaluate an intervention in which college students were taught the incremental theory of intelligence. As a way to internalise the information, the participants were instructed to write a letter that explains the theory to a low-performing middle school student. At the end of the next academic quarter, the average GPA in the treatment group was modestly higher than in the control group. There was no follow-up.

---

[1] Barnett, Jung, Yarosz, Thomas, Hornbeck, Stechuk, and Burns (2008); Barnett, Yarosz, Thomas, and Hornbeck (2006); Bodrova and Leong (2001, 2007); Diamond, Barnett, Thomas, and Munro (2007).

[2] See the study by Farran, Lipsey, and Wilson (2011).

[3] Motivation research (Deci, 1971; Deci and Ryan, 1985) suggests that intrinsic rewards have longer lasting consequences than extrinsic rewards (e.g., payments or punishment).

Good, Aronson, and Inzlicht (2003) study an intervention in which seventh-grade students were taught the incremental theory of intelligence. At the end of the year, students who learned about the incremental theory of intelligence performed somewhat better on the maths and reading sections of the Texas Assessment of Academic Skills (TAAS) achievement test. The effect was strongest for girls.

Blackwell, Trzesniewski, and Dweck (2007) evaluate a program in which seventh graders are taught about the incremental theory of intelligence through a series of class sessions. The treatment group and control group participated in eight 25-minute classroom sessions that had similar course material about the physiology of the brain, study skills, and antistereotypic thinking. Only the treatment group was taught the incremental theory of intelligence. After one term, the treatment group improved its maths GPAs relatively more than the control group.

While these studies are promising, their follow-ups are very short. The longest follow-up is a year after the beginning of the programme. For example, it is possible that the interventions increased motivation in the short run, much like the studies in which children are incentivised using M&M candies. In addition, while the effects are statistically significant, they are relatively small. It is unknown whether the benefits outweigh the costs. None of the studies of the "mindset" intervention report a rate of return. The mechanism by which reported results are produced is not analysed.[4]

Prevention vs. Remediation

Targeted prevention is often more effective than remediation. However, it is not true that in all cases prevention is better than remediation. The untargeted early childhood programmes favoured by many policy advocates unwisely spend resources on children from strong family environments who already have substantial early investments. In doing so they generate dead weight loss. Even if the targeting of a prevention programme is efficient, remediation may still be a better strategy. Remediation programmes delay expenditure until problems manifest themselves. The effectiveness of a prevention strategy versus a remediation strategy depends, in part, on the relative benefits of conducting the respective programmes including the costs of detecting problems early in life and the errors that result from incorrectly treating children from perfectly normal families.[5]

[4] Hunt (2012), a leading authority on intelligence, provides a trenchant assessment of this group of interventions and in particular the work of Dweck.

[5] See Heckman (2014).

THE EFFECTS OF EDUCATION AND PARENTAL INVESTMENT ON COGNITIVE AND NON-COGNITIVE SKILL

In this section, we discuss other evidence on the effect of schooling and parental investment on promoting cognitive and non-cognitive skills. We largely draw on evidence from observational studies but many of them address causality.

Heckman, Stixrud, and Urzua (2006) and Heckman et al. (2014) estimate the effect of schooling in creating cognitive and non-cognitive skills, controlling for the problem of reverse causality that schooling may be caused by skills. Specifically, they estimate the effect of schooling on Self-Esteem and Locus of Control, non-cognitive skills related to Neuroticism.[1]

Figures 12 and 13 show the effect of education on measures of these skills: Schooling improves both non-cognitive and cognitive skills. In results that are not displayed, they show that these traits, in turn, boost a variety of labour market and social outcomes at age 30.[2]

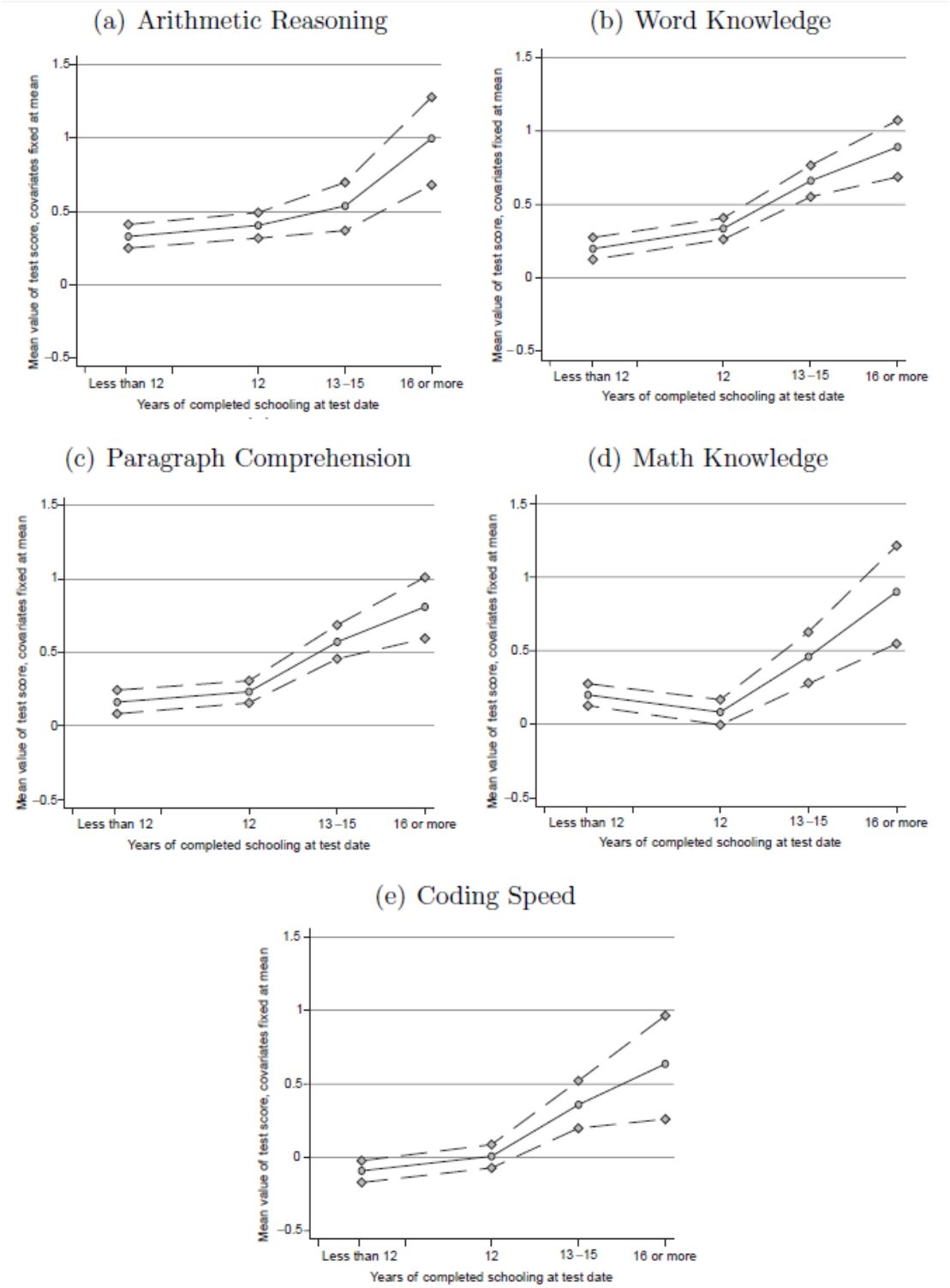
Cunha, Heckman, and Schennach (2010) estimate a causal model of parental investment using longitudinal data on the development of children with rich measures of parental investment and child skills. It estimates the model underlying Figure 9. They show that skills are self-productive and exhibit dynamic complementarity; levels of skills at one age affect the productivity of future investments at later ages and hence help determine the evolution of future skills through direct and cross effects. For example, children with higher levels of motivation learn more.

The authors find that it is more difficult to compensate for the effects of adverse environments on cognitive skills in adolescence than it is to build cognitive skills at earlier ages. Thus the early years are sensitive periods for cognitive skills. This finding is consistent with the high-rank stability of cognitive skills after age 11 or so; it also 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).

[1] The Rosenberg Self-Esteem Scale assesses the degree of approval or disapproval of oneself (Rosenberg, 1965). The relationship between these measures and the Big Five traits of Neuroticism is discussed in Almlund et al. (2011). See Table 1.

[2] Both Heckman, Humphries, and Veramendi (2014) and Heckman, Stixrud, and Urzu´a (2006) use an identification strategy employing matching on proxies for unobserved skills that corrects for measurement error and the endogeneity of schooling.

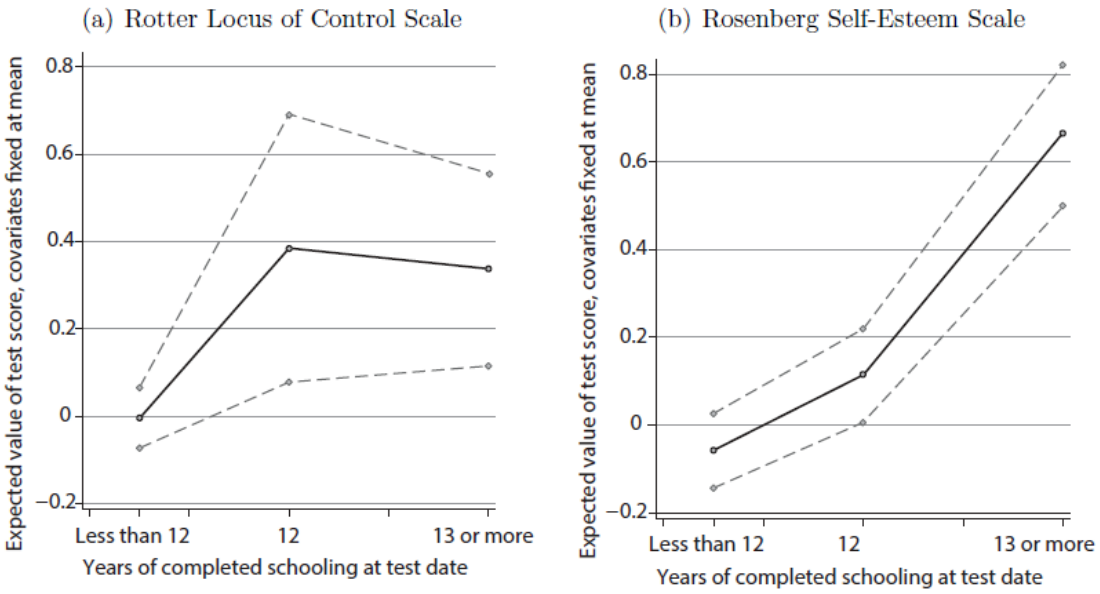
Figure 12. Casual Effect of Schooling on ASVAB Measures of Cognition



Note: Mean effect of schooling on components of the ASVAB. The first four components are averaged to create males with average ability. We standardise 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. Regressors are fixed at means.

Source: Heckman et al. (2006, Figure 4).

Figure 13. Casual Effect of Schooling on Two Measures of Non-Cognitive Skill



Note: Effect of schooling on socio-emotional 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 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 standardise 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. Regressors are fixed at means.

Source: Heckman et al. (2006, Figure 5).

Non-cognitive skills foster cognitive development. Greater malleability of non-cognitive skills is found over longer stretches of the life cycle than for cognitive skills. This occurs in part because new aspects of non-cognitive skills emerge with maturity and can be influenced.[3] Investment in non-cognitive skills in the early years has a higher economic return than investment in the later years because it builds the base for subsequent investment. Nonetheless, the productivity of later-age investment in non-cognitive skills is substantial. If the early years have been compromised, it is more effective in the adolescent years to focus on developing non-cognitive skills rather than on cognitive skills.[4]

Heckman et al. (2014) estimate a sequential model of education to study the effects of education on a variety of outcomes, controlling for cognitive and non-cognitive skills and the endogeneity of education. Correcting for selection into education, they find that early cognitive and non-cognitive skills have substantial causal effects on schooling, labour market outcomes, adult health, and social outcomes and that education at most levels causally produces benefits in labour market, health, and social outcomes. They also study the effects of schooling on measures of adult non-cognitive skills and find substantial impacts. They show that both cognitive and non-cognitive skills strongly determine schooling attainment but that

[3] Borghans et al. (2008); Roberts and Mroczek (2008); Roberts, Walton, and Viechtbauer (2006).

[4] Cunha, Heckman, and Schennach (2010) report that 16% of the variation in educational attainment is explained by adolescent cognitive skills, 12% is due to adolescent socioemotional traits (non-cognitive skills), and 15% is due to measured parental investments.

for many outcomes there are additional effects of both cognitive and non-cognitive skills on outcomes beyond their effects on education and the effect of education on outcomes.

Jackson (2013) estimates a valued-added model of the effect of ninth-grade maths and English teachers on student cognitive and non-cognitive skills as measured by absences, suspensions, grades, and grade progression.[5] Similar models have been adopted by school districts to assess the teacher's impact on student achievement test scores and to determine teacher bonuses. Jackson estimates a bigger effect of teachers on non-cognitive skills than on cognitive skills.[6] Estimated teacher effects on cognitive and non-cognitive skills are only weakly correlated with each other, suggesting that some teachers improve one dimension of skill without improving the other. These findings imply that using achievement tests alone to assess teacher effectiveness misses important dimensions of teacher quality.

Coleman, Hoffer, and Kilgore (1982), Neal (1997), and Bryk, Lee, and Holland (1993) document the benefits of Catholic education in terms of scores on achievement tests.[7] It has long been speculated that Catholic schools mold non-cognitive skills through creating a more disciplined environment and through teaching moral values. However, to our knowledge, there are no direct estimates of the effects of religious education on non-cognitive skills. Typical of the entire literature on evaluation of schools, these studies focus on the effects of Catholic schools on test scores and schooling.

Arum (2005) documents the decline of disciplinary practices in American schools arising in part from legal and legislative developments originating in the 1960s that increased the rights of students and circumscribed teachers from using traditional measures of enforcing discipline. He cites evidence showing that greater discipline improves test scores and behaviour and promotes high school graduation.[8]

Some modern schools have shown how it is possible to integrate non-cognitive skill education into schools. The Knowledge Is Power Program (KIPP) is a group of public charter middle schools that are designed to improve educational outcomes for low-income families. Tough (2012) discusses how the first wave of KIPP students excelled at taking achievement tests but dropped out of college at disappointingly high rates. In response, KIPP schools started to emphasis

e non-cognitive skills, "grit, self-control, social intelligence (including self-advocacy), zest, optimism, and gratitude, that enable students to stick with college even in the face of considerable obstacles" (The KIPP Foundation, 2011). Their motto is: Work hard. Be nice.

Evidence on the efficacy of KIPP is lacking. An evaluation of KIPP schools suggests that it has improved achievement test scores three years after students enter the programme.[9] A recent report from KIPP found that 33% of the students graduate from a four-year college, which is less than their goal of 75% but is four times higher than the rate for students with comparable backgrounds (The KIPP Foundation, 2011). However, these evaluations are based on comparing KIPP students to students in other schools; some have argued that KIPP students have higher levels of starting achievement and that KIPP selects more motivated students and parents into their schools.[10] A recent randomised trial evaluation of

[5] See, e.g., the Teacher Advancement Program in Chicago Public Schools (Glazerman and Seifullah, 2012).

[6] The estimates are from teacher "fixed-effect" models.

[7] Neal (1997) corrects for the selection effect that Catholic schools can expel unruly and undisciplined students. Uncontrolled comparisons could show a spurious positive effect of Catholic schools that do not reflect their true effect on any given person.

[8] DiPrete, Muller, and Shaeffer (1981).

[9] Tuttle, Teh, Nichols-Barrer, Gill, and Gleason (2010).

[10] Carnoy, Jacobsen, Mishel, and Rothstein (2005).

KIPP (Tuttle, Gill, Gleason, Knechtel, Nichols-Barrer, and Resch, 2013) found that KIPP boosted achievement test scores but had mixed effects on non-cognitive skills. For example, participants in KIPP were more likely to lie to their parents or quarrel with them (Tuttle et al., 2013).

Similar to KIPP, a recent intervention in Chicago attempts to foster non-cognitive skills. The OneGoal programme selects and trains high school teachers to help students apply to colleges, improve grades and test scores, and, importantly, persist through college. It serves low-income schools in Chicago, most of which have college enrolment rates of less than 50%. Like KIPP, it attempts to cultivate non-cognitive skills (Tough, 2012). So far, OneGoal appears to be successful: ninety-four percent of OneGoal participants attend college, and 85% of the college enrollees complete the first year.[11] In preliminary work, Kautz and Zanoni (2014) use non- experimental methods to evaluate OneGoal. They find that OneGoal selects students with near average cognitive skill but with above average non-cognitive skill. Nevertheless, after controlling for pre-existing skill, OneGoal appears to improve academic indicators in high school and boost high school graduation and college enrolment.

While there have been no published long-term evaluations of KIPP or OneGoal that account for selection bias, these programmes show that it is possible to openly emphasise non-cognitive skills in schools without infringing on the rights of students or families or violating the boundaries between church and state. These efforts can be viewed as secular versions of moral and non-cognitive skill education that do not introduce religion into schools.

[11] See http://www.onegoalgraduation.org/onegoal-results/. The OneGoal students are not old enough to estimate college graduation rates.

SUMMARY

This paper reviews the recent literature on the economics and psychology of non-cognitive skills and how interventions can develop non-cognitive skills. Non-cognitive skills predict later-life outcomes with the same, or greater, strength as measures of cognition. They have strong effects on educational attainment but have additional effects on important life outcomes beyond their effects on schooling.

Non-cognitive skills can be enhanced, and there are proven and effective ways to do so. Non-cognitive skills are shaped by families, schools, and social environments. At any age, non-cognitive skills are stable across different tasks, but performance on any task depends on multiple skills as well as the effort expended on it. Effort, in turn, depends on incentives to perform the task. Since all measures of cognitive and non-cognitive skills are measures of performance on some task, it is necessary to standardise for incentives, effort, and other skills in measuring any particular non-cognitive or cognitive skill, although this is not yet widely done. Despite these difficulties, reliable measures of non-cognitive skills have been developed, and there is an active literature on improving what is available.

Though stable at any age, skills are not set in stone over the life cycle. Both cognitive and non-cognitive skills can change. Parents, schools, and social environments shape them, although there are important genetic influences. Skill development is a dynamic process. The early years are important in laying the foundation for successful investment in the later years. While there is hard evidence on the importance of the early years in shaping all skills, many non-cognitive skills are more malleable than cognitive skills at later ages.

This paper reviews a variety of interventions targeted to different stages of the life cycle. Long-term evaluations of early childhood programmes are more numerous. There is substantial evidence that high-quality early childhood programmes have lasting and beneficial effects on non-cognitive skills. The evidence on interventions in elementary schools shows lasting benefits of interventions that primarily operate through boosting non-cognitive skills.

There are few long-term evaluations of adolescent interventions. The available evidence suggests a much greater benefit from programmes that target non-cognitive skills compared to the benefits of programmes that mainly target cognition and academic learning. Workplace-based programmes that teach non-cognitive skills appear to be the most effective remediation interventions for adolescents. They motivate acquisition of work-relevant skills and provide discipline and guidance for disadvantaged youth that is often missing in their homes or high schools. Successful interventions at any age emulate the mentoring and attachment that successful families give their children. Recent work suggests that programmes that provide information can be beneficial in promoting educational attainment.

The available evidence suggests that the most successful adolescent remediation programmes are not as effective as the most successful early childhood and elementary school programmes. The successful early childhood programmes are targeted toward disadvantaged children. Building an early base of skills that promote later-life learning and engagement in school and society appears to be a good strategy. Based on the available evidence, early childhood programmes tend to have higher rates of return than adolescent programmes. It is not, however, necessarily true that prevention is better than remediation.[1] The evidence to date may be an artefact of the particular types of adolescent programmes evaluated thus far.

[1] See Heckman (2014) for one analysis.

REFERENCES

Abbott, R. D. et al. (1998), "Changing teaching practices to promote achievement and bonding to school", *American Journal of Orthopsychiatry 68* (4), pp.542–552.

Achenbach, T. M. and L. A. Rescorla (2001), *Manual for ASEBA School-Age Forms and Profiles*. Burlington, VT: Research Center for Children, Youth and Families.

Ackerman, P. L. and E. D. Heggestad (1997). "Intelligence, Personality, and Interests: Evidence for overlapping traits", *Psychological Bulletin 121* (2), pp.219–245.

Adda, J. et al. (2013). Career progression, economic downturns, and skills. Unpublished manuscript, Institute for Fiscal Studies, London, UK.

Algan, Y. et al. (2014, February). The long-term impact of non-cognitive skills training at school entry. Unpublished manuscript, Sciences Po.

Almlund, M. et al. (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, and L. Wößmann (Eds.), *Handbook of the Economics of Education*, Volume 4, pp. 1–181. Amsterdam: Elsevier.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV* (4 ed.). Washington, DC: American Psychiatric Association.

Aos, S. et al. (2004). Benefits and costs of prevention and early intervention programs for youth. Report 04-07-3901, Washington State Institute for Public Policy, Olympia, WA.

Aronson, J., C. B. Fried, and C. Good (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology 38* (2), 113–125.

Arum, R. (2005). *Judging School Discipline: The Crisis of Moral Authority*. Cambridge, MA: Harvard University Press.

Ayllon, T. and K. Kelly (1972). Effects of reinforcement on standardised test performance. *Journal of Applied Behaviour Analysis 5* (4), 477–484.

Barnett, W. S. et al. (2008). Educational effects of the Tools of the Mind curriculum: A randomized trial. *Childhood Research Quarterly 23* (3), 299–313.

Barnett, W. S. et al. (2006). Educational effectiveness of a Vygotskian approach to preschool education: A randomized trial. Technical report, National Institute for Early Education Research, Rutgers, The State University of New Jersey.

Barrick, M. R. and M. K. Mount (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology 44* (1), 1–26.

Barton, P. (2006). *High School Reform and Work: Facing Labour Market Realities*. Princeton, NJ: Educational Testing Service.

Becker, A. et al. (2012, July). The relationship between economic preferences and psychological personality measures. *Annual Review of Economics 4*, 453–478.

Benda, B. B. (2005). The robustness of self-control in relation to form of delinquency. *Youth & Society 36* (4), 418–444.

Besharov, D. J. et al. (2011). Houston Parent-Child Development Center. Working Paper 17, University of Maryland, School of Public Policy, Welfare Reform Academy.

Bettinger, E. P. et al. (2012). The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *Quarterly Journal of Economics 127* (3), 1205–1242.

Blackwell, L. S., K. H. Trzesniewski, and C. S. Dweck (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development 78* (1), 246–263.

Bloom, D., A. Gardenhire-Crooks, and C. L. Mandsager (2009). Reengaging high school dropouts: Early results of the National Guard Youth ChalleNGe program evaluation. Report, MDRC. Last accessed online February 11, 2013. http://www.mdrc.org/sites/default/files/full_491.pdf.

Bodrova, E. and D. J. Leong (2007). *Tools of the Mind: The Vygotskian Approach to Early Childhood Education*. Upper Saddle River, NJ: Pearson Education, Inc.

Bodrova, E. and D. J. Leong (2001). *Tools of the Mind: A Case Study of Implementing the Vygotskian Approach in American Early Childhood and Primary Classrooms.* Geneva: International Bureau of Education, UNESCO.

Borghans, L., H. Meijers, and B. ter Weel (2008, January). The role of non-cognitive skills in explaining cognitive test scores. *Economic Inquiry 46* (1), 2–12.

Borghans, L. et al. (2011a). Identification problems in personality psychology. *Personality and Individual Differences 51* (3: Special Issue on Personality and Economics), 315–320.

Borghans, L. et al. (2011b). Reinterpreting estimated effects of cognition on social outcomes. Unpublished manuscript, Department of Economics, University of Chicago.

Borghans, L. et al. (2009, April). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association 7* (2–3), 649–658.

Borghans, L. et al. (2008, Fall). The economics and psychology of personality traits. *Journal of Human Resources 43* (4), 972–1059.

Bouchard, T. J. and J. C. Loehlin (2001, May). Genes, evolution and personality. *Behaviour Genetics 31* (3), 243–273.

Bougheas, S. and Y. Georgellis (2004). Early career mobility and earnings profiles of German apprentices: Theory and empirical evidence. *Labour 18* (2), 233–263.

Bound, J., C. Brown, and N. Mathiowetz (2001). Measurement error in survey data. In J. J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5 of *Handbooks in Economics*, Chapter 59, pp. 3705–3843. Amsterdam: Elsevier Science.

Bowen, W. G., M. M. Chingos, and M. S. McPherson (2009). Test scores and high school grades as predictors. In *Crossing the Finish Line: Completing College at America's Public Universities*, pp. 112–133. Princeton, NJ: Princeton University Press.

Bowlby, J. (1951). Maternal Care and Mental Health. Bulletin, World Health Organization, Geneva, Switzerland.

Bowles, S. and H. Gintis (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.

Bowles, S., H. Gintis, and M. Osborne (2001, December). The determinants of earnings: A behavioural approach. *Journal of Economic Literature 39* (4), 1137–1176.

Boyle, S. H. et al. (2005). Hostility, age, and mortality in a sample of cardiac patients. *American Journal of Cardiology 96* (1), 64–66.

Breitmayer, B. J. and C. T. Ramey (1986, October). Biological nonoptimality and quality of postnatal environment as codeterminants of intellectual development. *Child Development 57* (5), 1151–1165.

Breuning, S. E. and W. F. Zella (1978). Effects of individualized incentives on norm- referenced IQ test performance of high school students in special education classes. *Journal of School Psychology 16* (3), 220–226.

Bridgeman, B., J. B. Blumenthal, and S. R. Andres (1981). Parent Child Development Center: Final evaluation report. Report, ETS Research.

Bryk, A. S., V. E. Lee, and P. B. Holland (1993). *Catholic Schools and the Common Good*. Cambridge, MA: Harvard University Press.

Buechtemann, C. F., J. Schupp, and D. J. Soloff (1993). Roads to work: school-to-work transition patterns in Germany and the United States. *Industrial Relations Journal 24* (2), 97–111.

Campbell, F. et al. (2014). Early childhood investments substantially boost adult health. *Science 343* (6178), 1478–1485.

Campbell, F. A. et al. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science 6* (1), 42–57.

Campbell, F. A. et al. (2001, March). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology 37* (2), 231–242.

Canli, T. (2006). *Biology of Personality and Individual Differences*. New York: Guilford Press.

Card, D. and D. R. Hyslop (2009). The dynamic effects of an earnings subsidy for long- term welfare recipients: Evidence from the Self-Sufficiency Project applicant experiment. *Journal of Econometrics 153* (1), 1–20.

Carneiro, P. and R. Ginja (2012, January). Long-term impacts of compensatory preschool on health and behaviour: Evidence from Head Start. IZA Discussion Paper 6315, Institute for the Study of Labour (IZA).

Carneiro, P. and R. Ginja (2011, January). Preventing behaviour problems in childhood and adolescence: Evidence from Head Start. Unpublished manuscript, University College London.

Carneiro, P. and J. J. Heckman (2003). Human capital policy. In J. J. Heckman, A. B. Krueger, and B. M. Friedman (Eds.), *Inequality in America: What Role for Human Capital Policies*?, pp. 77–239. Cambridge, MA: MIT Press.

Carnoy, M. et al. (2005). *The Charter School Dust-up*. Washington, DC: Economic Policy Institute.

Carrell, S. E. and B. Sacerdote (2013). Late interventions matter too: The case of college coaching New Hampshire. Working Paper 19031, National Bureau of Economic Research.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies.* New York: Cambridge University Press.

Chetty, R. et al. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Quarterly Journal of Economics 126* (4), 1593–1660.

Clark, D. and R. Fahr (2002, February). The promise of workplace training for non-college bound youth: Theory and evidence from German apprenticeship. CEP Discussion Paper 518, Centre for Economic Performance, London School of Economics and Political Science, London, UK.

Clingman, J. and R. L. Fowler (1976). The effects of primary reward on the I.Q. performance of grade-school children as a function of initial I.Q. level. *Journal of Applied Behaviour Analysis 9* (1), 19–23.

Cloninger, C. R. et al.(1999). Measurement of psychopathology as variants of personality. In C. R. Cloninger (Ed.), *Personality and Psychopathology*, pp. 33–66. Washington, DC: American Psychiatric Press.

Coleman, J. S. (1961). *The Adolescent Society: The Social Life of the Teenager and Its Impact on Education. New* York: The Free Press of Glencoe.

Coleman, J. S., T. Hoffer, and S. Kilgore (1982). *High School Achievement: Public, Catholic, and Private Schools Compared.* New York: Basic Books.

Conti, G. et al. (2014). The long-term health effects of early childhood interventions. Under revision, *Economic Journal.*

Cook, P. J. et al. (2014, January). The (surprising) efficacy of academic and behavioural intervention with disadvantaged youth: Results from a randomized experiment in Chicago. Working Paper 19862, National Bureau of Economic Research (NBER).

Cooke, L. (2003). A comparison of initial and early life course earnings of the German secondary education and training system. *Economics of Education Review 22* (1), 79–88.

Costa, P. T. and R. R. McCrae (1992a). Four ways five factors are basic. *Personality and Individual Difference 13* (6), 653–665.

Costa, P. T. and R. R. McCrae (1992b). *Revised NEO Personality Inventory (NEO PI- R) and the NEO Five-Factor Inventory (NEO-FFI) Professional Manual.* Odessa, FL: Psychological Assessment Resources.

Cunha, F. and J. J. Heckman (2009, April). The economics and psychology of inequality and human development. *Journal of the European Economic Association 7* (2–3), 320–364. Presented as the Marshall Lecture, European Economics Association, Milan, Italy, August 29, 2008.

Cunha, F. and J. J. Heckman (2008, Fall). Formulating, identifying and estimating the technology of cognitive and non-cognitive skill formation. *Journal of Human Resources 43* (4), 738–782.

Cunha, F. and J. J. Heckman (2007, May). The technology of skill formation. *American Economic Review 97* (2), 31–47.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010, May). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78* (3), 883–931.

Cunha, F. et al. (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Chapter 12, pp. 697–812. Amsterdam: North-Holland.

Currie, J. and D. Thomas (1995, June). Does Head Start make a difference? *American Economic Review 85* (3), 341–364.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology 18*, 105–115.

Deci, E. L. and R. M. Ryan (1985). *Intrinsic Motivation and Self-Determination in Human Behaviour*. New York: Plenum Press.

DeLong, J. and K. Magin (2009, Winter). The U.S. equity return premium: Past, present and future. *Journal of Economic Perspectives 23* (1), 193–208.

Deming, D. (2009, July). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics 1* (3), 111–134.

Devlin, B., M. Daniels, and K. Roeder (1997). The heritability of IQ. *Nature 388* (6641), 468–471.

DeYoung, C. G. et al. (2010). Testing predictions from personality neuroscience: Brain structure and the Big Five. *Psychological Science 21* (6), 820–828.

Diamond, A. et al. (2007). Preschool program improves cognitive control. *Science 318* (5855), 1387–1388.

DiPrete, T. A., C. Muller, and N. Shaeffer (1981). *Discipline and order in American high schools.* Washington, DC: National Center for Education Statistics. Contractor Report from NORC.

Duckworth, A. (2012, November). Reference bias. Presented at the University of Chicago.

Duckworth, A. L. et al. (2011). Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences 108* (19), 7716–7720.

Duckworth, A. L. et al. (2007, June). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology 92* (6), 1087–1101.

Durlak, J. A. et al. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development 82* (1), 405–432.

Dweck, C. S. (2007). *Mindset: The New Psychology of Success.* New York, NY: Ballantine Books.

Eckenrode, J. et al. (2010, January). Long-term effects of prenatal and in- fancy nurse home visitation on the life course of youths: 19-year follow-up of a randomized trial. *Journal of the American Medical Association 164* (1), 9–15.

Edlund, C. V. (1972). The effect on the behaviour of children, as reflected in the IQ scores, when reinforced after each correct response. *Journal of Applied Behaviour Analysis 5* (3), 317–319.

Engle, P. L. et al. (2011, October). Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *The Lancet 378* (9799), 1339–1353.

Engle, P. L. et al. (2007, January). Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet 369* (9557), 229–242.

Epstein, S. (1979). The stability of behaviour: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology 37* (7), 1097–1126.

Farran, D. C., M. W. Lipsey, and S. Wilson (2011, August). Experimental evaluation of the Tools of the Mind pre-k curriculum. Technical report, Peabody Research Institute Report.

Fersterer, J., J.-S. Pischke, and R. Winter-Ebmer (2008). Returns to apprenticeship training in Austria: Evidence from failed firms. *Scandinavian Journal of Economics 110* (4), 733–753.

Flynn, J. R. (2007). *What is Intelligence?: Beyond the Flynn Effect.* New York: Cambridge University Press.

Franz, W. and D. W. Soskice (1995). The German apprenticeship system. In *Institutional Frameworks and Labour Market Performance*, Chapter 8, pp. 208–234. New York, NY: Routle.

Friedman, D. and A. H. Streicher (1985). Reliability of scores for Fiscal Year 1981 Army applicants: Armed Services Vocational Aptitude Battery forms 8, 9, and 10. Technical report, Defense Technical Information Center.

Friedman, H. S. and L. R. Martin (2011). *The Longevity Project: Surprising Discoveries for Health and Long Life from the Landmark Eight-Decade Study.* New York, NY: Hudson Street Press.

Garces, E., D. Thomas, and J. Currie (2002, September). Longer-term effects of Head Start. *American Economic Review 92* (4), 999–1012.

Geel, R. and U. Backes-Gellner (2009). Occupational mobility within and between skill clusters: An empirical analysis based on the skill-weights approach. Economics of Education Working Paper Series 0047, University of Zurich, Institute for Strategy and Business Economics (ISU).

Gensowski, M. (2014). Personality, IQ, and lifetime earnings. Discussion Paper 8235, IZA.

Gertler, P. and S. Martinez (2012). Vocational and life skills in youth training: A randomized experiment in the Dominican Republic. Presented at the IZA/World Bank/OECD Conference on Activation and Employment Support Policies, Bahcesehir University, Istanbul, Turkey, May 01, 2012.

Gertler, P. et al. (2014). Labour market returns to an early childhood stimulation intervention in Jamaica. *Science 344* (6187), 998–1001.

Gertler, P. et al. (2013). Labour market returns to early childhood stimulation: A 20-year follow-up to an experimental intervention in Jamaica. Working Paper 19185, NBER.

Glazerman, S. and A. Seifullah (2012). An evaluation of the Chicago Teacher Advancement Program (Chicago TAP) after four years. Final Report 06736-520, Mathematica Policy Research.

Goldammer, C. (2010). Skills, investment, and race: Evidence from the ECLS-K. Unpublished manuscript, University of Chicago, Department of Economics.

Goldin, C. and L. F. Katz (2008). *The Race between Education and Technology*. Cambridge, MA: Belknap Press of Harvard University Press.

Goldschmidt, P. and D. Huang (2007). The long-term effects of after-school programming on educational adjustment and juvenile crime: A study of the LA's BEST after-school program. Technical report, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles, CA.

Good, C., J. Aronson, and M. Inzlicht (2003). Improving adolescents' standardised test performance: An intervention to reduce the effects of stereotype threat. *Applied Developmental Psychology 24* (6), 645–662.

Gottfredson, M. R. and T. Hirschi (1990). *A General Theory of Crime*. Palo Alto, CA: Stanford University Press.

Gottschalk, P. (2005). Can work alter welfare recipients' beliefs? *Journal of Policy Analysis and Management 24* (3), 485–498.

Gough, H. G. and A. B. Heilbrun (1983). *The Adjective Check List Manual*. Palo Alto, CA: Consulting Psychologists Press.

Grantham-McGregor, S. M. et al. (2007). Developmental potential in the first 5 years for children in developing countries. *The Lancet 369* (9555), 60–70.

Grantham-McGregor, S. M. et al. (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: The Jamaican study. *The Lancet 338* (8758), 1–5.

Grek, S. (2009). Governing by numbers: The PISA 'effect' in Europe. *Journal of Education Policy 24* (1), 23–37.

Griffin, M., J. J. Heckman, and S. H. Moon (2013). Preschool curricula. Unpublished manuscript, University of Chicago.

Grobe, T., E. Rosenblum, and T. Weissman (2010). *Dollars and Sense: How "Career First" Programs Like Year Up Benefit Youth and Employers.* Boston, MA: Jobs for the Future.

Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics 19* (3), 403–420.

Haggerty, K. et al. (1999). Preparing for the drug free years. Juvenile justice bulletin, US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention, Washington, DC.

Halpern, R. (2009). *The Means to Grow Up: Reinventing Apprenticeship as a Developmental Support in Adolescence* (1 ed.). Critical youth studies. New York, NY: Routledge.

Hansen, K. T., J. J. Heckman, and K. J. Mullen (2004, July–August). The effect of schooling and ability on achievement test scores. *Journal of Econometrics 121* (1–2), 39–98.

Harhoff, D. and T. J. Kane (1997). Is the German apprenticeship system a panacea for the US labour market? *Journal of population economics 10* (2), 171–196.

Hawkins, J., E. Von Cleve, and R. Catalano Jr (1991). Reducing early childhood aggression: Results of a primary prevention program. *Journal of the American Academy of Child & Adolescent Psychiatry 30* (2), 208–217.

Hawkins, J. D. et al. (2008). Effect of social development intervention in childhood 15 years later. *Archives of Pediatrics and Adolescent Medicine 162* (12), 1133–1141.

Hawkins, J. D. et al. (2005). Promoting positive adult functioning through social development intervention in childhood. *Archives of Pediatrics and Adolescent Medicine 159* (1), 25–31.

Hawkins, J. D. et al. (1999). Preventing adolescent health-risk behaviours by strengthening protection during childhood. Archives of Pediatrics and Adolescent Medicine 153 (3), 226–234.

Heckman, J. J. (2014). Prevention versus remediation. Unpublished manuscript, University of Chicago.

Heckman, J. J. (2008, July). Schools, skills and synapses. *Economic Inquiry 46* (3), 289–324.

Heckman, J. J., J. E. Humphries, and T. Kautz (Eds.) (2014a). *The Myth of Achievement Tests: The GED and the Role of Character in American Life.* Chicago: University of Chicago Press.

Heckman, J. J., J. E. Humphries, and T. Kautz (2014b). The economic and social benefits of GED certification. In J. J. Heckman, J. E. Humphries, and T. Kautz (Eds.), *The Myth of Achievement Tests: The GED and the Role of Character in American Life.* Chicago: University of Chicago Press.

Heckman, J. J., J. E. Humphries, and T. Kautz (2014c). Who are the GEDs? In J. J. Heckman, J. E. Humphries, and T. Kautz (Eds.), *The Myth of Achievement Tests: The GED and the Role of Character in American Life.* Chicago: University of Chicago Press.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2014). Education, health and wages. Unpublished manuscript, University of Chicago, Department of Economics.

Heckman, J. J. and T. Kautz (2014a). Fostering and measuring skills: Interventions that improve character and cognition. In J. J. Heckman, J. E. Humphries, and T. Kautz (Eds.), *The Myth of Achievement Tests: The GED and the Role of Character in American Life*, pp. 341–430. Chicago, IL: University of Chicago Press.

Heckman, J. J. and T. Kautz (2014b). Fostering and measuring skills: Interventions that improve character and cognition. Technical report, IZA Discussion Paper No. 7750. J. Heckman, J.E. Humphries, and T. Kautz (eds.), The Myth of Achievement Tests: The GED and the Role of Character in American Life, Chicago: University of Chicago Press, 2014.

Heckman, J. J. and T. Kautz (2014c). Achievement tests and the role of character in American life. In J. J. Heckman, J. E. Humphries, and T. Kautz (Eds.), *The Myth of Achievement Tests: The GED and the Role of Character in American Life.* Chicago: University of Chicago Press.

Heckman, J. J. and T. Kautz (2012, August). Hard evidence on soft skills. *Labour Economics 19* (4), 451–464. Adam Smith Lecture.

Heckman, J. J. and P. A. LaFontaine (2010, May). The American high school graduation rate: Trends and levels. *Review of Economics and Statistics 92* (2), 244–262.

Heckman, J. J. and D. V. Masterov (2007). The productivity argument for investing in young children. *Review of Agricultural Economics 29* (3), 446–493.

Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annual Review of Economics 6* (1).

Heckman, J. J., S. H. Moon, and R. Pinto (2014). The effects of early intervention on abilities and social outcomes: Evidence from the Carolina Abecedarian study. Unpublished manuscript, University of Chicago.

Heckman, J. and R. Pinto (2015). Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs. Forthcoming, *Econometric Reviews.*

Heckman, J. J., R. Pinto, and P. A. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review 103* (6), 1–35.

Heckman, J. J., J. Stixrud, and S. Urzua (2006, July). The effects of cognitive and noncog- nitive abilities on labour market outcomes and social behaviour. *Journal of Labour Economics 24* (3), 411–482.

Heckman, J. J. et al. (2011). The effects of educational choices on labour market, health, and social outcomes. Unpublished manuscript, University of Chicago, Department of Economics.

Heckman, J. J. et al. (2010a, February). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics 94* (1–2), 114–128.

Heckman, J. J. et al. (2010b, August). Analyzing social experiments as implemented: A reexamination of the evidence from the High Scope Perry Preschool Program. *Quantitative Economics 1* (1), 1–46.

Heine, S. J., E. E. Buchtel, and A. Norenzayan (2008). What do cross-national comparisons of personality traits tell us? The case of Conscientiousness. *Psychological Science 19* (4), 309–313.

Heine, S. J. et al. (2002, June). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference-group effect. *Journal of Personality and Social Psychology 82* (6), 903–918.

Heine, S. J. et al. (2001). Divergent consequences of success and failure in Japan and North America: An investigation of self-improving motivations and malleable selves. *Journal of Personality and Social Psychology 81* (4), 599–615.

Herrnstein, R. J. and C. A. Murray (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.

Hieronymus, A. N., E. F. Lindquist, and H. D. Hoover (1980). *Iowa Tests of Basic Skills: Primary Battery*. Iowa City, IA: Houghton.

Hillage, J. et al. (2002). Employers skill survey: 2002. Research Report RR372, Department for Education and Skills.

Hirschi, T. and M. Gottfredson (1993). Commentary: Testing the general theory of crime. *Journal of Research in Crime and Delinquency 30* (1), 47–54.

Hogan, J. and B. Holland (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology 88* (1), 100–112.

Hollenbeck, K. (2008). State use of workforce system net impact estimates and rates of return. Technical report, Upjohn Institute. Presented at the Association for Public Policy Analysis and Management (APPAM) Conference, Los Angeles, CA.

Holmlund, H. and O. Silva (2009). Targeting non-cognitive skills to improve cognitive out- comes: Evidence from a remedial education intervention. Discussion Paper 4476, IZA.

Holt, M. M. and T. R. Hobbs (1979). The effects of token reinforcement, feedback and response cost on standardised test performance. *Behaviour Research and Therapy 17* (1), 81–83.

Holzer, H. (1997). Is there a gap between employer skill needs and the skills of the work force? In A. Lesgold, M. Feuer, and A. Black (Eds.), Transitions in *Work and Learning: Implications for Assessment*, Chapter 2, pp. 6–33. Washington, DC: National Academy Press.

Howard, K. S. and J. Brooks-Gunn (2009). The role of home-visiting programs in preventing child abuse and neglect. *The Future of Children 19* (2), 119–146.

Huang, D. et al. (2005). Keeping kids in school: An LA's BEST example—a study examining the long-term impact of LA's BEST on students' dropout rates. Technical report, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles, CA.

Huang, D. et al. (2000). A decade of results: The impact of the LA's BEST after school enrichment program on subsequent student achievement and performance. Technical report, UCLA Center for the Study of Evaluation, Graduate School of Education and Information Studies, Los Angeles, CA.

Humphries, J. E. (2014). Growth in GED testing. In J. J. Heckman, J. E. Humphries, and T. Kautz (Eds.), *The Myth of Achievement Tests: The GED and the Role of Character in American Life.* Chicago: University of Chicago Press.

Hunt, E. (2012). Improving intelligence. Presented at the American Psychological Association Annual Meeting, Sunday, May 27, 2012, Chicago, IL.

Jablonka, E. and G. Raz (2009). Transgenerational epigenetic inheritance: Prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly Review of Biology 84* (2), pp. 131–176.

Jackson, C. K. (2013). Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina. Working Paper 18624, NBER.

Jencks, C. (1979). *Who Gets Ahead? The Determinants of Economic Success in America*. New York: Basic Books.

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

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement. *Harvard educational review 39* (1), 1–123.

John, O. (2000). The big five personality test. http://www.outofservice.com/bigfive/?score-bigfive.

John, O. P. and S. Srivastava (1999). The Big Five trait taxonomy: History, measurement and theoretical perspectives. In L. A. Pervin and O. P. John (Eds.), *Handbook of Personality: Theory and Research*, Chapter 4, pp. 102–138. New York: The Guilford Press.

John, O. P. et al. (1994). The "little five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development 65* (1), 160–178.

Johnson, A. W. (1999). Assessing the impact of the Sponsor-A-Scholar youth mentoring program on student performance. Technical Report 355, Mathematica Policy Research.

Johnson, D. and T. Walker (1991). A follow-up evaluation of the Houston Parent-Child Development Center: School performance. *Journal of Early Intervention 15* (3), 226–236.

Jolly, R. (2007). Early childhood development: The global challenge. *The Lancet 369* (9555), 8–9.

Kaestle, C. F. (1984, May). Moral education and common schools in America: A historian's view. *Journal of Moral Education 13* (2), 101–111.

Kagitcibasi, C., D. Sunar, and S. Bekman (2001). Long-term effects of early intervention: Turkish low-income mothers and children. *Journal of Applied Developmental Psychology 22* (4), 333–361.

Kagitcibasi, C. et al. (2009). Continuing effects of early enrichment in adult life: The Turkish Early Enrichment Project 22 years later. *Journal of Applied Developmental Psychology 30* (6), 764–779.

Kahne, J. and K. Bailey (1999). The role of social capital in youth development: The case of "I Have a Dream" programs. *Educational Evaluation and Policy Analysis 21* (3), 321–343.

Kautz, T. and W. Zanoni (2014). Measuring and fostering non-cognitive skills in adolescents: Evidence from chicago public schools and the onegoal program. Unpublished manuscript, University of Chicago, Department of Economics.

Kemple, J. and J. C. Snipes (2000). Career Academies: Impacts on student engagement and performance in high school. Report 03/2000, MDRC.

Kemple, J. and C. Willner (2008). Career Academies: long-term impacts on labour market outcomes, educational attainment, and transitions to adulthood. Report 06/2008, MDRC.

Kern, M. L. and H. S. Friedman (2008). Do conscientious individuals live longer? *Health Psychology 27* (5), 505–512.

Kettner, N. (1976). Armed Services Vocational Aptitude Battery (ASVAB form 5): Comparison with GATB and DAT tests: Final report for period May 1975–October 1976. Technical report, DTIC Document: Department of Defense, Department of the Air Force, Air Force Systems Command, Air Force Human Resources Labouratory.

Kitzman, H. et al. (2010, May). Enduring effects of prenatal and infancy home visiting by nurses on children: Follow-up of a randomized trial among children at age 12 years. *Journal of the American Medical Association 164* (5), 412–418.

Knudsen, E. I. et al. (2006, July). Economic, neurobiological, and behavioural perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences 103* (27), 10155–10162.

Krueger, A. B. (2003, February). Economic considerations and class size. *Economic Journal 113* (485), F34–F63.

Kubzansky, L. D. et al. (2001). Is the glass half empty or half full? A prospective study of optimism and coronary heart disease in the normative aging study. *Psychosomatic Medicine 63* (6), 910–916.

Lake, A. (2011). Early childhood development—global action is overdue. *The Lancet 378* (9799), 1277–1278.

Lally, J. R., P. L. Mangione, and A. S. Honig (1987). The Syracuse University Family Development Research Program: Long-Range Impact of an Early Intervention with Low- Income Children & their Families. San Francisco, CA: Center for Child and Family Studies, Far West Labouratory for Educational Research and Development.

Lang, F. R. et al. (2011). Short assessment of the Big Five: Robust across survey methods except telephone interviewing. *Behaviour research methods 43* (2), 548–567.

Lapsley, D. and D. Yeager (2012). Moral-character education. In I. Weiner, W. Reynolds, and G. Miller (Eds.), *Handbook of Psychology*, Volume 7 of *Educational Psychology*, Chapter 7, pp. 117–146. New York, NY: Wiley.

Larson, G. E., D. P. Saccuzzo, and J. Brown (1994). Motivation: Cause or confound in information processing/intelligence correlations? *Acta Psychologica 85* (1), 25–37.

Lemann, N. (1999). *The Big Test: The Secret History of the American Meritocracy*. New York: Farrar, Straus and Giroux.

Lerman, R. I. (2013). Are employability skills learned in U.S. youth education and training programs? Under review, *IZA Journal of Labour Policy.*

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology 22* (140), 55.

Lillard, A. and N. Else-Quest (2006). The early years: Evaluating Montessori. *Science 313* (5795), 1893–1894.

Lleras, C. (2008). Do skills and behaviours in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. *Social Science Research 37* (3), 888–902.

Ludwig, J. and D. L. Miller (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity approach. *Quarterly Journal of Economics 122* (1), 159–208.

Lutz, M. N., J. Fantuzzo, and P. McDermott (2000). *Adjustment Scales for Preschool Intervention.* Philadelphia, PA: University of Pennysylvania.

Mallar, C. et al. (1982). Evaluation of the impact of the Job Corps program: Third follow-up report. Technical Report PR82-05, Mathematica Policy Research, Inc., Princeton, NJ.

Martin, L. R., H. S. Friedman, and J. E. Schwartz (2007). Personality and mortality risk across the life span: The importance of conscientiousness as a biopsychosocial attribute. *Health Psychology 26* (4), 428–436.

Martins, P. S. (2010). Can targeted, non-cognitive skills programs improve achievement? Discussion Paper 5266, IZA.

McAdams, D. P. (2006). *The Person: A New Introduction to Personality Psychology* (4 ed.). Hoboken, N.J.: J. Wiley and Sons.

McCord, J. (1978). A thirty-year follow-up of treatment effects. *American Psychologist 33* (3), 284–289.

McCord, J. and W. McCord (1959). A follow-up report on the Cambridge-Somerville Youth Study. *Annals of the American Academy of Political and Social Science 322* (1), 89–96.

McCormick, M. C. et al. (2006, March). Early intervention in low birth weight premature infants: Results at 18 years of age for the Infant Health and Development Program. *Pediatrics 117* (3), 771–780.

McDowell, J. et al. (2011), A shared responsibility: Apprenticeships for the 21st century, http://www.australianapprenticeships.gov.au/publications/shared-responsibility-apprenticeships-21st-century, last accessed February 11, 2013.

McLanahan, S. (2004, November). Diverging destinies: How children are faring under the second demographic transition. *Demography 41* (4), 607–627.

Michalopoulos, C. et al. (2002). Final Report on the Self-Sufficiency Project for Long-term Welfare Recipients. Report, Social Research and Demonstration Corporation.

Millenky, M., D. Bloom, and C. Dillon (2010). Making the transition: Interim results of the National Guard Youth ChalleNGe evaluation. Report 05/2010, MDRC.

Millenky, M. et al. (2011). Staying on course: Three- year results of the National Guard Youth ChalleNGe evaluation. Report 06/2011, MDRC.

Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.

Mischel, W. et al. (2011). 'Willpower' over the life span: decomposing self-regulation. *Social Cognitive and Affective Neuroscience 6* (2), 252–256.

Moon, S. H. (2014a). Decomposing racial skill gaps in the U.S. Unpublished manuscript, University of Chicago, Department of Economics.

Moon, S. H. (2014b). Multi-dimensional human skill formation with multi-dimensional parental investment. Unpublished manuscript, University of Chicago, Department of Economics.

Moon, S. H. (2012, July). Time to invest in disadvantaged young children. *Samsung Economic Research Institute Quarterly 5* (3), 50–59.

Mroczek, D. K. and A. Spiro (2007). Personality change influences mortality in older men. *Psychological Science 18 (5)*, 371–376.

Murray, H. A. (1938). *Explorations in Personality: A Clinical and Experimental Study of Fifty Men of College Age.* New York: Oxford University Press.

Myers, R. (1992). *The Twelve Who Survive: Strengthening Programmes of Early Childhood Development in the Third World.* New York, NY: Routledge/UNESCO.

National Apprenticeship Service (2012). History of apprenticeships. http://www.apprenticeships.org.uk, last accessed February 11, 2013.

National Centre for Vocational Education Research (2012). Historical time series of apprenticeships and traineeships in Australia from 1963. http://www.ncver.edu.au/publications/2329.html.

Naumann, L. P. and O. P. John (2011). Toward a domain-specific approach to cultural differences: The influence of cultural values and reference-group standards on self-reported personality. Unpublished manuscript, Sonoma State University, Department of Psychology.

Neal, D. A. (1997, January). The effects of Catholic secondary schooling on educational achievement. *Journal of Labour Economics 15* (1, Part 1), 98–123.

Nelsen, B. (1997). Should social skills be in the vocational curriculum? Evidence from the automotive career field. In A. M. Lesgold, M. J. Feuer, and A. M. Black (Eds.), *Transitions in Work and Learning: Implications for Assessment, Papers and Proceedings.* Washington, DC: National Academy Press.

Niles, M. D., A. J. Reynolds, and M. Nagasawa (2006). Does early childhood intervention affect the social and emotional development of participants? *Early Childhood Research and Practice 8* (1).

Nisbett, R. E. (2009, February). *Intelligence and How to Get It: Why Schools and Cultures Count.* New York, NY: W. W. Norton and Company.

Nisbett, R. E. et al. (2012). Intelligence: New findings and theoretical developments. *American Psychologist 67* (2), 130–159.

Nyhus, E. K. and E. Pons (2005). The effects of personality on earnings. *Journal of Economic Psychology 26* (3), 363–384.

OECD (2014), *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014)*, PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264201118-en.

OECD (2013), *PISA 2012 Results: Ready to Learn: Students' Engagement, Drive and Self-Beliefs (Volume III)*, PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264201170-en.

OECD (2011), *Doing Better for Families*, OECD Publishing, http://dx.doi.org/10.1787/9789264098732-en.

OECD (2001), OECD employment and labour market statistics, Data available from http://www.oecd-ilibrary.org/employment/data/oecd-employment-and-labour-market-statistics_lfs-data-en.

Olds, D. L. (2006). The Nurse-Family Partnership: An evidence-based preventive intervention. *Infant Mental Health Journal 27* (1), 5–25.

Olds, D. L. et al. (2010, May). Enduring effects of prenatal and infancy home visiting by nurses on maternal life course and government spending: Follow-up of a randomized trial among children at age 12 years. *Journal of the American Medical Association 164* (5), 419–424.

Olds, D. L. et al. (2007, December). Effects of nurse home-visiting on maternal and child functioning: Age 9 follow-up of a randomized trial. *Pediatrics 120* (4), 832–845.

Olds, D. L. et al. (2004, December). Effects of nurse home-visiting on maternal life course and child development: Age 6 follow-up results of a randomized trial. *Pediatrics 114* (6), 1550–1559.

Olds, D. L. et al. (2004, December). Effects of home visits by paraprofessionals and by nurses: Age 4 follow-up results of a randomized trial. *Pediatrics 114* (6), 1560–1568.

Oreopoulos, P., R. S. Brown, and A. Lavecchia (2014). Pathways to education: An integrated approach to helping at-risk high school students. Unpublished manuscript, University of Toronto, Department of Economics.

Orr, L. et al. (1994). *The National JTPA Study. Impacts, Benefits, and Costs of Title II-A.* Bethesda, MD: Abt Associates.

Peng, K., R. E. Nisbett, and N. Y. Wong (1997). Validity problems comparing values across cultures and possible solutions. *Psychological Methods 2* (4), 329–344.

Perez-Arce, F. et al. (2012). A cost-benefit analysis of the national guard youth ChalleNGe program. Technical Report 1193, RAND Corporation.

Pratt, T. C. and F. T. Cullen (2000). The empirical status of Gottfredson and Hirschi's general theory of crime: A meta-analysis. *Criminology 38* (3), 931–964.

Quinn, L. M. (2014). An institutional history of the GED. In J. J. Heckman, J. E. Humphries, and T. Kautz (Eds.), *The Myth of Achievement Tests: The GED and the Role of Character in American Life.* Chicago: University of Chicago Press.

Ramey, C. T. et al. (1976). The Carolina Abecedarian Project: A longitudinal and multidisciplinary approach to the prevention of developmental retardation. In T. Tjossem (Ed.), *Intervention*

Strategies for High-Risk Infants and Young Children, pp. 629–655. Baltimore, MD: University Park Press.

Rauner, F. (2007). Vocational education and training—A European perspective. In A. Brown, S. Kirpal, and F. Rauner (Eds.), *Identities at Work*, Volume 5 of *Technical and Vocational Education and Training: Issues, Concerns and Prospects*, pp. 115–144. Springer.

Raven, J., J. C. Raven, and J. Court (1988). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* San Antonio, TX: Harcourt Assessment.flynn

Raven, J. C. (1962). *Advanced Progressive Matrices: Sets I and II (*Revised ed.). London: H.K. Lewis.

Reed, D. et al. (2012). An effectiveness assessment and cost-benefit analysis of registered apprenticeship in 10 states. Mathematica Final Report 06689.090 and 40096, Mathematica Policy Research.

Reynolds, A. J. (2000). *Success in Early Intervention: The Chicago Child-Parent Centers*. Lincoln: University of Nebraska Press.

Reynolds, A. J. (1995). One year of preschool intervention or two: Does it matter? *Early Childhood Research Quarterly 10* (1), 1–31.

Reynolds, A. J. (1994). Effects of a preschool plus follow-on intervention for children at risk. *Developmental Psychology 30* (6), 787–804.

Reynolds, A. J. and J. A. Temple (1998, February). Extended early childhood intervention and school achievement: Age 13 findings from the Chicago Longitudinal Study. *Child Development 69* (1), 231–246.

Reynolds, A. J. et al. (2011, July). School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. *Science 333* (6040), 360–364.

Reynolds, A. J. et al. (2011, January– February). Age 26 cost-benefit analysis of the Child-Parent Center early education program. *Child Development 82* (1), 379–404.

Reynolds, A. J. et al. (2002). Age 21 cost-benefit analysis of the Title I Chicago Child-Parent Centers. *Educational Evaluation and Policy Analysis 24* (4), 267–303.

Roberts, B. W. (2009). Back to the future: Personality and assessment and personality development. *Journal of Research in Personality 43* (2), 137–145.

Roberts, B. W. and D. Mroczek (2008). Personality trait change in adulthood. Current Directions in Psychological Science 17 (1), 31–35.

Roberts, B. W., K. E. Walton, and W. Viechtbauer (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin 132* (1), 1–25.

Roberts, B. W. et al. (2007, December). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives in Psychological Science 2* (4), 313–345.

Roberts, R. D. et al. (2000). The Armed Services Vocational Aptitude Battery (ASVAB): Little more than acculturated learning (Gc)!? *Learning and Individual Differences 12* (1), 81–103.

Roder, A. and M. Elliott (2014). Sustained gains: Year Up's continued impact on young adults' earnings. Technical report, Economic Mobility Corporation.

Roder, A. and M. Elliot (2011). A promising start: Year-Up's initial impacts on young adults' careers. Technical report, Economic Mobility Corporation.

Rodríguez-Planas, N. (2012, October). Longer-term impacts of mentoring, educational services, and learning incentives: Evidence from a randomized trial in the United States. *American Economic Journal: Applied Economics 4* (4), 121–139.

Rodríguez-Planas, N. (2010). Mentoring, educational services, and economic incentives: Longer-term evidence on risky behaviours from a randomized trial. *Discussion Paper 4968*, IZA.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image.* Princeton, NJ: Princeton University Press.

Rothstein, R., R. Jacobsen, and T. Wilder (2008). *Grading Education: Getting Accountability Right.* New York: Economic Policy Institute and Teachers College Press.

Rutter, M. (2006). Implications of resilience concepts for scientific understanding. *Annals of the New York Academy of Sciences 1094* (1), 1–12.

Ryan, R. M. and E. L. Deci (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology 25* (1), 54–67.

Salgado, J. F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology 82* (1), 30–43.

Schmidt, F. L. and J. Hunter (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology 86* (1), 162–173.

Schmitt, D. P. (2002). Are sexual promiscuity and relationship infidelity linked to different personality traits across cultures? Findings from the International Sexuality Description Project. *Online Readings in Psychology and Culture 4* (Unit 4), 1–22. Retrieved from http://scholarworks.gvsu.edu/orpc/vol4/iss4/4.

Schmitt, D. P. et al. (2007). The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology 38* (2), 173–212.

Schochet, P. Z., J. Burghardt, and S. Glazerman (2001). *National Job Corps Study: The Impacts of Job Corps on Participants' Employment and Related Outcomes.* Princeton, NJ: Mathematica Policy Research, Inc.

Schochet, P. Z., J. Burghardt, and S. McConnell (2008, December). Does Job Corps work? Impact findings from the National Job Corps Study. *American Economic Review 98* (5), 1864–1886.

Schochet, P. Z., J. Burghardt, and S. McConnell (2006). *National Job Corps Study and Longer-Term Follow-Up Study: Impact and Benefit-Cost Findings Using Survey and Summary Earnings Records Data.* Princeton, NJ: Mathematica Policy Research, Inc.

Schulz, R. et al. (1996). Pessimism, age, and cancer mortality. *Psychology and Aging 11* (2), 304–309.

Schumacher, R. (2003). Family support and parent involvement in Head Start: What do Head Start program performance standards require? Technical report, Center for Law and Social Policy.

Schweinhart, L. J., H. V. Barnes, and D. Weikart (1993). *Significant Benefits: The High- Scope Perry Preschool Study Through Age 27.* Ypsilanti, MI: High/Scope Press.

Scribner, J. P. and D. Wakelyn (1998). Youth apprenticeship experiences in Wisconsin: A stakeholder-based evaluation. *The High School Journal 82* (1), 24–34.

Secretary's Commission on Achieving Necessary Skills (1992). *Learning a Living: A Blueprint for High Performance: A SCANS Report for America 2000.* Washington, DC: United States Government Printing Office.

Segal, C. (2012, August). Working when no one is watching: Motivation, test scores, and economic success. *Management Science 58* (8), 1438–1457.

Segal, C. (2008, October). Working when no one is watching: Motivation, test scores, and economic success. Economics working papers, Department of Economics and Business, Universitat Pompeu Fabra, Barcelona, Spain.

Shonkoff, J. P. and D. Phillips (2000). *From Neurons to Neighborhoods: The Science of Early Child Development.* Washington, DC: National Academy Press.

Shure, M. B. and G. Spivack (1988). Interpersonal cognitive problem solving. In R. H. Price, E. L. Cowen, R. P. Lorion, and J. Ramos-McKay (Eds.), *14 Ounces of Prevention: A Casebook for Practitioners,* Chapter 6, pp. 69–82. Washington, DC: American Psychological Association.

Social and Character Development Research Consortium (2010, October). Efficacy of school- wide programs to promote social and character development and reduce problem behaviour in elementary school children. Research Report NCER 20112001, National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Sroufe, L. A. (1997, Spring). Psychopathology as an outcome of development. *Development and Psychopathology 9* (2), 251–268.

Sroufe, L. A. et al. (2005). *The Development of the Person: The Minnesota Study of Risk and Adaptation from Birth to Adulthood.* New York, NY: Guilford Press.

Stasz, C. (2001). Assessing skills for work: Two perspectives. *Oxford Economic Papers 53* (3), 385–405.

Sylva, K. (1997). The quest for quality in curriculum. In L. J. Schweinhart and D. P. Weikart (Eds.), *Lasting Differences: The High/Scope Preschool Curriculum Comparison Study through Age 23*, pp. 89–93. Ypsilanti: High/Scope Press.

Temple, J. A. and A. J. Reynolds (2007). Benefits and costs of investments in preschool education: Evidence from the Child-Parent Centers and related programs. *Economics of Education Review 26* (1), 126–144.

Terman, L. M. and M. A. Merrill (1960). *Stanford-Binet Intelligence Scale: Manual for the Third Revision Form L-M.* Boston: Houghton Mifflin.

Terman, L. M. et al. (1947). *Genetic Studies of Genius: The Gifted Child Grows Up: Twenty-Five Years' Follow-Up of a Superior Group*, Volume 4. Stanford University, CA: Stanford University Press.

Terman, L. M. et al. (1925). *Genetic Studies of Genius: Mental and Physical Traits of a Thousand Gifted Children*, Volume 1. Stanford University, CA: Stanford University Press.

Thaler, R. H., S. Mullainathan, and D. Kahneman (2008, July). A short course in behavioural economics. Edge Master Class, Sonoma, CA, July 25–27, 2008.

The KIPP Foundation (2011). The promise of college completion: KIPP's early successes and challenges. Report, The KIPP Foundation, San Francisco, CA.

Tierney, J. P., J. Baldwin-Grossman, and N. L. Resch (1995). Making a difference: An impact study of Big Brothers Big Sisters. Report, Public/Private Ventures.

Tough, P. (2012). *How Children Succeed: Grit, Curiosity, and the Hidden Power of Character*. Boston, MA: Houghton Mifflin Harcourt.

Tuttle, C. C. et al. (2013). KIPP middle schools: Impacts on achievement and other outcomes: Final report. Report 06441.910, Mathematica Policy Institute.

Tuttle, C. C. et al. (2010). Student characteristics and achievement in 22 KIPP middle schools. Report 06441.900, Mathematica.

Tyler, R. W. (1973). Assessing educational achievement in the affective domain. *Measurement in Education 4* (3), 1–8.

Walker, G. C. and F. Vilella-Velez (1992). *Anatomy of a Demonstration: The Summer Training and Education Program (STEP) from Pilot through Replication and Postprogram Impacts*. Philadelphia, PA: Public/Private Ventures.

Walker, S. et al. (2011). Inequality in early childhood: risk and protective factors for early child development. *The Lancet 378* (9799), 1325–1338.

Walker, S. P. et al. (2011). Early childhood stimulation benefits adult competence and reduces violent behaviour. *Pediatrics 127* (5), 849–857.

Walker, S. P. et al. (2007). Early childhood stunting is associated with poor psychological functioning in late adolescence and effects are reduced by psychosocial stimulation. *The Journal of Nutrition 137* (11), 2464–2469.

Walker, S. P. et al. (2007, January). Child development: Risk factors for adverse outcomes in developing countries. *The Lancet 369* (9556), 145–157.

Walker, S. P. et al. (2005). Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: Prospective cohort study. *The Lancet 366* (9499), 1804–1807.

Walker, T. and D. L. Johnson (1988). A follow-up evaluation of the Houston Parent-Child Development Center: Intelligence test results. *Journal of Genetic Psychology 149* (3), 377–381.

Walsh, D. A. (2005). *Why Do They Act That Way? A Survival Guide to the Adolescent Brain for You and Your Teen* (1 ed.). New York: Free Press.

Washington Workforce Training Board (2008). Washington state employers workforce needs and practices survey. Statewide report.

Westat (2010). Head Start Impact Study: Final report. Contract 282-00-0022, U.S. Department of Health and Human Services.

Westwood, A. (2004). Skills that matter and shortages that don't. In C. Warchust, I. Grugulis, and E. Keep (Eds.), *The Skills that Matter*. New York, NY: Palgrave- Macmillan.

Winkelmann, R. (1996, July). Employment prospects and skill acquisition of apprenticeship- trained workers in Germany. *Industrial and Labour Relations Review 49* (4), 658–672.

Wolter, S. C. and P. Ryan (2011). Apprenticeship. In S. M. Eric A. Hanushek and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 3 of *Handbooks in Economics*, Chapter 11, pp. 521–576. Elsevier.

Young, M. E. (1996). *Early Childhood Development: Investing in the Future*. Washington, DC: World Bank.

Zabel, J., S. Schwartz, and S. Donald (2013, February). An analysis of the impact of the Self-Sufficiency Project on wages. *Empirical Economics 44* (1), 231–259.

Zabel, J., S. Schwartz, and S. Donald (2010). The impact of the Self-Sufficiency Project on the employment behaviour of former welfare recipients. *Canadian Journal of Economics 43* (3), 882–918.

Zemsky, R. (1997). Skills and the economy: An employer context for understanding the school-to-work transition. In A. Lesgold, M. Feuer, and A. Black (Eds.), *Transitions in Work and Learning: Implications for Assessment.* Washington, DC: National Academy Press.

Directorate for Education and Skills
Centre for Educational Research and Innovation (CERI)

Education and Social Progress
www.oecd.org/edu/ceri/educationandsocialprogress.htm