Web Appendix to Fostering and Measuring Skills: Interventions That Improve Character and Cognition

James J. Heckman and Tim Kautz*

This draft, November 9, 2013.

*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; University College Dublin, Ireland. Tel: 773-702-0634; Fax: 773-702-8490. Email: jjh@uchicago.edu. Kautz: Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago IL 60637. Tel: 773-702-0634. Email: tkautz@uchicago.edu. This research was supported in part by the University of Chicago, A New Science of Virtues: A Project of the University of Chicago, the American Bar Foundation, a conference series from the Spencer Foundation, the JB & MK Pritzker Family Foundation, Susan Thompson Buffett Foundation, the Geary Institute, University College Dublin, Ireland, NICHD R37 HD065072 and R01 HD054702, a European Research Council grant hosted by University College Dublin, DEVHEALTH 269874, a grant from the Institute for New Economic Thinking (INET), and an anonymous funder. The views expressed in this paper are those of the authors and not necessarily those of the funders or commentators mentioned here. We thank Linor Kiknadze and Edward Sung for valuable research assistance. We received helpful comments from Richard Boyle, Ron Diris, Maryclare Griffin, Robert Lerman, Seong Hyeok Moon, Dan Moran, Maria Rosales, and Indra Wechsberg.

Contents

Appendix for Summary of Effects Table

This document discusses how we categorized the interventions summarized in the "Summary of Effects Table." First, we explain the general methodology and approach. Second, we summarize the evidence.

Methodology

The summary table reviews a number of interventions that could potentially affect cognitive or non-cognitive skills. It is not a comprehensive review of the literature. We focus on interventions that have long-term follow-ups, offer important lessons, or are relatively well-cited. The table summarizes the components of the interventions and the effects on various outcomes.

It is difficult to compare the interventions for three primary reasons. First, evaluations have different lengths of follow-up. Some long-term interventions have found that early benefits fade. Comparing these evaluations with short-run evaluations can be misleading, because it is unclear whether the short-run increases would be sustained.

Second, different evaluations measure different outcomes within the categories of outcomes that we consider. For example, some evaluations only measure one or two non-cognitive skills, while others measure hundreds. It is possible that an intervention with fewer measures improved measures of other non-cognitive skills. In our summary, we do not address this issue.

Third, based on written reports, it is difficult to compare the components of the interventions. We attempt to classify the types of interventions into broad categories, e.g. whether the intervention attempted to affect health or whether the intervention grouped participants together. These comparisons might be misleading because interventions containing that we place in the same

category could still be very different from each other. Some interventions that have a health component provide nutritional supplements to young children, some for elementary school children include lessons about health behavior, and some of the interventions targeted towards adults include comprehensive medical coverage.

These issues make creating the summary table inherently subjective. In the remainder of this document, we provide our justification for how we have classified the components of the intervention and the effect on outcomes. Below we add additional notes about each column of the table.

Participant Characteristics

Age: Age indicates the earliest age at which participants entered the program. Some programs targeted students in particular grade levels, rather than ages. In these cases, we use age ranges that are typical of those grades, e.g. if the program targeted kindergarteners, we assume that children were age 5 to 6 at entry. Some programs allowed older students to enter later in the evaluation. For these evaluations, we report the youngest age but note these cases in this document.

Duration: Duration indicates the length of the intervention. Some evaluations included the baseline measurements or follow-ups when reporting the duration of the intervention. In these cases, we only use the length of the intervention itself. In cases where interventions are reported in months, we round to the nearest year.

Target: Target refers to the demographic targeted by the intervention. We categorize the target populations into one of 6 categories:

1. “SES” indicates that the program targeted generally disadvantaged populations. In some cases, “SES” indicates that the program had specific requirements for participation (e.g. incomes below a certain level) and in other cases it meant the programs targeted populations in particularly poor neighborhoods or regions.

2. “IQ” indicates that the program had specific IQ requirements as part of the screening.

3. “School” Indicates that the program selected participants based on performance in school, either based on teacher reports, test scores, or grades.

4. “Crime” indicates that the program targeted areas or populations with particularly high crime rates.

5. “Behav” indicates that the program targeted people with particular behavioral problems as reported by third parties such as teachers or community members. “Health” indicates that the program targeted populations with particular health problems, such as low birth weight or stunting.

6. “Drop” indicates that the program targeted people who dropped out of school.

Selection: Selection describes who initiated selection into the program. In some cases, programs or evaluators actively recruited participants and in other cases, participants applied to the program and were admitted. This type of selection might affect the interpretation of the results, e.g. it is possible that participants who actively applied to the program might have more

motivation to succeed or be more receptive to the program. We categorize selection into the program in five ways:

1. "Prgrm" indicates that the program actively recruited the participants.

2. "Refer" indicates that the participants were referred to the program by a third party, e.g. a hospital might refer patients with a certain health condition.

3. "Self" indicates that the individual actively applied to the program

4. "Schl" indicates that the school identified or recruited the participant.

5. "Par" indicates that parents applied to the program for their child.

Follow-Up: Follow-up refers to the duration of follow-up in the studies we review. Some programs have evaluations different follow-ups. In these cases, we use the longest available. We define follow-up relative to the time that the program started (not finished).

Sample: This indicates whether the sample size used for analysis in the study. Some of the programs have multiple evaluations with different sample sizes. In these cases, we report the one with the largest sample size. When possible, we use the sample size actually used for analysis. Some studies do not clearly indicate the analysis sample size.

<u>Intervention Components</u>

Home: Home refers to programs that include a substantial home-visit component. Several programs report that the staff occasionally visits homes, e.g. when a participant is having family problems. When these visits are rare and not part of the program design, we do not count the intervention as having a home visit component.

Health: Health indicates that the programs provides healthcare or teaches about health-related topics. In some cases, programs will devote only a small portion of the curriculum to health. In these cases, we do not count them as having a health component because health is not a primary focus of the program.

Parental: Parental indicates that the program actively engages parents of participants in a systematic way, either through home visits, parenting classes, or by involving parents in activities at a center. Some programs will contact parents if the participant is misbehaving or struggling, but when these instances are rare we do not count these programs as having a parental component.

On Site: On Site indicates that the program takes place at a facility such as a school, a job training center, or preschool center.

Group: Group indicates that the children or primary participants in the program are grouped together. We include this category because several studies have suggested that grouping participants together can induce negative peer effects.

Effects on Outcomes

We categorize the effect on outcomes in six different ways.

1) “Not measured” indicates that the outcome was not measured in any of the studies that we reviewed.

2) “Negative” means that the intervention had a negative effect on outcomes.

3) “Weak” indicates that the effect on a particular outcome is weak in one of three ways. First, weak could indicate any effects on the outcome might have faded out over time. Several interventions boosted outcomes directly after the end of the program but the effects were temporary. Second, weak could indicate that the effect of the program was economically small even if statistically significant. Third, weak could indicate that the intervention improved some outcomes within the category without changing other, e.g. we might classify an effect on health as weak if the intervention improved drinking habits without affecting other drug habits, weight, mental health, or self-reported health.

4) “Positive” indicates that the effect in a particular category outcomes was improved in a consistent and statistically significant way. If the intervention had a long term follow-up, then a positive rating means that the effect lasted throughout the intervention. It also indicates that the effect is likely economically significant.

5) “None” indicates that the intervention had no statistically significant effect on an outcome.

6) "Mixed" indicates that the intervention affected a subgroup but not the overall sample. For example, many interventions primarily improved outcomes for girls but not boys (although, there are a few exceptions).

We categorize outcomes as follows:

IQ: IQ includes any measures of a true IQ test (as opposed to a standardized achievement test).

School: School includes measures of school performance including GPA, achievement test scores, and retention. Some of these measures could conceivably belong in the educational attainment category, e.g. retention relates both to school performance and educational attainment.

Non-Cognitive: Includes any measurement of non-cognitive skills. The measures include self-reports, teacher reports, or in some cases behaviors that relate directly to non-cognitive skills but do not fit into any of the other categories.

Education: Education includes any measures of total educational attainment, such as high school dropout rates, college credits earned, and degrees earned. We include grade retention in the school category.

Health: Health includes a range of health outcomes, including self-reported health and direct measures of health such as BMI. We include measures of mental health in this category, even though some measures of mental health are related to non-cognitive skills. We also include measures of risky behaviors such as drinking and smoking in the health category. At some ages,

some of the measures such as drinking could conceivably be classified as crime, but we have chosen to include them in the health category.

Crime: Crime includes any measures of criminal activity besides drug and substance abuse.

Earnings: Earnings includes reported earnings, hourly wages, and employment. We do not include other labor market outcomes in the table but summarize them in the rest of this document when applicable.

Return/Benefits

Rate of Return: Rate of return is the estimated annual rate of return to the program.

Benefit-Cost Ratio: The benefit cost ratio is the estimated benefits of the program divided by the total costs of the program.

Summary of Evidence

In the remained of the document, we summarize the evidence and explain how we reached each of the conclusions in the summary of effects table. For each intervention, the categories correspond to the columns in the intervention table. The parentheses indicate the value that we entered into the table, and the remaining text explains our rational.

Nurse Family Partnership (NFP)

Participant Characteristics

Age: (0) The program began pre-birth (Olds [2006]).

Duration: (2Y) The program lasted through the first 2 years of the child's life (Olds [2006]).

Target: (SES) Each program has focused on women who have not previously given birth, are young, and have relative low incomes. Memphis and Denver trials targeted especially high-risk mothers, after it was found that in the Elmira sample high-risk mothers benefited the most (Olds [2006]).

Selection: (Program) The program actively recruits eligible participants (Olds et al. [2004a]).

Follow-Up (19Y): Eckenrode et al. [2010] follow participants through age 19.

Sample Size: (641)

- Eckenrode et al. [2010] consider 310 youths.
- Kitzman et al. [2010] consider a sample of 613 children.
- Olds et al. [2010] considers a sample of 594 people.
- Olds et al. [2004a] considers a final sample of 635 women and 605 children.

- Olds et al. [2007] considers a final sample of 627 women. Different numbers of children responded to different numbers of questions. 558 have complete information on teacher reports. 604 have complete information on school records. 570 have complete records of achievement tests.

- Olds et al. [2004b] considers a final sample of 641 women and 615 children.

Intervention Components

Home: (Yes) The program primarily consisted of home visitations (Olds [2006]).

Health: (Yes) The program attempted to change the health behaviors of the mothers (Olds [2006]).

Parental: (Yes) The program targeted mothers (Olds [2006]).

On Site: (No) The program consisted of home visits and children did not interact in a group (Olds et al. [2004a]).

Group: (No) The program consisted of home visits, so it did not group children together (Olds [2006]).

Effects on Outcomes

IQ: (Positive) IQ is measured through age 6. In the Memphis sample, the program boosted mental processing and PPVT. In the Elmira sample at age 4, the effect on IQ was similar in

magnitude to the effect in the Memphis sample but the effect was not statistically significant (Olds et al. [2004a]).

School: (Mixed) The program primarily improved school outcomes for a sample of participants with low resources. At age 6, there are positive but generally insignificant effects on achievement tests. The only significant ($p<0.10$) effect was on arithmetic achievement but only for the low resource sample (Olds et al. [2004a]). At age 9, there the program boosted achievement test scores and grades, but the effect was only statistically significant for the low resource group (Olds et al. [2007]). At age 12, achievement tests and grades increased, but the effect was only statistically significant for the subgroup with low resources. Grade retention is higher in the treatment group but the difference is not statistically significant (Kitzman et al. [2010]).

Non-Cognitive: (Positive) At age 4 in the Denver sample, there was no effect on externalizing behavior but there were improvements in behavioral adaptation in testing for the low resource group (Olds et al. [2004b]). At age 6, the program had no statistically significant effect on classroom social skills or dysregulated aggression. At age 9, the treatment group children had a lower incidence of conduct failure (grades 1-3 total) but are similar in other measures. At age 12, the program improved measures of internalizing behavior but had no effect on externalizing behavior or an index of total problem behavior (Kitzman et al. [2010]).

Education: (None) At age 19, high school completion is slightly lower in the treatment group, but the effect is not statistically significant. The follow-up is too short to know whether the program affected college outcomes (Kitzman et al. [2010]).

Health: (Weak) The program reduced child abuse and injury (Olds [2006]). For poor and unmarried mothers, the program reduced role impairment (performance at work or with family and friends) due to drug and alcohol use, even though it did not affect the amount of alcohol or drug use (Olds et al. [2010]). At age 12, the program reduced substance use of the children (Kitzman et al. [2010]).

Crime: (Mixed) Mother's crime rates were reduced in the Elmira site but not in the Memphis sample (Olds et al. [2010]). At age 19, girls of mother in the treatment group were much less likely to have been arrested (Eckenrode et al. [2010]).

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed, but the program did reduce welfare use of the mothers (Olds et al. [2010]).

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies we reviewed.

Benefit-Cost Ratio: (2.88) Karoly et al. [2005] estimate that the benefit cost ratio is 1.26 for the lower risk sample, 2.88 for the entire sample, and 5.70 for the higher risk sample.

Abecedarian (ABC) Project (Note: The program had a preschool and school age component. In this report, we focus on the preschool intervention. In general, the school-age intervention had few positive returns. The same group that evaluated ABC simultaneously evaluated another program called CARE. We do not discuss the CARE program in this summary, as it is a conceptually different intervention.)

Participant Characteristics

Age: (0) The preschool program started at birth and lasted until age 5 (Heckman et al. [2012b]).

Duration: (5Y) The preschool program lasted for 5 years (Heckman et al. [2012b]).

Target: (SES) The program screened applicants using the High-Risk Index (HRI), which is based on 13 socio-economic factors. Selection was based in part on an interview of the mother and the mother's score on an intelligence test (Heckman et al. [2012b]).

Selection: (Referral) The program received referrals from local community organizations such as prenatal clinics, hospitals, and social services (Breitmayer and Ramey [1986]).

Follow-Up (30Y): Heckman et al. [2012b] follow participants through age 30.

Sample Size: (92)

- Heckman et al. [2012b] reports that 90 of the children had non-missing information at the 96^{th} month evlaution and first-folowup and 92 at the final follow-up.

Intervention Components

Home: (No) The preschool component did not include home visits. Although the CARE intervention included home visits (Heckman et al. [2012b]).

Health: (Yes) The children in the treatment group received their primary medical care in the child care center and received much of their nutrition at the child care center. However, children in the control group also received some iron-fortified formula (Heckman et al. [2012b]).

Parental: (Yes) The program involved parents, especially for the health component (Ramey et al. [1976]).

On Site: (Yes) The program took place at a day care center (Heckman et al. [2012b]).

Group: (Yes) Children were grouped together at the day care center (Heckman et al. [2012b]).

Effects on Outcomes

IQ: (Positive) The program had long-term effects on IQ, especially for girls (Heckman et al. [2012b]). The program improved IQ for boys but the effects are only borderline statistically significant after age 8 (Heckman et al. [2012b]). After accounting for multiple hypothesis testing, the effects on IQ are not longer statistically significant after age 8.

School: (Positive) At age 8, the program improved scores on achievement tests for both boys and girls (Heckman et al. [2012b]).

Non-Cognitive: (Mixed) At ages 6-8, treated children tend to be more creative, have better verbal intelligence, but also be more hostile (Heckman et al. [2012b]). At age 12 and 15, treated girls have lower internalizing and externalizing behavior (Heckman et al. [2012b]). At age 12 and 15, boys did not have improved internalizing and externalizing behavior (Heckman et al. [2012b]).

Education: (Mixed) For females the program increased rates of college attendance and of high school graduation (Heckman et al. [2012b]). For males the program did not have a statistically

significant effect on educational outcomes after accounting for multiple hypothesis testing (Heckman et al. [2012b]).

Health: (Weak) After correcting for multiple hypothesis testing, the program had little effect on health for girls. The program increased boy's BMIs before age 8. Treated boys were also more likely to be in a normal range of BMI. In adult years males were moderately less likely to use drugs but were more likely to drink. Females were less likely to have a sexually transmitted disease (Heckman et al. [2012b]).

Crime: (Mixed) After accounting for multiple hypothesis testing, there are no statistically significant impacts on criminal outcomes for girls. Boys have a lower number of probations ($p<0.10$) (Heckman et al. [2012b]).

Earnings: (Mixed) The program had little effect on earnings or employment for females (Heckman et al. [2012b]). At age 30, the program improved employment for males (Heckman et al. [2012b]).

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies we reviewed.

Benefit-Cost Ratio: (3.78) Using a discount rate of 3% and a real wage growth rate of 2%, Temple and Reynolds [2007] estimate a cost-benefit ratio of 3.78.

Infant Health and Development Program (IHDP)

Participant Characteristics

Age: (0) The program started at birth (McCormick et al. [2006]).

Duration: (3Y) The program lasted through the first 36 months of life (McCormick et al. [2006]).

Target: (Low Birth) The program targeted low-birth weight, preterm infants (McCormick et al. [2006]).

Selection: (Program) Children meeting the health eligibility requirements were asked to participate in the study (McCormick et al. [2006]).

Follow-Up (18Y): McCormick et al. [2006] study participants through age 18.

Sample Size: (636) McCormick et al. [2006] analyze a sample 636 youths.

Intervention Components

Home: (Yes) The program consisted of home visits during the first two years of the program (McCormick et al. [2006]).

Health: (No) The program included high-quality pediatric care and referral, but both the treatment and control groups received the services so we do not count it as part of the intervention (McCormick et al. [2006]).

Parental: (Yes) The program included group meetings for parents (McCormick et al. [2006]).

On Site: (Yes) The program included an onsite daycare component (McCormick et al. [2006]).

Group: (Yes) Children were grouped together in the daycare component (McCormick et al. [2006]).

Effects on Outcomes

IQ: (Mixed) At age 18, the intervention group had a higher full IQ score (WASI), but only for the higher birth weight group and before accounting for attrition. At age 18, the intervention group had a higher verbal IQ (PPVT), but this was only true for the higher birth weight children (McCormick et al. [2006]).

School: (Mixed) At age 18, the higher weight intervention group had higher scores on the Woodcock Johnson test for reading and math, but this was only true for reading for the lower birth weight sample (McCormick et al. [2006]).

Non-Cognitive: (Mixed) At age 18, the intervention group had better scores on the Youth Risk Behavior Surveillance System, but this was only true of the higher birth weight sample (McCormick et al. [2006]).

Education: (None) At age 18, there were no statistically significant effects on high school graduation (McCormick et al. [2006]).

Health: (None) At age 18, the program did not impact physical health scales reported by either the caretaker or the student (McCormick et al. [2006]).

Crime: (None) At age 18, there were no statistically significant effects on arrest rates or rates of having gone to jail (McCormick et al. [2006]).

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed.

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies we reviewed.

Benefit-Cost Ratio: (Not reported) The benefit-cost ratio was not estimated in the studies we reviewed.

Syracuse Family Development Research Program (FDRP)

Participant Characteristics

Age: (0) Mothers were recruited during their last trimester of pregnancy (Lally et al. [1987]).

Duration: (5Y) The families received 5 years of continuous day care (Lally et al. [1987]).

Target: (Low SES) All families had less than $5,000 income, and mothers had less than a high school education and had no work or semi-skilled work history (Lally et al. [1987]).

Selection: (Program) The program actively recruited mothers and then the mothers consented (Lally et al. [1987]).

Follow-Up (15Y): The program followed participants for 10 years after the 5-year intervention (Lally et al. [1987]).

Sample Size: (109) By the 15-year follow-up, the program had 109 members in the treatment and control group (Lally et al. [1987]).

Intervention Components

Home: (Yes) Weekly contact in the home was an emphasis of the intervention (Lally et al. [1987]).

Health: (Yes) The program included a nutrition and health and safety component (Lally et al. [1987]).

Parental: (Yes) The program viewed parental contact as the primary intervention, whereas childcare was less important (Lally et al. [1987]).

On Site: (Yes) Day care services were provided at Syracuse University's Children's Center (Lally et al. [1987]).

Group: (Yes) Children were grouped together at the day care center (Lally et al. [1987]).

Effects on Outcomes

IQ: (None) At 48 months, there was no effect on Stanford-Binet IQ test scores (Lally et al. [1987]).

School: (Mixed) The program improved grades for girls, but not boys. In 7^{th} and 8^{th} grade, none of the program girls were failing classes, whereas about 16% of the control group was failing a class. Girls also had better school attendance beginning in year 3 (Lally et al. [1987]).

Non-Cognitive: (Weak) Based on teacher reports, program girls were rated as having more positive attitudes toward themselves compared to the control girls. Program girls were also rated as having more impulse control and better academic achievement than treatment girls. There were no differences for boys (Lally et al. [1987]).

Education: (Not measured) Education was not measured in the studies that we reviewed.

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (Positive) The program reduced juvenile delinquency in the program group. 22% of the control group and 6% of the treatment group were processed as probation cases. The authors do not report differences by gender (Lally et al. [1987]).

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed.

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) The benefit-cost ratio was not estimated in the studies that we reviewed.

Houston Parent-Child Development Center

Participant Characteristics

Age: (1) The program began at age 1 (Johnson and Walker [1991]).

Duration: (2Y) The program lasted for two years (Johnson and Walker [1991]).

Target: (SES) The program targets low-income Mexican-American families (Johnson and Walker [1991]).

Selection: (Program) Program staff conducted a door-to-door survey to identify eligible families. Families then were contacted and asked invited them to join the program (Johnson and Walker [1991]).

Follow-Up (15Y): The longest follow-up was 13 years after the end of the program, but the length differed for different participants (Besharov et al. [2011]).

Sample Size: (170)

- Walker and Johnson [1988] used a sample size of 170.
- Johnson and Walker [1991] used a sample size of 137.

Intervention Components

Home: (Yes) The first year of the program included 25 home visits from a paraprofessional (Johnson and Walker [1991]).

Health: (No) The program did not focus on health. The program did have a section on "health and safety at home," but compared to most other interventions, this was a minor part of the program (Johnson and Walker [1991]).

Parental: (Yes) The program actively engaged mothers (Johnson and Walker [1991]).

On Site: (Yes) During the second year of the program, women and children attended a project center 4 mornings each week (Johnson and Walker [1991]).

Group: (Yes) During the second year of the program, the mothers and children participated in group discussions (Johnson and Walker [1991]).

Effects on Outcomes

IQ: (Weak) There was an initial increase in Stanford-Binet IQ scores at the end of the program when children were 3 (Bridgeman et al. [1981]). When children were 4 to 6 and 6 to 9, there were no statistically significant differences in overall IQ (Walker and Johnson [1988]). However, there were some program effects in some of the subscales (Walker and Johnson [1988]). We classify the effect as weak, because it fades over time and the only relatively long-term effects apply to some of the subscales.

School: (Weak) At ages 8-11, there were no statistically significant differences in grades between the treatment and control groups, but the treatment group performed better on the Iowa Test of Basic Skills (Johnson and Walker [1991]). At ages 8-11, the children in the treatment group were less likely to be retained, but the difference was not statistically significant (Johnson and Walker [1991]).

Non-Cognitive: (Positive) At ages 8-11, the program reduced hostility for both girls and boys. The program reduced dependency for boys (Johnson and Walker [1991]).

Education: (Not measured) Education was not measured in the studies that we reviewed.

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (Not measured) Crime was not measured in the studies that we reviewed.

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed.

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Jamaican Supplementation Study (JSS)

Participant Characteristics

Age: (1-2) The children were aged 9-24 months at entry (Grantham-McGregor et al. [1991]).

Duration: (2Y) The program lasted 2 years (Grantham-McGregor et al. [1991]).

Target: (Health) The program targeted stunted children (low height for age) (Grantham-McGregor et al. [1991]).

Selection: (Program) The program recruited participants through a house-to-house survey (Grantham-McGregor et al. [1991]).

Follow-Up (22Y): The longest follow-up is 20 years after the end of the program (Gertler et al. [2012]).

Sample Size: (159)

- Grantham-McGregor et al. [1991] use a sample size of 159.
- Walker et al. [2005] use a sample size of 103.
- Walker et al. [2007] use a sample size of 103.
- Gertler et al. [2012] uses a sample of between 94 and 115 for most of their analyses.

Intervention Components

Home: (Yes) The children were visited once per week (Grantham-McGregor et al. [1991]).

Health: (Yes) The intervention included nutritional supplementation (Grantham-McGregor et al. [1991]).

Parental: (Yes) Part of the intervention included a stimulation component in which mothers were taught how to play with their children (Grantham-McGregor et al. [1991]).

On Site: (No) The program consisted of home visits and did not have onsite meetings (Grantham-McGregor et al. [1991]).

Group: (No) The program consisted of individualized home-visits, so children were not grouped together (Grantham-McGregor et al. [1991]).

Effects on Outcomes

IQ: (Positive) Both supplementation and stimulation lead to higher IQ's at age 2 (Grantham-McGregor et al. [1991]). For stimulation this effect persists through age 18 and comprises more than half of a standard deviation (Walker et al. [2005]).

School: (Positive) There is a positive impact of stimulation on reading achievement. Math achievement test scores were not affected (Walker et al. [2005]).

Non-Cognitive: (Positive) The program led to better internalizing behavior (lower anxiety and depression, higher self-esteem). The effect on externalizing behavior was insignificant (Gertler et al. [2012]).

Education: (Weak) The program had modest effects on education. Dropout rates are slightly lower in the treatment group. The impact on years of education is positive but insignificant at age 20-22, but significantly more treatment group individuals are attending additional education (Gertler et al. [2012]).

Health: (Positive) The program improved measures of mental health, including anxiety and depression. Hyperactivity is not affected and still higher than in the non-stunted group. All other

measures of mental health are not significantly different from those of the non-stunted children (Walker et al. [2007]).

Crime: (Not measured) Crime was not measured in the studies that we reviewed.

Earnings: (Positive) 20 years after the intervention, the treatment group had much higher earnings compared to the control group, especially in recent jobs and permanent jobs. Employment is higher in the treatment group, but the difference is not statistically signficant (Gertler et al. [2012]).

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) The benefit-cost ratio was not estimated in the studies that we reviewed.

Perry Preschool Program

Participant Characteristics

Age: (3) The program targeted 3 year olds, although some children began the program at age 4 (Heckman et al. [2010a]).

Duration: (2Y) The program lasted for 2 years (Heckman et al. [2010a]).

Target: (SES, Low-IQ) The program targeted disadvantaged African American children with low IQs (Heckman et al. [2010a]).

Selection: (Program) Candidate families were identified and then were asked to participate (Heckman et al. [2010a]).

Follow-Up (37Y): The longest follow-up was at age 40, 37 years after the start of the program (Heckman et al. [2010a]).

Sample Size: (123) Most of the analyses in Heckman et al. [2010a] had 108-123 observations.

Intervention Components

Home: (Yes) The program included weekly home visits by the teacher (Heckman et al. [2010a]).

Health: (No) The program did not focus on improving health (Heckman et al. [2010a]).

Parental: (Yes) The home visits included parental interaction (Schweinhart et al. [1993]).

On Site: (Yes) The program took place at a preschool facility (Heckman et al. [2010a]).

Group: (Yes) The program grouped together children at the facility (Heckman et al. [2010a]).

Effects on Outcomes

IQ: (Weak) The program improved IQ for both girls and boys in the short run but the effects faded by age 10 (Heckman et al. [2012a]).

School: (Positive) The program improved GPAs for girls but not boys (Heckman et al. [2010a]). The program improved scores on the California Achievement Test (CAT) for both boys and girls (Heckman et al. [2012a]).

Non-Cognitive: (Positive) The program improved externalizing behavior for boys and girls and academic motivation for girls (Heckman et al. [2012a]).

Education: (Mixed) The program improved high school graduation rates for girls but not boys (Heckman et al. [2010a]).

Health: (None) The program reduced alcohol use for boys but the effect was not statistically significant once adjusting for multiple hypothesis testing (Heckman et al. [2010a]).

Crime: (Positive) The program reduced arrest rates for both girls and boys (Heckman et al. [2010a]).

Earnings: (Mixed) The program improved earnings for boys (Heckman et al. [2010a]).

Return/Benefits

Rate of Return: (8.1-9.2) Using a 3% discount rate, Heckman et al. [2010b] estimate the rate of return to be between 8.1 and 9.2 (depending on whether murder costs are high or low).

Benefit-Cost Ratio: (7.1-12.2) Using a 3% discount rate, Heckman et al. [2010b] estimate the rate of return to be between 7.1 and 12.2 (depending on whether murder costs are high or low).

Head Start

Participant Characteristics

Age: (3) Students may enter the program after age 3, but the earliest age is 3 (Currie and Thomas [1995]). Early Head Start starts at birth (Carneiro and Ginja [2012]).

Duration: (2Y) The program lasts for 2 years (Currie and Thomas [1995]).

Target: (Low SES) Federal funding requires that 90% of children in Head Start be below the federal poverty level (Deming [2009]).

Selection: (Parent) Head Start is a national program and parents enroll their children (Currie and Thomas [1995]).

Follow-Up (23Y): Garces et al. [2002] follows participants through age 25.

Sample Size: (4,165)

- Garces et al. [2002] use a sample of a little less than 3,255 people.
- Currie and Thomas [1995] use 4,165 observations in their main analyses but use fewer for some of their sub-analyses.
- Deming [2009] uses a final sample size of 3,698 for the main analysis.
- Carneiro and Ginja [2012] use a final sample size of 1,676 children, but some of the analyses use a sub-sample.
- Westat [2010] uses different sample sizes depending on the analysis. They range from 724 to 2110.
- Ludwig and Miller [2007] use a sample size of 1,296.

Intervention Components

Home: (Yes) Head start programs can offer both center-based and home based care (Carneiro and Ginja [2012]).

Health: (Yes) Head Start provides medical care, dental care, and mental health care (Deming [2009]).

Parental: (Yes) The home-based option includes the parents (Carneiro and Ginja [2012]).

On Site: (Yes) Head start programs can offer both center-based and home based care (Carneiro and Ginja [2012]).

Group: (Yes) The center-based option would group together children care (Carneiro and Ginja [2012]).

Effects on Outcomes

IQ: (Weak) Effects of Head Start on IQ typically fade away by grade 3 (See the discussion in Currie and Thomas [1995]).

School: (Weak) There are initial gains in test scores, but they fade away by age 11-15. Persistent gains are found for the subgroup of males and those with mothers who have low AFQT scores. The program also lowers grade retention for whites but not blacks (Currie and Thomas [1995]).

Non-Cognitive: (None) The Head Start Impact Study finds initial positive effects on behavior that fade away by the end of pre-school (Westat [2010]). Other studies do not report the program's impact on non-cognitive skills.

Education: (Positive) Several studies found that the program improves educational attainment, but they find that different sub-groups benefit. Deming [2009] report that boys and blacks are the largest benefactors (Deming [2009]). Garces et al. [2002] finds that the effect is only significant for whites (Garces et al. [2002]).

Health: (Positive) Currie and Thomas [1995] find that participants have better access to preventative health care, but the effect is equal to those attending other preschools. Ludwig and Miller [2007] find that the program lowers mortality rates. Deming [2009] reports improvements in self-reported health.

Crime: (Mixed) Deming [2009] finds no effects on criminal outcomes. Garces et al. [2002] find that the program benefits blacks.

Earnings: (Positive) Garces et al. [2002] report find that whites have higher earnings in their 20s.

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) The benefit-cost ratio was not estimated in the studies that we reviewed.

Chicago Child-Parent Center Program (Note: The intervention contained a preschool component, a kindergarten component, and a grade school component. We focus our discussion on the preschool component.)

Participant Characteristics

Age: (3&4) The preschool program admits 3 and 4 year old students.

Duration: (2Y) The preschool program lasted for 2 years (Reynolds [1994]).

Target: (SES) The program is available to people who live in particular residential areas in Chicago and applicants are selected based on a most-in-need basis. Children are eligible due to economic and educational disadvantage (Reynolds and Temple [1998]).

Selection: (Parent) Parents living in eligible regions of Chicago apply to participate in the program (Reynolds [1994]).

Follow-Up (25Y): Reynolds et al. [2011a] includes a follow-up though age 28.

Sample Size: (1,286)

- Niles et al. [2006] used a study sample of 1,379 people.
- Reynolds [1994] used a study sample of 1,106 children.
- Reynolds and Temple [1998] used a study sample of 559 children.
- Reynolds et al. [2011a] used study samples ranging from 971 to 1,473.
- Reynolds et al. [2011b] used study samples ranging from 1,233 to 1,473.
- Reynolds et al. [2002] used a study sample of 1,286

Reynolds et al. [2011a] includes a sample size of 1,539.

Intervention Components

Home: (No) While implementation varies across sites, the program does not include home visits (Reynolds [1994]).

Health: (Yes) The program has a health component, including free breakfasts and lunches and health screenings (Reynolds [1994]).

Parental: (Yes) Parental involvement is an important part of the program. The program requires one half day per week of parental involvement (Reynolds [1994]).

On Site: (Yes) The program takes place at a center (Reynolds [1994]).

Group: (Yes) The program groups children together at the centers (Reynolds [1994]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Positive) Compared to the control group, participants in the preschool program scored better on components of the Iowa Test of Basic Skills (ITBS) at ages 5, 6, 9, and 14 (Reynolds et al. [2002]). Participants in the preschool program were less likely to be retained by age 15 and were less likely to be in special education by age 18 (Reynolds et al. [2002]).

Non-Cognitive: (Positive) Compared to children who were not in the CPC program, participants in the preschool and kindergarten group had improved social-emotional competence (Niles et al. [2006]).

Education: (Positive) At age 20 and 21, participants in the preschool group had a 11.2 and 10.5 percentage point higher high school completion (Reynolds et al. [2002]). At age 25, participants in the preschool group had 6.8 percentage point higher high school graduation rate and were more likely to have completed at least 0.5 credits at a 4-year college (Reynolds et al. [2011a]). The program did not have a statistically significant effect on receipt of a AA or BA degree by age 28 (Reynolds et al. [2011a]).

Health: (Positive) The program reduces child abuse. Participants reported higher use of health insurance, reduced substance abuse, and reduced depression (Reynolds et al. [2011b]). The program also reduced substance abuse at age 28 (Reynolds et al. [2011a]).

Crime: (Positive) Participants in the preschool program had fewer juvenile arrests and adult arrests by age-26 (Reynolds et al. [2011b]). At age 28, the program improved arrests (Reynolds et al. [2011a]).

Earnings: (Positive) The program improved age-28 income (Reynolds et al. [2011a]).

Return/Benefits

Rate of Return: (18) Reynolds et al. [2011b] estimate a rate of return of 18% using a discount rate of 3% and a real wage growth rate of 2%.

Benefit-Cost Ratio: (10.8) Reynolds et al. [2011b] estimate a rate of return of 10.8 using a discount rate of 3% and a real wage growth rate of 2%.

Turkish Early Enrichment Project (TEEP)

Participant Characteristics

Age: (3,5) Children were 3 and 5 at entry (Kagitcibasi et al. [2001]).

Duration: (2Y) The study lasted for a total of 4 years, but the first and last years were used to assess the child, so we do not include them in the duration of the intervention (Kagitcibasi et al. [2001]).

Target: (SES) The study selected children from centers serving low income families (Kagitcibasi et al. [2001]).

Selection: (Program) The program recruited mothers and then requested them to attend the mother training (Kagitcibasi et al. [2001]).

Follow-Up: The longest follow-up is 22-years (Kagitcibasi et al. [2009]).

Sample Size: (255)

- Kagitcibasi et al. [2001] has a sample size of 255.
- Kagitcibasi et al. [2009] has a sample size of 131.

Intervention Components (Note: The sample consisted of several treatment groups. First, children were chosen from three separate types of daycare programs: custodial daycare, educational day care, and home day care. This selection was not based on random assignment. Next families in each of the groups were randomly assigned to mother training or a control group. For the purposes of this report, we consider the intervention to be the mother training, not the different types of daycare that the children were already enrolled in at the start of the study.)

Home: (Yes) The mother training included short home visits in which educational materials were given to the mothers (Kagitcibasi et al. [2001]).

Health (Yes) The mother training contained discussions of nutrition (Kagitcibasi et al. [2001]).

Parental: (Yes) Mother training was provided to a randomly selected sub-sample of participants (Kagitcibasi et al. [2001]).

On Site: (Yes) Part of the mother training took place on site (Kagitcibasi et al. [2001]).

Group: (No) The mother training took place in groups, but the children themselves were not placed in groups during this time (Kagitcibasi et al. [2001]).

Effects on Outcomes

IQ: (Weak) In the fourth year (the year after the end of the program, IQ was higher in the mother training group Seven years after the program, IQ was still higher in the mother training group (Kagitcibasi et al. [2001]). There were not significant differences in vocabulary tests at age 19 (Kagitcibasi et al. [2009]).

School: (Positive) In the fourth year, mother training had an effect on a general ability achievement test but not on Turkish or mathematics scores. In the fourth year for students entering at age 5, mother training had an effect on Turkish grades. Seven years after the program, students in the mother training group had higher grades in math, Turkish, and overall (Kagitcibasi et al. [2001]).

Non-Cognitive: (Positive) In the fourth year, mother training children were rated as less aggressive, higher in self-concept, and better in deportment scores (Kagitcibasi et al. [2001]).

Education: (Positive) Seven years after the program, children in the mother training group were more likely to still be in school (Kagitcibasi et al. [2001]). At age 19, children in the mother training sample were more likely to have attended college (Kagitcibasi et al. [2009]).

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (Not measured) Crime was not measured in the studies that we reviewed.

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed.

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Project STAR

Participant Characteristics

Age: (5-6) Participants could enter the program in kindergarten (Chetty et al. [2010]). We assume that they were 5-6 years old. However, some students entered the program later.

Duration: (4Y) The program lasted through grade 3 (Chetty et al. [2010]).

Target: (SES) The program targeted schools with low SES (Chetty et al. [2010]).

Selection: (Program) The program selected schools to participate in the program (Chetty et al. [2010]).

Follow-Up (22): Chetty et al. [2010] follow students through age 27.

Sample Size: (10,992)

- Chetty et al. [2010] analyze a sample of 10,992.
- Dee and West [2008] analyze a sample of 2,212 and 2,978.

Intervention Components

Home: (No) The intervention reduced the class size of the students (Chetty et al. [2010]).

Health: (No) The intervention reduced the class size of the students (Chetty et al. [2010]).

Parental: (No) The intervention reduced the class size of the students (Chetty et al. [2010]).

On Site: (Yes) The program worked through schools (Chetty et al. [2010]).

Group: (Yes) Students in the treatment groups were put together (Chetty et al. [2010]).

Effects on Outcomes (Note: The intervention itself reduced class size. Chetty et al. [2010] also use the experimental design consider the effects of allocating students to different quality kindergarten classes (as measured by their performance on achievement tests at the end of kindergarten). In the summaries below, we consider the effects of class size and class quality.)

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Weak) Effects on test scores fade out by grade 8 for both class quality and class size (Chetty et al. [2010]).

Non-Cognitive: (Positive) Dee and West [2011] find that lower grade 8 class sizes lead to better non-cognitive skills. The effects persist for at least two years. Chetty et al. [2010] find that better class quality is associated with better non-cognitive skills.

Education: (Positive) Students in small classes are more likely to be enrolled in college at age 20 (Chetty et al. [2010]). Students enrolled in kindergarten classes that are higher quality (as measured by achievement test scores at the end of the class) are more likely to attend college and attend higher quality college (Chetty et al. [2010]).

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (Not measured) Crime was not measured in the studies that we reviewed.

Earnings: (Positive) At age 27, a standard deviation increase in class quality was associated with a $1,520 (9.6%) increase in earnings. Class size had a modest effect on earnings (Chetty et al. [2010]).

Return/Benefits

Rate of Return: (Not reported) Krueger [2003] estimates a 6.2% internal rate of return to the reductions in class size.

Benefit-Cost Ratio: (Not reported) The cost-benefit ratio was not estimated in the studies that we reviewed.

LA's Best

Participant Characteristics

Age: (5-6) The program begins as early as kindergarten and goes through 5th grade (Huang et al. [2000]). Students might enter the program after kindergarten.

Duration: (6Y) The program is an after school program that could run for up to 6 years if students start in kindergarten and remain in the program for the entire time (Huang et al. [2000]).

Target: (SES, Crime)The schools are chosen based on "low academic performance and their location in low-income, high-crime neighborhoods" (Goldschmidt and Huang [2007]).

Selection: (School) Schools offer the program and students choose to enter the program (Goldschmidt and Huang [2007]).

Follow-Up: The follow-up differed by cohort, but the longest was 12 years after the start of the program (Goldschmidt and Huang [2007]).

Sample Size: (19,322)

- Huang et al. [2000] use a sample of 19,322.
- Huang et al. [2005] use a sample size of 11,643.
- Goldschmidt and Huang [2007] use a sample size of 6,570, although some variables were missing in different years

Intervention Components

Home: (No) The program is an afterschool program and does not explicitly include visit the homes of the children (Huang et al. [2000]).

Health: (Yes) The curriculum includes teaching about drug prevention, exercise, and eating healthfully (Goldschmidt and Huang [2007]).

Parental: (Yes) The program includes programs for parents, field trip supervision, celebrations, etc… (Goldschmidt and Huang [2007]).

On Site: (Yes) The program takes place on school grounds after school (Huang et al. [2000]).

Group: (Yes) The program groups students together after school (Huang et al. [2000];Huang et al. [2005];Goldschmidt and Huang [2007]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in any of the studies (Huang et al. [2000];Huang et al. [2005];Goldschmidt and Huang [2007]).

School: (Weak) Goldschmidt and Huang [2007] find that the treatment group performs better on achievement tests than a control group that does not attend the same school.

Non-Cognitive: (Not measured) Non-cognitive skills were not measured in the studies that we reviewed.

Education: (Positive) Huang et al. [2005] find that the program reduced high school dropout rates by approximately 10% .

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (None) Goldschmidt and Huang [2007] claim that the program reduced crime for those with more intensive treatment. This finding neglects that treatment intensity is selective and that

those with lower than average intensity had higher crime rates than the control group. Crime rates for the overall treatment and control groups were virtually the same.

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed.

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (0.86) Goldschmidt and Huang [2007] conduct a cost-benefit analysis in which the benefits are assumed to come through the reduction in crime. For the sample average, they estimate a benefit-cost ratio of 0.86.

Cambridge –Somerville Program (CSP)

Participant Characteristics

Age: (5-13) The participants were 5-13 when the program began (McCord [1978]).

Duration: (5Y) The treatment program lasted for approximately 5 years (McCord [1978]).

Target: (Behav) The boys were chosen based on their behavior and tendency for delinquency (McCord [1978]).

Selection: (Referral) Churches, welfare agencies, police, and schools recommended candidates to participate in the program (McCord [1978]).

Follow-Up: (35Y) The follow-up lasted 30 years after the end of the program (McCord [1978]).

Sample Size: (506) The sample size is 506 (McCord [1978]).

Intervention Components

Home: (Yes) Counselors would visit the families approximately twice per month (McCord [1978]).

Health: (Yes) Over 100 received medical or psychiatric attention through the program (McCord [1978]).

Parental: (Yes) They strongly encouraged the families to reach out to the program for assistance (McCord [1978]).

On Site: (Yes) About ¼ of the children were sent to summer camps (McCord [1978]).

Group: (Yes) The children were often sent to camp with other children, but they were not necessarily joining children in the treatment group (McCord [1978]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Not Measured) School performance was not measured in the studies that we reviewed (McCord [1978]).

Non-Cognitive: (None) The survey included some measures of the beliefs and attitudes. The treatment and control groups did not differ on their responses to this question (McCord [1978]).

Education: (Not measured) Educational attainment was not measured in the studies that we reviewed.

Health: (Negative) More members of the treatment group showed signs of alcoholism, died at younger ages, were more likely to have severe mental illnesses, and were more likely to have stress-related diseases.

Crime: (None) There were no differences in criminal behavior, except that the members of the treatment group were more likely to commit multiple crimes (McCord [1978]).

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed (McCord [1978]).

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) The benefit-cost ratio was not estimated in the studies that we reviewed.

Seattle Social Development Project (SSDP)

Participant Characteristics

Age: (6-7) The full program began in first grade, so we assume that most students were 6 or 7. The program also had a late program that began in fifth grade (Hawkins et al. [1999]).

Duration: (6Y) The full program lasted through sixth grade for a total of 6 years (Hawkins et al. [1999]).

Target: (Crime) The program targeted high-crime areas in Seattle, Washington (Hawkins et al. [1999]).

Selection: (Program) 18 public schools were selected and then individuals were asked to join (Hawkins et al. [1999]).

Follow-Up: (21Y) Hawkins et al. [2008] follows students through age 27.

Sample Size: 605)

- Hawkins et al. [2008] uses a sample size of 598.
- Hawkins et al. [1999] uses a sample size of 598.
- Hawkins et al. [2005] uses a sample size of 605.

Intervention Components

Home: (No) The program did not include any explicit visits to the home (Hawkins et al. [1999]).

Health: (Yes) The program included a class aimed at parents to reduce drug use (Hawkins et al. [1999]).

Parental: (Yes) The program offered several voluntary parental classes (Hawkins et al. [1999]).

On Site: (Yes) The program primarily trained teachers, which would have affected children while they were at school (Hawkins et al. [1999]).

Group: (Yes) Entire schools participated in the program, so the participants would have been grouped together (Hawkins et al. [1999]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Positive) At age 18, the program increased GPA and reduced grade retention (Hawkins et al. [1999]).

Non-Cognitive: (Positive) The program improved self-efficacy at age 21 and 24 (Hawkins et al. [2005];[2008]).

Education: (Positive) It improved the rate of getting associates degrees at age 24 and age 27 (Hawkins et al. [2008]).

Health: (Positive) The program reduced the Mental Disorder Index at age 24 and age 27 (Hawkins et al. [2008]). The program reduced depression at age 21 (Hawkins et al. [2005]).

Crime: (Weak) The program reduced the amount of violent crime committed at the age 18 follow-up. It, however, did not reduce arrest rates in the long-run (through age 27) (Hawkins et al. [1999];Hawkins et al. [2005];[2008]).

Earnings: (None) The program improved earnings but the difference was not statistically significant (Hawkins et al. [2008]).

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (3.14) Aos et al. [2004] estimate a benefit-cost ratio of 3.14.

Big Brothers Big Sisters (BBBS)

Participant Characteristics

Age: (10-16) The age range in the study was 10 to 16 (Tierney et al. [1995]).

Duration: (1 Y) The follow-up for the study was 18 months, so that at the time of follow-up, participants would have received 18 months of treatment. It is possible for people to participate in the program for a longer period of time, but we use 18 months because that is the relevant amount of treatment for the outcomes analyzed in the evaluation (Tierney et al. [1995]).

Target: (SES) The participants frequently come from low-income families, broken families, and households with substance abuse or violence (Tierney et al. [1995]).

Selection: (Child) Children applied to the program. Then some were randomly selected of the waitlist to be in the treated group and some were randomly selected to be in the control group (Tierney et al. [1995]).

Follow-Up (1Y): The follow-up was 18 months (Tierney et al. [1995]).

Sample Size: (959) The evaluation had a sample size of 959 (Tierney et al. [1995]).

Intervention Components

Home: (No) The program is very flexible, as mentors and mentees can choose their own activities. It is possible that mentors would have visited the home of the mentee, but home visits are not emphasized in the program (Tierney et al. [1995]).

Health: (No) It is possible that mentors discussed drugs or other health issues but it is not a focus of the program (Tierney et al. [1995]).

Parental: (No) Parental engagement is not a standard practice for the program (Tierney et al. [1995]).

On Site: (No) The program does not involve a particular site like a school or other facility (Tierney et al. [1995]).

Group: (No) The program features one-on-one mentoring (Tierney et al. [1995]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed (Tierney et al. [1995]).

School: (Mixed) The program improved GPA and perceived ability to complete school work for girls but not boys (Tierney et al. [1995]).

Non-Cognitive: (None) The program did not improve Global Self-Worth, Social Acceptance, or Self-Confidence (Tierney et al. [1995]).

Education: (Not measured) Educational attainment was not measured in the studies that we reviewed.

Health: (Mixed) The program reduced the probability of starting drug use for boys but not for girls (Tierney et al. [1995]).

Crime: (None) The program did not affect whether students damaged property or stole something. As noted above, it did reduce the probably of drug use for boys (Tierney et al. [1995]).

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed.

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: Aos et al. [2004] estimates a benefit cost ratio of 1.01.

I Have a Dream (IHAD)

Participant Characteristics

Age: (11-12) The program began in 6th grade, so we assume that students were between 11 and 12 when entering the program (Kahne and Bailey [1999]).

Duration: (7Y) The program started in 6th grade and continued through HS. The students formed relationships with the staff that might have lasted longer (Kahne and Bailey [1999]).

Target: (SES) The program is targeted at low-SES schools (Kahne and Bailey [1999]).

Selection: (Program) The program randomly picks a 6th grade class in a disadvantaged school (Kahne and Bailey [1999]).

Follow-Up (8Y): The last interview occurred on year after graduating from high school for one sample and the year another sample began college (Kahne and Bailey [1999]).

Sample Size: (181) Kahne and Bailey [1999] use a sample size of 181.

Intervention Components

Home: (No) Home visits are not explicitly part of the program, but it is possible that some of the mentors in the program did visit the home. The evaluation gives an example where a mentor visited the household to confront one of the parents about an alcohol problem (Kahne and Bailey [1999]).

Health: (Yes) The program includes access to health services (Kahne and Bailey [1999]).

Parental: (Yes) Mentors would interact with parents to convince them of the importance of college (Kahne and Bailey [1999]).

On Site: (Yes) The evaluated programs were based in a youth organization and church space (Kahne and Bailey [1999]).

Group: (Yes) Participants engaged in activities with each other (Kahne and Bailey [1999]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Not measured) School performance was not measured in the studies that we reviewed.

Non-Cognitive: (Not measured) Non-cognitive skills were not measured in the studies that we reviewed.

Education: (Positive) The program doubled high school graduation rates and tripled college attendance rates (Kahne and Bailey [1999]).

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (Not measured) Crime was not measured in the studies that we reviewed.

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed.

Return/Benefits

Rate of Return: (Not reported) The rate of return of was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) The benefit-cost ratio was not estimated in the studies that we reviewed.

EPIS

Participant Characteristics

Age: (13-15) The participants are typically 7th and 8th graders who were between 13 and 15 (Martins [2010]).

Duration: (3Y) The program lasted for 3 years (Martins [2010]).

Target: (Academics) Students are recruited into the program based on academic performance (Martins [2010]).

Selection: (School) Students are recommended by teachers and assessed by the program staff (Martins [2010]).

Follow-Up: The follow-up was 3 years (Martins [2010]).

Sample Size: (45,069) The main results use a sample of 45,069 (Martins [2010]).

Intervention Components

Home: (No) The program took place entirely at school (Martins [2010]).

Health: (No) The program focused on academic and non-cognitive skills (Martins [2010]).

Parental: (No) Parents had to consent for their child to participate, but parental involvement was not a primary focus (Martins [2010]).

On Site: (Yes) The program takes place within schools (Martins [2010]).

Group: (Yes) The program includes one-on-one or small group meetings (Martins [2010]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Positive) The program improved class performance and reduced grade retention by at least 10 percentage points (Martins [2010]).

Non-Cognitive: (Not measured) Non-cognitive skills were not measured in the studies that we reviewed.

Education: (Not measured) Education was not measured in the studies that we reviewed.

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (Not measured) Crime was not measured in the studies that we reviewed.

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed.

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (0.9-3.0) Martins [2010] estimates a benefit-cost ratio of 0.9-3.0, depending on the value used for the number of students who are not retained as a result of the program.

Xl Club

Participant Characteristics

Age: (14) Participants were age 14 when they entered the program (Holmlund and Silva [2009]).

Duration: (2 Y) The program began at age 14 and lasted until age 16 (Holmlund and Silva [2009]).

Target: (School) The program focused on students who were at-risk of educational exclusion (Holmlund and Silva [2009]).

Selection: (School) Selection occurred in two stages. First, the organization that runs the program, "Princes Trust," selected schools. Then teachers and xl advisors selected students to participate (Holmlund and Silva [2009]).

Follow-Up: (2Y) The follow-up test was given at age-16, 2 years after the start of the program (Holmlund and Silva [2009]).

Sample Size: (261,422) There are 2,233 treated youths and 259,189 potential controls (Holmlund and Silva [2009]).

Intervention Components

Home: (No) The instruction takes place within school like a class and there is no mention of home visits (Holmlund and Silva [2009]).

Health: (No) Health was not explicitly listed as one of the main goals of the program (Holmlund and Silva [2009]).

Parental: (No) The program did not actively engage parents (Holmlund and Silva [2009]).

On Site: (Yes) The instruction takes place within school like a class (Holmlund and Silva [2009]).

Group: (Yes) Participants are grouped together (Holmlund and Silva [2009]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (None) The program did not improve achievement test scores (Holmlund and Silva [2009]).

Non-Cognitive: (Not measured) Non-cognitive skills were not measured in the studies that we reviewed.

Education: (Not measured) Education was not measured in the studies that we reviewed.

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (Not measured) Crime was not measured in the studies that we reviewed.

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed.

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the evaluations that we reviewed.

Benefit-Cost Ratio: (Not reported) The benefit-cost ratio was not estimated in the evaluations that we reviewed.

Sponsor-a-Scholar (SAS)

Participant Characteristics

Age: (14-15) The participants entered in 9th grade, which we assume corresponds to age 14-15 (Johnson [1999]).

Duration: (5Y) The program lasts from 9th grade to the year after high school (Johnson [1999]).

Target: (School, SES) The program targets disadvantaged but middle-achieving (grades in the B-C range) and motivated students (Johnson [1999]).

Selection: (School) Students are nominated by teachers and staff (Johnson [1999]).

Follow-Up: (6Y) The duration varied depending on the cohort, but the longest was two years after high school (Johnson [1999]).

Sample Size: (434) The evaluation contains 434 people (Johnson [1999]).

Intervention Components

Home: (No) Some of the mentors visited homes, but this was not a requirement or primary component of the program (Johnson [1999]).

Health: (No) The program was a mentoring program that focused on academic performance. It did not emphasize health (Johnson [1999]).

Parental: (No) The mentors might engage with the parents, but this was not a primary part of the program (Johnson [1999]).

On Site: (No) While the students occasionally meet as a group, there is not a consistent site where all students met (Johnson [1999]).

Group: (No) While all the students in a class meet periodically, the group meetings are not the main component of the intervention (Johnson [1999]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Weak) The program increased 10^{th} and 11^{th} grade GPAs but the effects faded by 12^{th} grade (Johnson [1999]).

Non-Cognitive: (None) The program did not increase measured Self-Esteem (Johnson [1999]).

Education: (Positive) Participants were more likely to attend college during the first and second years after high school (Johnson [1999]).

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (Not measured) Crime was not measured in the studies that we reviewed.

Earnings: (Not measured) Earnings were not measured in the studies that we reviewed.

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) The benefit-cost ratio was not estimated in the studies that we reviewed.

The Summer Training and Education Program (STEP)

Participant Characteristics

Age: (14-15) The program targeted 14-15 year olds (Walker and Vilella-Velez [1992]).

Duration: (2Y) The program lasted for two summers but also had some components during the school year (Walker and Vilella-Velez [1992]).

Target: (SES, School) The program targeted poor, inner-city children with academic problems (Walker and Vilella-Velez [1992]).

Selection: (Self) Half of the participants enrolled from announcements another half from personal outreach specially targeted mailings and intensive in-school recruitment (Walker and Vilella-Velez [1992]).

Follow-Up: (4Y) The follow-up is 42-54 months (Walker and Vilella-Velez [1992]).

Sample Size: (4,800) The sample is 4,800 (Walker and Vilella-Velez [1992]).

Intervention Components

Home: (No) P/PV site coordinators, for example, routinely called participants at home (sometimes weekly) and sometimes made home visits to persuade them to reenroll. Home visitation was, however, not a primary part of the program (Walker and Vilella-Velez [1992]).

Health: (Yes) As part of the “life-skills” component, the program covered material relating to sex and health (Walker and Vilella-Velez [1992]).

Parental: (No) The program does not explicitly attempt to affect interactions between parents and the children. Although, it is possible that the site coordinators would have interacted with the parents to some extent (Walker and Vilella-Velez [1992]).

On Site: (Yes) The program partially took place at training and employment agencies (Walker and Vilella-Velez [1992]).

Group: (Yes) The program had a remediation and life skills component that would have grouped the students together (Walker and Vilella-Velez [1992]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (None) The program had short-run impacts on reading and math score performance that faded (Walker and Vilella-Velez [1992]).

Non-Cognitive: (Not measured) Non-cognitive skills were not measured in the studies that we reviewed.

Education: (None) The dropout rates and college attendance rates were similar to non-participants (Walker and Vilella-Velez [1992]).

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (Not measured) Crime was not measured in the studies that we reviewed.

Earnings: (None) Earnings were not measured but the program had no effect on employment (Walker and Vilella-Velez [1992]).

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) The benefit-cost ratio was not estimated in the studies that we reviewed.

Quantum Opportunity Program (QOP)

Participant Characteristics

Age: (14-15) The program began in 9th grade, so we assume that most participants were between 14 and 15 (Rodríguez-Planas [2012]).

Duration: (5Y) The program lasted 5 years and was available year-round (Rodríguez-Planas [2012]).

Target: (School) To be eligible for the program, students must be enrolled in 9th grade, were not overly disabled, had a 8th grade GPA below the 67th percentile (Rodríguez-Planas [2012]).

Selection: (Program) The QOP staff contacted students and parents within the group of eligible students (Rodríguez-Planas [2012]).

Follow-Up: (10Y) The follow-up was 10 years after the start of the program started (Rodríguez-Planas [2012]).

Sample Size: (1,069) The sample size is 1,069 (Rodríguez-Planas [2012]).

Intervention Components

Home: (No) The program was offered after school and there is no mention of home visits in the evaluations (Rodríguez-Planas [2010];[2012])

Health: (Yes) The program included discussions about family planning, nutrition, personal hygiene, and avoiding drugs (Rodríguez-Planas [2010]).

Parental: (No) The program required parental consent but did not explicitly engage the parents in any of the activities. Rodríguez-Planas [2010] discusses the possibility that parents spent less time with their children, because they believed that the program would take care of some of their responsibilities.

On Site: (Yes) The program is described as an after-school program and featured tutoring and mentoring (Rodríguez-Planas [2010]).

Group: (Yes) The program included group discussions. Rodríguez-Planas [2012] discusses the possibility that the group aspect of the program could have induced negative peer effects.

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (None) Participants did not have statistically significantly higher scores on math or reading test scores at the 5-year follow-up. Test scores were not measured during other follow-ups. (Rodriguez-Planas [2012]).

Non-Cognitive: (Not measured) Non-cognitive skills were not measured in the studies that we reviewed.

(Rodríguez-Planas [2012]).

Education: (Weak) At the 5-year follow-up, females in the treatment were more likely to have graduated from high school, but these effects were not present in the 7-year or 10-year follow-ups. At the 10-year follow up males in the treatment group were more likely to have completed two years of college or training, but were not more likely to have earned a bachelor's degree (Rodríguez-Planas [2012]).

Health: (None) The program reduced binge drinking for males at the 7-year follow-up but not at either the 5-year follow-up or the 10-year follow-up. It had no effect on binge drinking for females at any year. It had no statistically significant effect on illegal drug use at any age for males or females. The effect on binge-drinking at the 7-year follow-up for males seems to be anomalous, so we categorize the effects on health as "none" (Rodríguez-Planas [2012]).

Crime: (Negative) At the 5-year follow-up, males in the treatment group were less likely to have been arrested or charged, but this effect reversed by the 10-year follow up. For this reason, we classify the effect as negative (Rodríguez-Planas [2012]).

Earnings: (None) The program had a small but negative impact on male wages at the 7-year follow-up, but there were no other statistically significant differences between the treatment and control groups (Rodríguez-Planas [2012]).

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in any of the studies that we reviewed.

Benefit-Cost Ratio: (0.42) Aos et al. [2004] estimate that the Benefit-to-Cost Ratio is 0.42. They consider the benefits of the program through high school graduation and through public assistance only. They also rely on evaluations that occurred before the long-term evaluation of Rodríguez-Planas [2012].

Career Academies

Participant Characteristics

Age: (13-16) The program typically starts in 9th or 10th grade (Kemple and Willner [2008]). We assume that this corresponds to ages between 13 and 16.

Duration: (4Y) The program starts in 9th grade and ends in 12th grade (Kemple and Willner [2008]).

Target: (School, SES) The academies target students who are at risk of being disengaged and dropping out of school and tend to be in relatively poor areas with high dropout rates (Kemple and Willner [2008]).

Selection: (Self) Students first applied to the program. There were more applicants than spots and some applicants were randomly assigned to treatment, while others were assigned to be in the control group (Kemple and Willner [2008]).

Follow-Up: (12Y) The follow-up was 8 years after the scheduled graduation (Kemple and Willner [2008]).

Sample Size: (1,458) The sample size for the long-term analysis was 1,458 (Kemple and Willner [2008]).

Intervention Components

Home: (No) Career academies are located within high schools and the program does not include explicit home visits (Kemple and Willner [2008]).

Health: (No) The main focus of career academies is to expose students to work-related experiences (Kemple and Willner [2008]).

Parental: (No) The evaluation does not mention parental involvement (Kemple and Willner [2008]).

On Site: (Yes) Part of the intervention takes place at high schools (Kemple and Willner [2008]).

Group: (Yes) Participants are grouped together in classes and other activities (Kemple and Willner [2008]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Weak) The academies did not impact scores on standardized achievement tests, but it did increase the number of credits earned (Kemple and Snipes [2000]).

Non-Cognitive: (None) There were no direct measures of non-cognitive traits, but the program did not affect risky behaviors during high school (Kemple and Snipes [2000]).

Education: (None) Career academies had no effect on high school completion or post-secondary educational attainment (Kemple and Willner [2008]).

Health: (None) The only health outcome measured was access to health insurance, which did not differ between the treatment and control groups (Kemple and Snipes [2000]).

Crime: (Weak) By 12^{th} grade high-risk students in the treatment group were less likely to have been arrested, but not for the medium- or low- risk groups (Kemple and Snipes [2000]).

Earnings: (Mixed) The program increased earnings for males in the long-run (Kemple and Willner [2008]).

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in any of the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) The benefit-cost ratio was not estimated in any of the studies that we reviewed.

National Guard ChalleNGe Program (ChalleNGe)

Participant Characteristics

Age: (16-18) Candidates are required to be between age 16 and 18 (Millenky et al. [2011]).

Duration: (1Y) The program lasts for 17 months (Millenky et al. [2011]).

Target: (Dropout) The program targets high school dropouts (Millenky et al. [2011]).

Selection: (Indiv) Individuals apply to the program and then may be accepted. The program does not reach out to them first. The treatment and control groups for the evaluation were created by randomly assigning people on a waitlist to enter the program or a control group.

Follow-Up: (3Y) The longest follow-up was three years after the start of the program (Millenky et al. [2011]).

Sample Size: (1,196)

- Bloom et al. [2009] use a sample of 1,018.
- Millenky et al. [2010] use a sample of 1,196.
- Millenky et al. [2011] use a sample size of 1,173.

(Yes) For the 3-year follow-up, the analysis sample size was 1,173 (Millenky et al. [2011]).

Intervention Components

Home: (No) The program includes a postresidential phase but it does not include explicit home visits. This posresidential phase seems to take place onsite. Millenky et al. [2011] suggest that one way to improve the postresidential phase would be to make it more offsite.

Health: (Yes) The program includes components that focus on health, physical fitness, and hygiene (Millenky et al. [2011]).

Parental: (No) The program does not explicitly engage parents (Millenky et al. [2011]).

On Site: (Yes) Much of the program takes place in a residential facility (Millenky et al. [2011]).

Group: (Yes) The participants live together and participate in group activities at the residential facility (Millenky et al. [2011]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Not measured) Achievement tests and grades are not measured (Millenky et al. [2011]).

Non-Cognitive: (Weak) At the 9-month follow-up, participants were more likely to have a high rating for self-efficacy and less likely to have a low rating for self-efficacy (Bloom et al. [2009]). The later surveys did not contain direct measures of non-cognitive skills (Bloom et al. [2009];Millenky et al. [2010];Millenky et al. [2011]).

Education: (Positive) The program had an initial positive effect on HS diplomas but the effects faded by the 3-year survey (Bloom et al. [2009];Millenky et al. [2010];Millenky et al. [2011]). The program had a positive effect on GED receipt during all follow-ups, but this likely results in few labor market benefits (Bloom et al. [2009];Millenky et al. [2010];Millenky et al. [2011]). The program had positive effects for any college credit measured during the 21-month and 3-year survey (Millenky et al. [2010];Millenky et al. [2011]). The 3-year follow is too short to know whether the treatment group completed their post-secondary education.

Health: (Mixed) At the first follow-up participants were less likely to be obese (Bloom et al. [2009]). At the last follow-up participants were more likely to be overweight (Millenky et al. [2011]).

Crime: (Weak) At the first follow-up, the program led to lower-levels of arrests and convictions (Bloom et al. [2009]). At the second follow-up, the program led to lower levels of convictions (Millenky et al. [2010]). By the third follow-up, there were no differences in arrests or convictions (Millenky et al. [2011]). The early reduction in crime likely stems from the fact that the treatment group lived in a residential facility, which may have made it more difficult to commit crimes.

Earnings: (Positive) At both the second and third follow-ups there is a statistically significant increase in earnings. (Millenky et al. [2010];Millenky et al. [2011]).

Return/Benefits

Rate of Return: Perez-Acre et al. [2012] estimate that the program has a rate of return of 6.4. They assume a discount rate of 3% and use data from the NLSY79 to estimate the returns to education.

Benefit-Cost Ratio: Perez-Acre et al. [2012] estimate that the program has a cost-benefit ratio of 2.66. They assume a discount rate of 3% and use data from the NLSY79 to estimate the returns to education.

Job Corps

Participant Characteristics

Age: (16-24) Job Corps provides assistance to disadvantaged youth between age 16 and 24 (Schochet et al. [2001]).

Duration: (1Y) The program length varies by participant but the average is around 8 months (Schochet et al. [2008]).

Target: (SES) The program targets disadvantaged youths. To qualify for Job Corps, participants must satisfy several criteria relating to socioeconomic status (Schochet et al. [2008]).

Selection: (Self) Youths apply to the program and then are selected to participate (Schochet et al. [2001]).

Follow-Up: (9Y) For most, the enrollment period was in 1995 and the final year of earnings data was in 2003 (Schochet et al. [2008]).

Sample Size: (15,309)

- Schochet et al. [2001] have sample sizes of 10,448, 10,405, and 9,937 at the 12-month, 30-month, and 48-month interviews.
- Schochet et al. [2008] use a sample of 15,309.

Intervention Components

Home: (No) The program does not include home visits (Schochet et al. [2001]).

Health: (Yes) Job Corps provides comprehensive health services, including medical exams, dental examinations, and counseling for mental illness (Schochet et al. [2001]).

Parental: (No) The program does not actively engage the parents of the participants (Schochet et al. [2001]).

On Site: (Yes) The program primary takes place at a residential facility (Schochet et al. [2001])..

Group: (Yes) The participants are grouped together at the residential facility (Schochet et al. [2001]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Not measured) School performance was not measured in the studies that we reviewed.

Non-Cognitive: (Not measured) Non-cognitive skills were not measured in the studies that we reviewed.

Education: (None) Job Corps participants were less likely to attend ABE/ESL, GED, High school, and Vocational school. They were similarly likely to attend two-year and four-year colleges Schochet et al. [2001])

Health: (Weak) There were small but statistically significant effects on self-reported health outcomes during the 12-month and 30-month surveys but not for the 48 month survey (Schochet et al. [2001]).

Crime: (Weak) There were improvements in arrests but primarily during the program and for the residential participants, but the effects faded near when the program ended (Schochet et al. [2001]).

Earnings: (Weak) Any impacts on earnings faded by the 3-year follow-up. The earnings improvements only offset the earnings losses (Schochet et al. [2008]).

Return/Benefits

Rate of Return: (Not reported) The rate of return was not estimated in any of the studies that we reviewed.

Benefit-Cost Ratio: (Not reported) Using a 3% discount rate, Schochet et al. [2006] estimate a benefit cost ratio of 0.22.

Year-Up Program

Participant Characteristics

Age: (18-24) Year-Up provides training to young urban adults age 18-24 (Roder and Elliot [2011]).

Duration: (1Y) The program lasts for one year (Roder and Elliot [2011]).

Target: (SES) The program targets disadvantaged youth (Roder and Elliot [2011]).

Selection: (Self) Participants applied to the program and some on the waitlist were placed as controls (Roder and Elliot [2011]).

Follow-Up: (2Y) The follow-up was two years after the start of the program (Roder and Elliot [2011]).

Sample Size: (195) The sample size was 195 (Roder and Elliot [2011]).

Intervention Components

Home: (No) The program does not include home visits (Roder and Elliot [2011]).

Health: (No) The program did not cover health topics (Roder and Elliot [2011]).

Parental: (No) The program does not actively engage the parents of the participants (Roder and Elliot [2011]).

On Site: (Yes) Participants attended classes and internships (Roder and Elliot [2011]).

Group: (Yes) Participants were grouped together during classes (Roder and Elliot [2011]).

Effects on Outcomes

IQ: (Not measured) IQ was not measured in the studies that we reviewed.

School: (Not measured) School performance was not measured in the studies that we reviewed.

Non-Cognitive: (Not measured) Non-cognitive skills were not measured in the studies that we reviewed.

Education: (Not measured) Education was not measured in the studies that we reviewed.

Health: (Not measured) Health was not measured in the studies that we reviewed.

Crime: (Not measured) Crime was not measured in the studies that we reviewed.

Earnings: (Positive) The program increased earnings primarily by increasing hourly wage (Roder and Elliot [2011]).

Return/Benefits

Rate of Return: The rate of return was not estimated in any of the studies that we reviewed.

Benefit-Cost Ratio: The benefit-cost ratio was not estimated in any of the studies that we reviewed.

W 7. Early Life Interventions that Begin before Formal Schooling

W 8. Education and Interventions in Kindergarten and Elementary School

W 9. Education and Interventions Targeted Toward Adolescents and Young Adults

W 10. Apprenticeship Programs

W 11. Other Curricula that Have Been Applied to Multiple Age Groups

W 12. The Effects of Education and Parental Investment on Character and Cognition

W 13. Summary

References

Aos, S., Lieb, R., Mayfield, J., Miller, M. and Pennuci, A. (2004). *Benefits and Costs of Prevention and Early Intervention Programs for Youth*. Olympia, WA, Washington State Institute for Public Policy.

Besharov, Douglas J., Germanis, Peter and Call, Douglas M. (2011). "Houston Parent Child Development Center." University of Maryland, School of Public Policy

Bloom, Dan, Gardenhire-Crooks, Alissa and Mandsager, Conrad L. (2009). "Reengaging High School Dropouts: Early Results of the National Guard Youth Challenge Program Evaluation." *MDRC Report, New York*.

Breitmayer, Bonnie J. and Ramey, Craig T. (1986). "Biological Nonoptimality and Quality of Postnatal Environment as Codeterminants of Intellectual Development." *Child Development* **57**(5): 1151-1165.

Bridgeman, Brent, Blumenthal, Janet B and Andrews, Susan R (1981). "Parent Child Development Center: Final Evaluation Report." ETS Research ED 213 764.

Carneiro, Pedro and Ginja, Rita (2012). "Long Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start.

Chetty, Raj, Friedman, John N., Hilger, Nathaniel, Saez, Emmanuel, Diane, Schanzenbach Whitmore and Yagan, Danny (2010). "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star." NBER Working Paper No. 16381.

Currie, Janet and Thomas, Duncan (1995). "Does Head Start Make a Difference?" *American Economic Review* **85**(3): 341-364.

Dee, Thomas S and West, Martin R (2011). "The Non-Cognitive Returns to Class Size." *Educational Evaluation and Policy Analysis* **33**(1): 23-46.

Dee, Thomas and West, Martin (2008). "The Non-Cognitive Returns to Class Size." NBER Working Paper No. 13994.

Deming, David (2009). "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal* **1**: 111-134.

Eckenrode, John, Campa, Mary, Luckey, Dennis W., Henderson, Charles R., Cole, Robert, Kitzman, Harriet, Anson, Elizabeth, Sidora-Arcoleo, Kimberley and Olds, David L. (2010). "Long-Term Effects of Prenatal and Infancy 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.

Garces, Eliana, Thomas, Duncan and Currie, Janet (2002). "Longer-Term Effects of Head Start." *American Economic Review* **92**(4): 999--1012.

Gertler, Paul, Zanolini, Arianna, Pinto, Rodrigo, Heckman, James, Walker, Susan, Vermeersch, Christel, Chang, Susan and Grantham-McGreror, Sally (2012). "Labor Market Returns to Early Childhood Stimulation: A 20-Year Follow-up to the Jamaica Study.

Goldschmidt, Pete and Huang, Denise (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." National Center for Research on Evaluation, Standards, and Student Testing (CRESST)

Grantham-McGregor, Sally M., Powell, Christine A., Walker, Susan P. and Himes, J. H. (1991). "Nutritional Supplementation, Psychosocial Stimulation, and Mental Development of Stunted Children: The Jamaican Study." *The American Journal of Clinical Nutrition* **54**: 642-648.

Hawkins, J. David, Catalano, Robert F., Kosterman, Rick, Abbott, Robert D. and Hill, Karl G. (1999). "Preventing Adolescent Health-Risk Behaviors by Strengthening Protection During Childhood." *Archives of Pediatrics and Adolescent Medicine* **153**: 226-234.

Hawkins, J. David, Kosterman, Rick, Catalano, Robert F., Hill, Karl G. and Abbott, Robert D. (2005). "Promoting Positive Adult Functioning through Social Development Intervention in Childhood." *Archives of Pediatrics and Adolescent Medicine* **159**(1): 25-31.

Hawkins, J. David, Kosterman, Rick, Catalano, Robert F., Hill, Karl G. and Abbott, Robert D. (2008). "Effect of Social Development Intervention in Childhood 15 Years Later." *Archives of Pediatrics and Adolescent Medicine* **162**(12): 1133-1141.

Heckman, James J., Malofeeva, Lena, Pinto, Rodrigo and Savelyev, Peter A. (2012a). "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." Unpublished manuscript. Department of Economics, University of Chicago.

Heckman, James J., Moon, Seong H. and Pinto, Rodrigo (2012b). "The Effects of Early Intervention on Abilities and Social Outcomes: Evidence from the Carolina Abecedarian Study." University of Chicago.

Heckman, James J., Moon, Seong Hyeok, Pinto, Rodrigo, Savelyev, Peter A. and Yavitz, Adam Q. (2010a). "Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the Highscope Perry Preschool Program." *Quantitative Economics* **1**(1): 1-46.

Heckman, James J., Moon, Seong Hyeok, Pinto, Rodrigo, Savelyev, Peter A. and Yavitz, Adam Q. (2010b). "The Rate of Return to the Highscope Perry Preschool Program." *Journal of Public Economics* **94**(1-2): 114-128.

Holmlund, Helena and Silva, Olmo (2009). "Targeting Non-Cognitive Skills to Improve Cognitive Outcomes: Evidence from a Remedial Education Intervention." IZA Discussion Paper No. 4476.

Huang, Denise, Gribbons, Barry, Kim, Kyung Sung, Lee, Charlotte and Baker, Eva L. (2000). "A Decade of Results: The Impact of the La's Best after School Enrichment Program on Subsequent Student Achievement and Performance." UCLA Center for the Study of Evaluation, Graduate School of Education and Information Studies

Huang, Denise, Kim, Kyung Sung, Marshall, Anne and Pï¿½rez, Patricia (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." National Center for Research on Evaluation, Standards, and Student Testing (CRESST)

Johnson, Amy W. (1999). "Assessing the Impact of the Sponsor-a-Scholar Youth Mentoring Program on Student Performance." Mathematica Policy Research 355.

Johnson, D. and Walker, T. (1991). "A Follow-up Evaluation of the Houston Parent Child Development Center: School Performance." *{J}ournal of {E}arly {I}ntervention* **14**(3): 226--236.

Kagitcibasi, Cigdem, Sunar, Diane and Bekman, Sevda (2001). "Long-Term Effects of Early Intervention: Turkish Low-Income Mothers and Children." *Journal of Applied Developmental Psychology* **22**: 333-361.

Kagitcibasi, Cigdem, Sunar, Diane, Bekman, Sevda, Baydar, Nazli and Cemalcilar, Zeynep (2009). "Continuing Effects of Early Enrichment in Adult Life: The Turkish Early Enrichment Project 22 Years Later." *Journal of Applied Developmental Psychology* **30**: 764-779.

Kahne, Joseph and Bailey, Kim (1999). "The Role of Social Capital in Youth Development: The Case of ``I Have a Dream'' Programs." *Educational Evaluation and Policy Analysis* **21**: 321-343.

Karoly, Lynn A., Kilburn, M. Rebecca and Cannon, Jill S. (2005). *Early Childhood Interventions: Proven Results, Future Promise*. Santa Monica, CA, RAND.

Kemple, J. and Willner, Cynthia (2008). "Career Academies: Long-Term Impacts on Labor Market Outcomes, Educational Attainment, and Transitions to Adulthood." MDRC

Kemple, James.J. and Snipes, Jason.C. (2000). "Career Academies: Impacts on Students' Engagement and Performance in High School." MRDC

Kitzman, Harriet, Olds, David L., Cole, Robert, Hanks, Carole, Anson, Elizabeth, Arcoleo, Kimberley, Luckey, Dennis W., Knudtson, Michael, Henderson, Charles R. and Holmberg, John (2010).

"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.

Krueger, Alan B. (2003). "Economic Considerations and Class Size." *Economic Journal* **113**(485): F34-F63.

Lally, J. Ronald, Mangione, Peter L. and Honig, Alice S. (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 Laboratory for Educational Research and Development.

Ludwig, Jens and Miller, Douglas L. (2007). "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Approach." *Quarterly Journal of Economics* **122**: 159-208.

Martins, Pedro S. (2010). "Can Targeted, Non-Cognitive Skills Programs Improve Achievement?" IZA Discussion Paper No. 5266.

McCord, Joan (1978). "A Thirty-Year Follow-up of Treatment Effects." *American PSychologyst* **33**: 284-289.

McCormick, Marie C., Brooks-Gunn, Jeanne, Buka, Stephen L., Goldman, Julie, Yu, Jennifer, Salganik, Mikhail, Scott, David T., Bennett, Forrest C., Kay, Libby L., Bernbaum, Judy C., Bauer, Charles R., Martin, Camilia, Woods, Elizabeth R., Martin, Anne and Casey, Patrick H. (2006). "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.

Millenky, Megan, Bloom, Dan and Dillon, Colleen (2010). "Making the Transition: Interim Results of the National Guard Youth Challenge Evaluation." MDRC May 2010.

Millenky, Megan, Bloom, Daniel, Muller-Ravett, Sara and Broadus, Joseph (2011). "Staying on Course: Three-Year Results of the National Guard Youth Challenge Evaluation." Manpower Demonstration Research Corporation

Niles, Michael D., Reynolds, Arthur J. and Nagasawa, Mark (2006). "Does Early Childhood Intervention Affect the Social and Emotional Development of Participants?" *Early Childhood Research and Practice* **8**(1): 32-48.

Olds, David L. (2006). "The Nurse-Family Partnership: An Evidence-Based Preventive Intervention." *Infant Mental Health Journal* **27**: 5-25.

Olds, David L., Kitzman, Harriet, Cole, Robert, Hanks, Carole, Arcoleo, Kimberley, Anson, Elizabeth, Luckey, Dennis W., Knudtson, Michael, Henderson, Charles R., Bondy, Jessica and Stevenson, Amanda J. (2010). "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, David L., Kitzman, Harriet, Cole, Robert, Robinson, J., Sidora, Kimberly, Luckey, Dennis W., Henderson, Charles R., Hanks, C., Bondy, J. and Holmberg, J. (2004a). "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, David L., Kitzman, Harriet, Hanks, Carole, Cole, Robert, Anson, Elizabeth, Sidora-Arcoleo, Kimberley, Luckey, Dennis W., Henderson, Charles R., Holmberg, John, Tutt, Robin A., Stevenson, Amanda J. and Bondy, Jessica (2007). "Effects of Nurse Home-Visiting on Maternal and Child Funcrtioning: Age 9 Follow-up of a Randomized Trial." *Pediatrics* **120**(4): 832-845.

Olds, David L., Robinson, J., Pettitt, Lisa, Luckey, Dennis W., Holmberg, J., Ng, R. K., Isacks, K., Sheff, K. and Henderson, Charles R. (2004b). "Effects of Home Visits by Paraprofessionals and by Nurses: Age 4 Follow-up Results of a Randomized Trial." *Pediatrics* **114**(6): 1560-1568.

Perez-Acre, Franisco, Constant, Louay, Loughran, David S. and Karoly, Lynn A. (2012). "A Cost-Benefit Analysis of the National Guard Youth Challenge Program." RAND Corporation RAND Technical Report,

Ramey, C. T., Collier, A. M., Sparling, J. J., Loda, F. A., Campbell, F. A., Ingram, D. L. and Finkelstein, N. W. (1976). "The Carolina Abecedarian Project: A Longitudinal and Multidisciplinary Approach to the Prevention of Developmental Retardation." in *Intervention Strategies for High-Risk Infants and Young Children*. T. Tjossem, ed. Baltimore, University Park Press**:** 629-665.

Reynolds, Arthur J. (1994). "Effects of a Preschool Plus Follow-on Intervention for Children at Risk." *Developmental Psychology* **30**(6): 787-804.

Reynolds, Arthur J. and Temple, Judy A. (1998). "Extended Early Childhood Intervention and School Achievement: Age Thirteen Findings from the Chicago Longitudinal Study." *Child Development* **69**(1): 231-246.

Reynolds, Arthur J., Temple, Judy A., Ou, Suh-Ruu, Arteaga, Irma A. and White, Barry A. B. (2011a). "School-Based Early Childhood Education and Age-28 Well-Being: Effects by Timing, Dosage, and Subgroups." *Criminal Justice and Behavior* **333**(6040): 360-364.

Reynolds, Arthur J., Temple, Judy A., Robertson, Dylan L. and Mann, E. A. (2002). "Age 21 Cost-Benefit Analysis of the Title I Chicago Child-Parent Centers." *Educational Evaluation and Policy Analysis* **24**(4): 267-304.

Reynolds, Arthur J., Temple, Judy A., White, Barry A. B., Ou, Suh-Ruu and Robertson, Dylan L. (2011b). "Age 26 Cost–Benefit Analysis of the Child-Parent Center Early Education Program." *Child Development* **82**(1): 379-404.

Roder, Anne and Elliot, Mark (2011). "A Promising Start: Year-Up's Initial Impacts on Young Adults' Careers." Economic Mobility Corporation

Rodríguez-Planas, Núria (2010). "Mentoring, Educational Services, and Economic Incentives: Longer-Term Evidence on Risky Behaviors from a Randomized Trial." IZA Discussion Paper No. 4968.

Rodríguez-Planas, Núria (2012). "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.

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

Schochet, Peter Z., Burghardt, John and McConnell, Sheena (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.

Schochet, Peter Z., Burghardt, John and McConnell, Sheena (2008). "Does Job Corps Work? Impact Findings from the National Job Corps Study." *American Economic Review* **98**(5): 1864–1886.

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

Temple, Judy A. and Reynolds, Arthur J. (2007). "Benefits and Costs of Investments in Preschool Education: Evidence from the Child–Parent Centers and Related Programs." *Economics of Education Review* **26**: 126-144.

Tierney, Joseph P., Grossman, Jean B. and Resch, Nancy L. (1995). *Making a Difference: An Impact Study of Big Brothers/Big Sisters*. Philadelphia, PA, Public/Private Ventures.

Walker, Gary C. and Vilella-Velez, Frances (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, Susan P., Chang, Susan M., Powell, Christine A. and Grantham-McGregor, Sally M. (2005). "Effects of Early Childhood Psychosocial Stimulation and Nutritional Supplementation on

Cognition and Education in Growth-Stunted Jamaican Children: Prospective Cohort Study." *Lancet* **366**: 1804-1807.

Walker, Susan P., Chang, Susan M., Powell, Christine A., Simonoff, Emily and Grantham-McGregor, Sally M. (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**: 2464-2469.

Walker, Todd. and Johnson, Dale L. (1988). "A Follow-up Evaluation of the Houston Parent-Child Development Center: Intelligence Test Results." *The Journal of genetic psychology* **149**(3): 377-381.

Westat (2010). "Head Start Impact Study: Final Report." {U.S.} Department of Health and Human Services Contract no. 282-00-0022.