IEG WORKING PAPER
2015/3

Later Impacts of Early Childhood Interventions:

A Systematic Review

IEG
Independent Evaluation Group
WORLD BANK GROUP
World Bank • IFC • MIGA

WHAT WORKS

IEG Independent Evaluation Group | WORLD BANK GROUP World Bank • IFC • MIGA

Later Impacts of Early Childhood Interventions:
A Systematic Review

IEG Working Paper 2015/3

Jeffery C. Tanner, Tara Candland, and Whitney S. Odden

Independent Evaluation Group, The World Bank Group
1818 H St., NW
Washington, DC 20433

IEG: Improving World Bank Group Development Results Through Excellence in Evaluation

The Independent Evaluation Group is an independent unit within the World Bank Group; it reports directly to the Bank's Board of Executive Directors. IEG assesses what works, and what does not; how a borrower plans to run and maintain a project; and the lasting contribution of the Bank to a country's overall development. The goals of evaluation are to learn from experience, to provide an objective basis for assessing the results of the Bank's work, and to provide accountability in the achievement of its objectives. It also improves Bank work by identifying and disseminating the lessons learned from experience and by framing recommendations drawn from evaluation findings.

IEG Working Papers are an informal series to disseminate the findings of work in progress to encourage the exchange of ideas about development effectiveness through evaluation.

ISBN-10: 1-60244-261-4
ISBN-13: 978-1-60244-261-0

Contact: IEG Communication, Learning and Strategies (IEGCS)
e-mail: ieg@worldbank.org
Telephone: 202-458-4497
Facsimile: 202-522-3125
http:/www.worldbank.org/ieg

Contents

INCLUDED BOXES

Abbreviations

BMI	body mass index
BMIZ	BMI for age z-score
CCT	conditional cash transfer
CPC	Child Parent Center
CSG	Child Support Grant
DQ	development quotient
ECCE	Early Childhood Care and Education
ECD	early childhood development
HAZ	height-for-age z-score
HC	Hogares Comunitarios
IBRD	International Bank for Reconstruction and Development
IDA	International Development Assistance
IEG	Independent Evaluation Group
INCAP	Instituto de Nutrición de Centroamérica y Panamá
IQ	intelligence quotient
MUAC	mid-upper-arm circumference
SIEF	Strategic Impact Evaluation Fund
PROBIT	Promotion of Breastfeeding Intervention Trial
RCT	randomized controlled trial
RPS	Red de Protección Social
SDQ	Strengths and Difficulties Questionnaire
SMP	Safe Motherhood Program
TEEP	Turkish Early Enrichment Project
TVIP	Test de Vocabulario en Imagenes Peabody
UCT	unconditional cash transfer
UNICEF	United Nations Children's Fund
UNIT	Universal Nonverbal Intelligence Test
WAZ	weight-for-age z-score
WAIS	Wechsler Adult Intelligence Scale
WASI	Wechsler Abbreviated Scale of Intelligence
WISC	Wechsler Intelligence Scale for Children
WPPSI	Wechsler Preschool and Primary Scale of Intelligence
WHZ	weight-for-height z-score
WHO	World Health Organization

Glossary

Estimate — The empirical result of the application of an impact evaluation (IE) methodology to a set or subset of data for an intervention. Multiple estimates are possible for the same intervention and study if multiple IE methods are used.

Impact evaluation — A methodology of empirical analysis allowing causal inference through the use of a counterfactual. For the purposes of counting impact evaluations in this review, the number of impact evaluations within a study is the number of unique interventions within that study that reports an estimate for one of the included outcomes. Multiple arms are counted separately.

Intervention — The most disaggregated combination of policy or project components for which there is a unique impact evaluation estimate; an arm of a randomized control trial or pieces of an intervention whose effect can be separately estimated in a quasi-experimental IE.

Intervention family — The broadest categorization of interventions used for this review (see figure 1.1).

Intervention type — The more specific categorization of interventions used for this review (see figure 1.2).

Outcome — The construct measured by an estimate.

Project — The full bundle of interventions carried out for a population over a period of time.

Study — An article, working paper, or other publication that has at least one unique estimate for use in this systematic review. For counting purposes, a publication with several effectiveness for multiple arms of a project are still counted as a single study.

Acknowledgments

This study was prepared by a team led by Jeffery C. Tanner. The core writing team consisted of Tara L. Candland, Whitney S. Odden, and Jeffery C. Tanner. The team benefited from excellent research assistance provided by Ryotaro Hayashi. In addition to these individuals, the search and coding team also included Tomoko Kato, Elizabeth Koshe, Yunsun Li, Karol Acon Monge, and Maya Sherpa. Valuable topical expertise was provided by Patricia Kariger. Editorial support was given by Barbara Rice and Bruce Ross-Larson. Yezena Yimer offered administrative support.

The team thanks Harold Alderman, Lucy Basset, Susan Caceres, Leslie Elder, Julieta Trias, and Quentin Wodon for insightful discussions and comments at various stages of the review. Members of the Independent Evaluation Group management team also provided useful comments, as did participants of presentations of the work from the Japan International Cooperation Agency, the University of Tokyo, and the World Bank. The team is particularly grateful to the government of Japan for partially funding this work through a Policy and Human Resources Development Grant.

Finally, the team wishes to thank and acknowledge the authors and implementers of the evaluations contained herein, together with their funders, for their vision and dedication in breaking new ground to aid in the understanding of the later-life effects of early childhood interventions.

The work was conducted under the general guidance of Mark Sundberg (manager), Emmanuel Jimenez (former director over the course of the majority of the work), and Nicholas York (acting director). The team is grateful for comments proffered by its peer reviewers, Lia Fernald and Sophie Naudeau.

Overview

Highlights

The economic rationale for investing in young children goes beyond improving quality of life during early childhood; it hinges on the belief that the benefits of these investments persist into school age and beyond. This report is the first systematic review devoted exclusively to investigating this theory.

By identifying and analyzing all 55 studies that provide reliable causal estimates, the report provides the most complete, credible evidence to date on the post-early childhood effects of early childhood interventions. It serves three important functions. First, it provides analysis on early childhood interventions whose sustained effects have been evaluated across six areas of human development. Second, it examines how effects change—both within a population (yielding shared prosperity implications) and over time. Finally, the review aims to improve the quality and coverage of ECD knowledge by enumerating commonly observed evaluation challenges and identifying research gaps on the question of benefits beyond early childhood.

Four important findings emerge from this effort:

- Early childhood interventions can, but do not always, lead to benefits later in life in the areas of cognition, language, socioemotional health, education, and the labor market. Evaluated interventions have not demonstrated consistent lasting advantages for physical development, although these outcomes are less salient to adult welfare.
- Gender-neutrality dominates outcomes generally, but schooling does tend to improve for girls, the poor, and those who are in quality preschool and supplemental feeding programs for longer.
- Nutrition interventions may need to be in place throughout and beyond the first 1,000 days in order to leverage the window of opportunity from conception to age 2 and achieve sustained effects beyond early childhood.
- Sizeable knowledge gaps persist but can be closed with careful planning and design.

Early childhood development holds considerable promise for making progress on the World Bank's dual objectives of reducing poverty and increasing shared prosperity while encouraging economic growth. More investment and evaluation are required to enable interventions, and the children they serve, to realize their potential.

In developed countries, well-documented evidence shows that interventions aimed at improving early childhood development (ECD) can play a major role in shaping the arc of young children's lives and livelihoods. This evidence prompts many in the international development community—including the World Bank—to focus attention on the years before children enroll in primary school or reach the age of six as the first step in a sequenced strategy to build the skills needed for productivity and economic growth.

In developing countries, the ability of interventions to improve outcomes beyond early childhood is less well studied. Impact evaluations in middle- and low-income countries are scarce; it may be problematic to assume findings from high-income countries apply to middle- and low-income economies.

In an effort to bridge the evidence gap, the Independent Evaluation Group (IEG) prepared a systematic review that gathers and analyzes the available impact evaluation evidence in developing countries from 1990 to 2015 on whether early childhood interventions shape future outcomes. Its purpose is not to supplant existing evidence but rather to help practitioners understand how evidence from impact evaluations supports or challenges beliefs about interventions and can be used to inform development policy.

This review aims to answer two questions:

- What is the evidence of attributable effects on outcomes in primary school and beyond from early childhood interventions in low- and middle-income countries?
- How do the post-early childhood effects of early childhood interventions vary by socio-economic status, gender, age at intervention (particularly the first 1,000 days from conception to age two), and age at evaluation?

Methodology and Scope

Most systematic reviews are intervention based, meaning they track the outcomes from a narrow set of interventions. This report, on the other hand, reviews all interventions from developing countries that occur during early childhood for which impact evaluation estimates exist for effects observed at primary school age and older. From an initial search of thousands of studies, the search process—which included database searches, hand searches, and snowballing—identified more than 500 ECD-related impact evaluations written from 1990 to 2015. After excluding clinical trials as well as outcome estimates for which at least part of the sample was still in the early childhood period, 116 studies remained. The team then analyzed the risk of bias for these 116 studies, including conducting a rigorous check of the credibility of the causal estimates, identifying assumptions, and assessing construct validity and representativeness. Studies were given a quality rating of A, AA, or AAA. The 55 studies with ratings of AA or AAA were included in the review.

Although this review contains all of the causally robust evaluations within its inclusion criteria, the evidence base for any given combination of interventions and outcomes may still be thin. Furthermore, there are many reasons why a study may yield a null result besides the intervention having no impact, including challenges of statistical power, contamination, attrition, uptake, and implementation. Therefore, IEG's aim is to elucidate what is known about the long-term effects of early childhood interventions and identify the remaining research gaps.

Findings

This review covers six areas—or domains—of human development: physical development, cognitive development, language development, socioemotional development, schooling outcomes, and employment and labor market outcomes. These domains are commonly included in evaluations of early interventions targeting poor children because they are negatively affected by early poverty, can benefit from early intervention, and are important for overall well-being or adult productivity. Some outcomes can be measured repeatedly starting from early childhood (for example, height and weight) while others are measurable only later in life (for example, cognition, schooling, and employment and labor market status).

PHYSICAL DEVELOPMENT

Some improvement is seen in height but no effect is found on weight.

Taken as a whole, physical development benefits appear difficult to sustain past early childhood. Cohort studies find strong correlations between early anthropometrics and later cognition, schooling and labor market outcomes, and the broader literature contains many examples of impact evaluations that demonstrate the ability of early health and nutrition interventions to improve physical development among children under age 6 in the near term. However, those effects appear to dissipate quickly beyond early childhood.

Despite the review of seven different intervention types in five geographical Regions of the World Bank over a range of ages from six to 17 years old, none had a lasting effect on weight.

The ability of early childhood interventions to change height may also be limited. Only 2 of 10 nutritional supplementation interventions improved height, and both were only marginally significant: Providing iron and zinc supplements to infants in Thailand improved height at six to eight years old; giving pregnant Nepalese mothers folic acid, iron, and zinc slightly improved children's height at age nine. Yet the eight nutrition interventions affected height either negatively or not at all.

The cash transfer from Mexico's Progresa, and the family planning and maternal and child health program in Matlab, Bangladesh, improved beneficiaries' height-for-age, and female beneficiaries of the South African Child Support Grant were significantly taller. It is difficult to draw firm conclusions from these results because other evaluations of cash transfers and early stimulation programs found no effect.

The third physical outcome—fine motor skills—has a rather thin evidence base but was included for its intrinsic value as a measure of school readiness. Fine motor skills improved in a Mozambican preschool program, but the Bucharest Early Intervention Program had no lasting effect on fine motor skills.

Even so, physical outcomes such as stunting and wasting are important not for their intrinsic value but as functional correlates of impeded cognitive development, school achievement, and future economic activity. Therefore, the finding that physical development benefits are rarely sustained past early childhood is less alarming in light of the many sustained benefits in the functional outcomes more closely tied to lifelong well-being.

COGNITIVE DEVELOPMENT

Stimulation interventions most consistently impacted outcomes over the spectrum of cognitive outcome measures; "general cognition" was improved by a wide range of interventions.

Cognitive ability reflects an individual's problem-solving and analytical skills, memory functions, general knowledge and ability to apply logic, and reactions to new situations. Cognitive outcomes are of particular interest in ECD research because of their influence on an individual's future productivity in areas such as schooling and the labor market.

The Nutrition Enigma

Small-scale and near-term studies as well as others from high-income countries have shown that nutrition interventions can lead to significant improvements in child development, including better morbidity, mortality, and cognitive outcomes.

Longitudinal studies, including those in developing countries, indicate a strong correlation between early nutritional status and later-life outcomes.

Yet these later-life effects in developing countries are less well-established. Across human development domains, nutrition interventions had little impact on post-early childhood outcomes.

The way forward for nutrition interventions may lie in both timing and duration. The only nutrition project that demonstrated sustained effects was also the only one that had a sustained intervention throughout (and beyond) the critical period of the first 1,000 days of a child's life from conception through age two.

Early sustained nutrition promises later sustained results.

Quantifying outcomes in cognition can be difficult because of the lack of a universal, standardized measurement of cognitive ability. Therefore, the reviewed studies report outcomes from a range of tests, from brief screening assessments to comprehensive full-scale intelligence tests. Given that it is inaccurate to compare results across different types of cognitive tests, the review groups results from similar assessment tools to compare outcomes across intervention types and contexts.

Improvements in both full-scale and abbreviated measures of general cognition were caused by four different intervention types and six separate programs: breastfeeding promotion in Belarus, stimulation-related interventions in Jamaica and Romania, CCTs in Mexico and Nicaragua, and a deworming program in Kenya.

The breastfeeding promotion program in Belarus improved IQ outcomes at age 6.5 years old, though the effects were only marginally significant. The stimulation component of the Jamaica program improved cognition scores for low birthweight children at age 6 and stunted children at the ages of 11–12, 17–18, and 22 years. In Romania, children who received greater stimulation by being randomly assigned to foster care had marginally higher full-scale IQ scores at age 8 years than children who had remained in institutional care. The cash component of the Progresa CCT in Mexico produced a marginally significant effect on general cognition between the ages of 8 and 10, and Nicaragua's conditional cash transfer improved cognition in 10-year-old boys. Siblings of children who had participated in a deworming program in Kenya also had measurable improvements between the ages of 8 and 15 years old.

While several interventions proved effective in improving nonverbal cognition, the evidence base for each intervention type is too thin to be more than suggestive. Only a single nutritional intervention, Guatemala's Instituto de Nutrición de Centroamérica y Panamá (INCAP) program, improved nonverbal cognition in the post-early childhood period. Stunted children who participated in the stimulation component of the Jamaica program scored higher in nonverbal cognition at ages 11–12 and 17–18. Only one of two social protection interventions, the CCT in Nicaragua, improved nonverbal outcomes. The deworming program in Kenya also improved nonverbal cognition among siblings of children who participated in the intervention.

There is similarly little evidence around executive function, a measure of cognition that reflects how an individual responds to new or challenging situations. While stimulation for low birthweight babies in Jamaica improved short-term memory at age six years, related measures did not show improvement at age 11–12 (processing speed) or 17–18 years old (working memory). The Romanian foster care program did not result in noticeable improvements in executive function across a range of measurements assessed at eight years of age. The Nicaragua CCT program improved executive function in 10-year-old boys but not girls.

Most nutrition interventions did not yield lasting cognitive benefits. The one (of six) that demonstrated strong impacts was also the only one to have supplied nutrition in a sustained fashion through (and in this case beyond) the first thousand days from conception to age two. On the other hand, both stimulation programs, Jamaica and Bucharest, improved four different cognitive outcomes, as did health access in Indonesia and deworming in Kenya.

LANGUAGE DEVELOPMENT

Early childhood interventions can improve language outcomes, although the evidence is mixed across intervention types.

During the second year of life (12–24 months), children experience a vocabulary explosion. As they enter the preschool years, vocabulary, spoken grammar, and sentence structure become more sophisticated, and children develop the ability to identify letters and, later, words. These skills are important for enabling them to read and do well in school.

Language outcomes were measured in three subdomains: verbal ability, reading and literacy, and vocabulary. The existing evidence suggests that stimulation, nutrition, and social protection programs can improve language. In addition, three nontraditional ECD interventions—deworming, sanitation, and governance—all improved language outcomes, although only one study was found for each type.

Two of the three stimulation programs included in this review improved long-term language outcomes. Verbal abilities seemed to be particularly sensitive to early stimulation programs: children who participated in the Romanian foster care program had significantly higher verbal abilities at age eight years, and gains were observed among participants of the Jamaican stimulation program at ages 11–12, 17–18, and 22 years. The stimulation component of the Jamaica program also produced lasting gains in reading and vocabulary. A second stimulation program in Mozambique, however, did not find measurable effects on beneficiaries' vocabulary.

Only two of the five nutritional interventions that measured long-term language outcomes found significant effects, although a single nutritional intervention—breastfeeding promotion in Belarus—improved outcomes for 6.5 year old children in all three language subdomains. A supplementation program in Guatemala also improved reading and literacy. However, the supplementation-only arm of the Jamaica intervention did not have a measurable effect on reading, and a nutritional intervention in The Gambia did not significantly improve children's vocabulary. These results suggest that supplementation alone might not be enough to produce sustained linguistic effects, despite solid rationale for expecting that various types of nutritional interventions aimed at children who are at risk for or suffering from deficiencies such as chronic malnutrition and iron deficiency could positively impact cognitive and language outcomes.

Social protection programs appear to have inconsistent effects on language outcomes. Although Mexico's Progresa CCT increased verbal abilities among beneficiaries and a CCT in Nicaragua produced measurable gains in vocabulary among boys, none of the three social protection interventions that measured reading and literacy produced significant improvements.

Non-traditional ECD interventions also proved effective in improving language outcomes, although the evidence is thin and mixed. In India, both a sanitation intervention and a governance campaign improved reading and literacy. Although the governance campaign did not have measurable effects on children's vocabulary, a deworming program in Kenya did significantly improve vocabulary among 15 year olds.

SOCIOEMOTIONAL DEVELOPMENT

Delayed improvement occurs in externalizing behavior, but little improvement is seen in internalizing behavior.

Social and emotional functioning involves the acquisition of the skills and knowledge required by a person to understand and manage emotions, set and achieve goals, empathize with others, establish and maintain positive relationships, and make responsible decisions. It encompasses a broad range of internalizing and externalizing behaviors, which indicate how people view themselves and how they react to the world around them, respectively. These skills are important in learning to cope with difficulties and succeed in various endeavors.

Early childhood interventions appear to have a delayed effect on externalizing behavior; the review found no consistent effects on young children, but benefits from different early stimulation programs seem to show up more as children age into adolescence and beyond. No early intervention has yet been able to demonstrate later improvements in a child's ability to pay attention, notwithstanding the efforts of the early stimulation intervention in Jamaica and the micronutrients intervention in Thailand. However, the Environmental Enrichment Program in Mauritius and the early stimulation program in Jamaica did have a positive effect on some elements of externalizing behavior for teenagers including violent conduct.

The early stimulation program in Jamaica also had a sustained effect on internalizing behavior, while the nutrition arm of that intervention, foster care in Romania, and the early enrichment intervention in Mauritius did not. Those who benefited from early stimulation in Jamaica exhibited lower anxiety and depression and greater self-esteem at 17 years old, a finding that persisted at age 20, although by that time it was largely driven by improvements among women. When participants reached the age of 22 years, the effect on anxiety had disappeared, but there was still a significant decrease in depression among beneficiaries.

The Strength and Difficulties Questionnaire, which is a brief screening test, is an additional measure of both internalized and externalized socioemotional behavior. Its results can change due to early childhood interventions. Mexico's conditional cash transfer program, an early stimulation program for low birthweight infants in Jamaica, and the early stimulation foster care program in Romania significantly reduced post-early childhood behavior problems as measured by the questionnaire.

SCHOOLING OUTCOMES

Early stimulation, preschool, and cash transfers appear to be most effective in improving educational outcomes.

Early childhood interventions could affect schooling through a number of possible pathways. For instance, improved cognitive development could result in increased scholastic achievement, while healthier children are better able to attend classes. Indeed, there is evidence that early childhood interventions can improve various educational outcomes.

Preschool programs and cash transfers appear effective in promoting on-time primary school enrollment. Of the four interventions evaluated, Mexico's Progresa and the Mozambican preschool program had a significant beneficial effect, while South Africa's Child Support Grant had no overall effect but did decrease the probability of delayed enrollment for girls and children whose mothers were less educated. A micronutrient supplementation program in Thailand, however, had no effect on on-time enrollment.

While the heterogeneity of interventions evaluated make it difficult to find a clear pattern in differences in the years of schooling completed, some general trends emerge to suggest that early stimulation and cash transfers (though not nutrition programs) may be effective in promoting more education. For example, early stimulation in Jamaica and preschool attendance in Uruguay increased schooling among participants. Similar results were seen among some beneficiaries of the Honduran, Mexican, and South African cash transfer. For nutrition, the INCAP supplementary feeding program in Guatemala had a large effect, but no effect was detectable from either the Jamaican supplementary feeding intervention or the maternal supplementation program in The Gambia. Remarkably, the largest improvement in schooling came from a clean water program in China.

Nutrition programs also did not have a detectable effect on school performance, but participants in early stimulation programs may be more likely to perform well and attend post-primary school. Only two programs were evaluated for their effect on post-primary attendance—*Hogares Comunitarios* in Colombia and the early stimulation intervention in Jamaica—both of which caused a large increase in the probability of attending. Furthermore, beneficiaries of the early stimulation in Jamaica, as well as the children who participated in a preschool program in Argentina and Chile's Early Childhood Care and Education programs did significantly better on their subject matter and standardized achievement tests. Conversely, in three different nutrition interventions, participants' test scores were not significantly different than their peers.

EMPLOYMENT AND LABOR MARKET OUTCOMES

Early stimulation can help.

The goal of many early childhood interventions is to improve human capital, and labor market outcomes offer an important measure of its fulfillment. Early childhood services devoted to enriching an individual's environment by increasing inputs in education, health, and nutrition can determine the nature of these outcomes.

Indeed, there is evidence that early stimulation as well as good nutrition can positively influence participants' subsequent labor market outcomes. Psychosocial stimulation in Jamaica dramatically increased earnings, especially among women and people with full-time jobs. Interestingly, additional (though flawed) evidence of an early stimulation training provided to mothers through the Early Enrichment Program in Turkey did not cause its participants to be better employed but did delay the starting age of employment—commonly associated with lifetime earnings—likely because it also increased the probability of attending college.

Heterogeneous Effects

Later-life effects are generally gender-neutral, but girls and those from poor families or more educated parents do tend to have better schooling outcomes.

Evidence for heterogeneous effects—or differences in outcomes due to individual characteristics—is reported in fewer than half of the studies. Although evidence is thin for specific outcomes, a few noteworthy trends emerge for the broader outcome domains.

Based on the available evidence, the later-life effects of early childhood intervention appear to be mostly gender-neutral, especially in the physical and socioemotional domains and for nutrition and early learning or childcare interventions; in other words, there is usually no significant difference in the benefits accrued to girls versus boys. However, girls are much more likely than not to benefit from interventions that affect schooling outcomes, and neither gender is likely to enjoy lasting physical benefits from interventions occurring in early childhood.

Conversely, for socioeconomic status, there are some groups that are significantly more likely to benefit than others. Interventions affecting physical outcomes appear to benefit the rich and the poor equally when they affect them at all, but the poor and those with better-educated parents benefit significantly more from interventions that improve schooling than do children from richer families or those with less-educated parents.

Time Effects

The persistence of effects over time varies by outcome domain. Interventions lasting at least the full first 1,000 days may be more effective, and additional exposure to some programs can be beneficial.

Three dimensions of time are evaluated for their effect on post-early childhood outcomes: temporal trajectories, age at exposure, and length of exposure. These elements are critical to consider when designing an intervention, but few studies examine these important elements, making it impossible to draw firm conclusions. Nevertheless, the preliminary findings drawn from the available evidence can help target future research.

In the first dimension—temporal trajectories—outcomes within a given intervention are traced across time to determine whether and how they change as a child ages. It appears that socioemotional benefits, particularly those that pertain to internalizing behavior, may fade over time, but cognitive benefits from an early stimulation program in Jamaica not only remained significant from 11 to 22 years old but actually increased in magnitude.

The evidence regarding the temporal trajectory of physical outcomes suggests that they tend to remain constant over time. Evaluations of the six different programs with IE results at multiple points in time found null effects for most estimates for height and weight—both in early childhood and later in life. In a few cases, however, initial anthropometric benefits during early childhood disappeared by the post-early childhood evaluations, and in one instance, six- to eight-year-old Nepali children whose mothers had received prenatal micronutrients were taller than their peers despite no detectable difference in length at birth.

To evaluate age at exposure, this review examines the first 1,000 days, a period that is believed to be critical for a child's development. Four impact evaluations specifically isolated the effect of treatment during this period compared to later periods, but because the interventions are so varied, it is difficult to draw firm conclusions. Children who began to receive South Africa's Child

Support Grant before they were two years old were not significantly taller or less stunted than those who began the program between two and five years old. In China, however, children who had access to clean water during the first 1,000 days completed more schooling than those who gained access later in life. Indonesian children enjoyed higher cognitive, math and schooling outcomes if the Safe Motherhood Program was in place over their first 1,000 days compared to children who benefitted after age two. Finally, a review of four evaluations of the Bucharest Early Intervention Program was inconclusive; age at placement influenced some cognitive development outcomes but not others.

Impact evaluations of six nutrition programs starting at various ages and lasting for various lengths of time—but always starting and ending during the first 1,000 days— demonstrated few later-life effects. However, a seventh program—providing supplemental feeding in Guatemala—demonstrated that continuous exposure from pregnancy through the first two to three years of age was more important than at three to six years of age and caused larger and significant results for highest grade completed, reading comprehension, and nonverbal cognition. Taken together, this suggests that effective interventions may need to not only start early but also continue throughout and perhaps beyond the first 1,000 days.

In the final dimension—length of exposure—much of the evidence comes from dose response estimates for cash transfer programs or from evaluations of preschool interventions. Given how important length of exposure is to determining the optimal timing of an intervention, very little evidence is available to guide policy makers on the effect of longer participation in any given intervention. What evidence does exist, however, highlights two important areas in which longer exposure times may be helpful in producing benefits. For preschool or childcare programs, it appears that longer exposure can lead to higher school enrollment rates, while additional involvement in a cash transfer program during the early childhood period could help reduce behavioral problems through adolescence.

Knowledge Gaps

More causal evidence is needed to fill gaps and corroborate findings.

International attention around early childhood development is fairly new. Much of the scientific evidence supporting the need for ECD comes from work within the United States and other developed countries that has recently been able to thoroughly explore the post-early childhood outcomes of interventions. Many low- and middle-income countries, confronted by different development challenges, have focused their efforts primarily on child survival and growth, subsequently limiting their ability to invest in interventions such as preschool and other stimulation programs.

Although this picture is slowly changing, very little is known about the effectiveness of ECD programs across the full range of outcome domains in developing countries, particularly in the post-early childhood period. While many early studies in developing countries, such as the INCAP supplementary feeding program in Guatemala and the maternal biscuit intervention in The Gambia, made important discoveries about the short-term effects of nutritional interventions during the first 1,000 days, the ability of researchers to assess the longer-term effects of these programs is limited by the design of the initial studies, which was not intended for follow-up.

At the time these evaluations were implemented, it was not yet widely understood that early childhood interventions could have sustained effects on ECD outcome domains, and it will be particularly important for future evaluations of ECD programs to facilitate long-term follow-up. The logistical difficulty of conducting long-term follow-up studies, particularly in the developing country context, has further contributed to a dearth of research.

One goal of this review is to further clarify the existing knowledge base of long-term effects of early childhood interventions and help inform future evaluations. Early childhood interventions can impact a variety of cognitive, linguistic, socioemotional, physical, educational, and employment outcomes; however, this review identified just 55 impact evaluations across all possible intervention types and outcome domains that passed the quality check. The evidence in several domains was particularly thin. For example, only one study measured the effects of ECD programs on employment outcomes, and only one intervention type—stimulation—had long-term effects measured in each of the six outcome domains. Future research should aim not only to provide more evidence across the full range of possible outcomes throughout an individual's lifespan, but also to expand the scope of interventions evaluated for their effects.

Challenges

Long-term follow-up evaluations face logistical challenges that contribute to the knowledge gap.

Evaluations aiming to estimate impacts of early childhood programs after a prolonged period face four challenges: confounding variables, attrition, designing for follow-up, and external validity. These challenges are not unique to evaluations of early childhood interventions or to longer-term evaluations, but they may be compounded here. In particular, issues of attrition and confounding variables are primarily responsible for the exclusion of more than half of the impact evaluation studies otherwise eligible for this review.

The evaluations that today constitute the evidence base of later-life effects of ECD interventions were often not designed with that objective. Many study designs were implemented prior to the existence of strong evidence of effects across a range of outcomes in the post-early childhood years and were not designed to track participants into adolescence and adulthood. Additionally, universal, standardized measurement across a range of outcomes over the lifespan—cognition and socioemotional development in particular—are not yet established, making it difficult to know how best to assess these constructs longitudinally. Although in some cases researchers are able to apply econometric methods based on analysis of past performance to tease out lasting effects that can be attributed to the original intervention, the absence of prior planning for long-term follow-up at the implementation stage has complicated causal inference from these studies.

Finally, evaluations of all types, including impact evaluations, have challenges of external validity—the ability to apply results found in one study to a different scale, context, or time. Most of the interventions evaluated here are somewhat small; scaling up to a national level may present administrative or other challenges. Furthermore, the longer-term nature of the interventions means that interventions included here occurred in an era—sometimes 30 years ago—that may have influenced interactions with the project in very different ways than would be expected in the contemporary context, even if in the exact same location. Potential variation across location underscores the need to fill in Regional gaps.

Implications

While much has been written on early childhood development and the near-term benefits to children selected to participate in interventions, few studies look at the sustained impacts of these programs. At this point in the development of the literature, this systematic review aims to contribute to the field's progress by collecting those studies that offer high-quality, causal estimates, providing analysis on interventions shown to have sustained effects across a range of outcomes, and identifying research gaps to help guide future studies.

Design Challenges

Evaluation design is critical in yielding valid causal estimates. Design is a major determining factor in how similar the comparison group is to the treated group, which in turn is the basis for calculating attributable program effects. Problems in design can be exacerbated over time as they interact with other factors, compromising the comparability of the two groups.

Two well-known interventions, the program of the Instituto de Nutrición de Centroamérica y Panamá (INCAP) in Guatemala and the Turkish Early Enrichment Project, suffered from this problem. Because of weaknesses in the initial randomization, many of the INCAP studies rely on a comparison group that is not statistically equivalent to the treated group, while in Turkey, group comparability suffered from selection bias in one wing of the study and extremely high attrition rates overall. Such design challenges can undermine causal inference.

The results of this outcome-based systematic review imply that some domains may be easier to affect than others. In cognitive, linguistic, socioemotional, and employment domains, the evidence suggests that early stimulation interventions can result in sustained benefits to children, and various interventions were successful in improving subsequent schooling outcomes. Conversely, there was little evidence of a strong post-early childhood effect in physical outcomes across the range of evaluated interventions.

Despite these promising results, much work remains to be able to draw firm conclusions on the post-early childhood effects of ECD interventions. With 55 impact evaluation studies on 25 projects of 20 intervention types across 22 countries, the results presented are still indicative rather than conclusive. High-quality evaluations of interventions that could significantly impact a child's development, including nontraditional interventions such as clean water and sanitation or agriculture, are hard to find but are necessary to determine the most effective method of intervening. Furthermore, evaluations across Regions are important to capture context-specific variables and improve general external validity. The analysis of heterogeneous effects, especially by gender and socioeconomic status, can improve intervention targeting.

Early childhood interventions, like the children they serve, have transformative potential if properly supported. As the old proverb says, the best time to plant a tree is 20 years ago, but the second best time is today. That applies to investments in children as well as it applies to investments in evaluations that can track their progress.

Introduction: Review Questions and Strategy

While the later-life effects of interventions during the early childhood period are generally well documented in the developed world, far less evidence exists in developing countries. To address this knowledge gap, the review uses a comprehensive search strategy and a rigorous screening method to compile the causally robust evidence on the post-early childhood effects of early childhood interventions in developing countries. The evidence is organized around six outcome domains to determine which intervention types can effectively and consistently produce sustained effects in particular areas of child development.

In the *World Development Report 2006: Equity and Development,* the World Bank singled out early childhood development (ECD) interventions as a promising policy area to achieve both equity and efficiency objectives (World Bank 2005). More recently, in the *World Development Report 2015: Mind, Society, and Behavior,* the Bank again chose child development as a key facet of inequality, noting that children from developing nations have systematically lower socioemotional and cognitive stimulation in their early years, which together with the high stress of growing up poor can impair decision-making and deliberative abilities (World Bank Group 2015).

These flagship reports perceive that ECD can play a key role in achieving the Bank's twin goals to reduce extreme poverty and promote shared prosperity while encouraging economic growth. The realization of these aims is predicated on the ability of countries to "build human capital and increase long term productivity" through "access to early childhood development, health, nutrition, education, and basic infrastructure," which enhances "mobility on the economic and social ladder within and across generations" (World Bank 2013, p. 28). Programs targeted at ECD do just that—build human capital—by intervening during a critical period of development when it is suggested that interventions can improve both the starting point and trajectory of a child's life path and provide a longer time horizon over which the benefit stream is realized (Carneiro and Heckman 2003).

Early childhood development is an integrated construct influenced by many factors, such as nutrition, health, hygiene, early learning, and stimulation. For example, good nutrition during the first 1,000 days, from conception to the child's second birthday, is essential for normative linear growth (Victora and others 2008; Black and others 2013) and health brain formation (Couperus and Nelson 2006; Georgieff and Rao 1999). The plasticity of the young brain (that is, its capacity to change) allows young children to benefit from positive inputs such as stimulation and nutrition but also makes them vulnerable to negative external shocks including deprivation and abuse (Shonkoff and others 2012a; Fox and others 2010). Cognitive, language, fine motor, and socioemotional skills important for educational and social success emerge during these years (Heckman 2008; NRC and IOM 2000; Shonkoff 2011). Risk factors related to poverty (for example, undernutrition, poor sanitation, insensitive parenting) in early childhood are associated with delays in these skills as well as in school progress (Georgieff 2007; Grantham-McGregor and others 2007; Walker and others 2007; Glewwe, Jacoby, and King 2001).

Experimental evidence suggests that nutrition, health, early learning, and other factors can play a major role in shaping young children's subsequent school attainment, performance, and earnings (Heckman 2008; Naudeau and others 2011; Barnett 2011; Duncan and others 2007). Indeed, the benefits of a variety of early childhood interventions are well documented in developed countries.[1] Long-term evaluations of children who received these interventions in the United States found positive life outcomes in education, health, fertility, risky behaviors, and the labor market (Smith 2009; Cunha and Heckman 2009; Schweinhart 2007; Campbell and others 2002; Camilli and others 2010; Reynolds and others 2007; Anderson and others 2003; Bitler, Hoynes, and Domina 2014; Olds, Sadler, and Kitzman 2007; Sweet and Appelbaum 2004). This evidence led many in the international development community—including the World Bank—to promote ECD and to focus on interventions as the first step in a sequenced strategy to build the skills needed for productivity and economic growth (World Bank 2010).

In developing countries, numerous studies and reviews establish that early childhood interventions can improve early childhood outcomes (Maternal and Child Nutrition Study Group 2013). Yet the ability of ECD interventions to improve later outcomes—those occurring after the early childhood period[2]—is less well studied. Impact evaluations (IEs) that examine the post-early childhood effects from early childhood interventions in middle- and low-income countries are scarce, and it may be problematic to extend the findings from the United States to developing countries. While the challenges of developing countries differ from place to place, on average they face challenges that Organisation for Economic Co-operation and Development countries do not, such as weaker institutions and lower economic development. It follows that interventions that are effective in wealthy nations may not have the same results in the resource-constrained contexts of low- and middle-income countries, especially given that basic needs in these countries are often less well met.

The converse may also be true—interventions that are effective in developing countries may not yield the same results in wealthier contexts where basic needs are met.[3] This dynamic may explain differences in the set of evaluated interventions by national income.[4]Nevertheless, evidence from developed countries can help establish the physiological pathways between particular inputs and ECD outcomes. However, as was discovered in IEG's systematic review of maternal and child mortality (IEG 2013), even where what to do is known (and even this is not always the case with ECD), knowing how to do it under such different circumstances is a persistent challenge.

This review by the Independent Evaluation Group (IEG) tries to addresses that challenge by gathering and analyzing the available impact evaluation evidence on the post-early childhood effects of interventions conducted during the early childhood years in the developing world. The purpose of doing so is not to supplant existing evidence, but rather to help practitioners understand how evidence from IEs supports or challenges beliefs about ECD interventions and how this evidence can be used to inform development policy. Impact evaluations aim to overcome confounding factors inherent in other forms of evaluation to determine the causal impact of an intervention. This highly rigorous form of evaluation is particularly well suited to evaluate the claims of lasting effects from early childhood interventions, considering the number and scope of confounding factors that arise over time.

By taking such a specific approach—focusing exclusively on the post-early childhood effects of any ECD intervention—this review goes beyond the existing work done in this field, an anthology of which is found in appendix E.[5] For instance, a 2011 review published in *The Lancet* series on ECD in low- and middle-income countries included some evidence on post-early childhood effects, but these outcomes were a relatively minor part of the review (Engle and others 2011).

The present report differs from *The Lancet* piece in the breadth of the interventions and outcomes included, the specificity of the age at evaluation, and the types of studies accepted for inclusion. While Engle and others use an intervention-based approach to focus on psychosocial and educational interventions and the resulting range of outcomes, this review includes IEs of any type of intervention occurring during the early childhood period and presents them through an outcome-based framework (Waddington and others 2012). Additionally, unlike *The Lancet* review, which included outcomes measured during the early childhood period for both children and parents, this review only reports outcomes in the post-early childhood period for children. Finally, while *The Lancet* includes both efficacy[6] and effectiveness studies implemented using local or foreign capacity, this review is restricted to those interventions that use local capacity so as to provide evidence on interventions that could actually be replicated by low- and middle-income countries.

A systematic review published recently in the *Annals of the New York Academy of Sciences* examined the effect of integrated interventions (Grantham-McGregor and others 2014) in low- and middle-income countries, but again its scope differs in significant ways from this review. As with *The Lancet* review, Grantham-McGregor and others focused on a narrow group of interventions (stimulation and nutrition), included a broader range of outcomes and ages at evaluation (both parent and child outcomes measured during the early childhood period and after), and drew evidence from interventions implemented using local or foreign capacity. It should also be noted that their use of evidence from the post-early childhood period was limited.

Despite having received relatively little attention, post-early childhood benefits are quite important as they comprise a major economic argument for investing in young children based on the assumption that returns continue over time. Budget and time constraints, together with the estimation challenges inherent in follow-up IEs, may be responsible for the relatively few IEs that address later-stage outcomes of early childhood interventions. However, it is important to take stock of the causal evidence that does exist. With that in mind, this review uses evidence from all impact evaluations with credible causal estimates for interventions occurring before primary school age on outcomes occurring at or after primary school enrollment. It aims to answer the following questions:

- What is the evidence of attributable effects on outcomes in primary school and beyond from interventions in low- and middle-income countries that occur during the early childhood period?
- How do the post-early childhood effects of early childhood interventions vary by socioeconomic status, gender, age at intervention, and age at evaluation, particularly during the first 1,000 days from conception to the child's second birthday and from age three to primary school enrollment at age five to six?

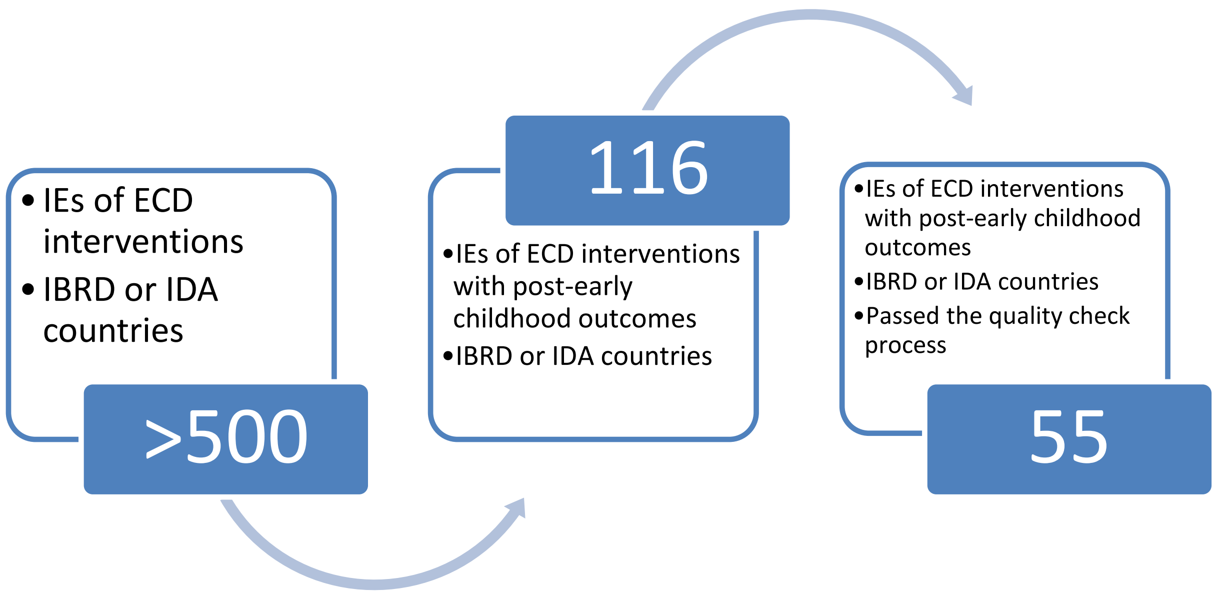
For the purposes of this review, the defined intervention period for early childhood runs from conception to primary school enrollment of the child. Accordingly, all early childhood interventions are considered that are either directly provided to the child or to the parent, caregiver, or pregnant or lactating mother. This includes a range of interventions such as preschool, conditional cash transfers, and behavior change interventions such as early stimulation by caregivers, promotion of exclusive breastfeeding, complementary feeding practices, and health and hygiene practices.

Search Strategy for Identification of Relevant Studies

Studies were identified through a detailed search strategy based on an approach used by a previous systematic review in IEG's Maternal and Child Health series (IEG 2013). Using search terms that encompassed the outcomes, methods, and the definition of low- and middle-income countries employed by the International Bank for Reconstruction and Development (IBRD) and the International Development Association (IDA), studies between 1990 and 2014 were collected from relevant bibliographic databases and the World Bank's impact evaluation database. This was supplemented by searches of relevant institutions and a hand search of top ECD authors, followed by the snowballing of references from included studies and related systematic reviews identified during the electronic search. The search included both published impact evaluations and unpublished grey literature, defined as working papers and studies soliciting review from the research community, in so far as they were identifiable by the established search procedure. (See appendix C for more information on the search protocol.)

Initial searches produced more than 500 potentially-relevant evaluations of ECD interventions in IBRD or IDA countries; 116 provided estimates for post-early childhood outcomes. A rigorous quality check against the inclusion criteria yielded 55 relevant studies of sufficient quality to include in the analysis (figure 1). These studies represent 20 intervention types (see figure 2) across 25 projects in 22 low- and middle-income countries.

Figure 1. Flow of Search Results



Note: ECD = early childhood development; IBRD = International Bank for Reconstruction and Development; IDA = International Development Association; IEs = impact evaluations.

Despite this relatively thin evidence base, these 116 studies constitute the complete body of IE knowledge on the post-early childhood effects of ECD programs. The number is expected to grow. It should be remembered that the field of early childhood development is still relatively new, especially for developing country contexts. The pioneering IEs that offer evidence of later-life effects of early childhood development interventions were designed decades ago. Considerable evolution of both ECD theory and empirical standards and design has occurred since the initiation of some of the earliest studies reviewed (for example, the supplementary feeding program of the Instituto de Nutrición de Centroamérica y Panamá was begun in the 1970s), and most were not designed to serve as platforms for understanding long-term effects. This is not to critique the garden of child development evidence for being newly sown. Rather, in light of this dynamic evidence base, this review seeks to take stock of what is known and point out where further cultivation might bear fruit.

Box 1. Analyzed Outcomes: The Rule of Three

This review reports outcomes across six different domains: physical development, cognitive development, language development, socioemotional development, schooling outcomes, and employment and labor market outcomes. Outcomes were selected for inclusion based on how often they were measured and reported and the extent to which they provided unique insight into measurements within the domain.

Outcomes that were reported in three or more independent studies were included. The high frequency of these outcomes was indicative of their usefulness in measuring and understanding early childhood development, and enabled comparison across contexts and intervention types.

In some cases, however, an outcome may have been reported in more than three studies, but was not included due to high overlap with other indicators. For example, both fat mass index and body mass index (BMI) were reported in many studies, but this review only included BMI measurements. Both indicators capture a similar measurement of a child's physical development, and including both would have provided little unique insight.

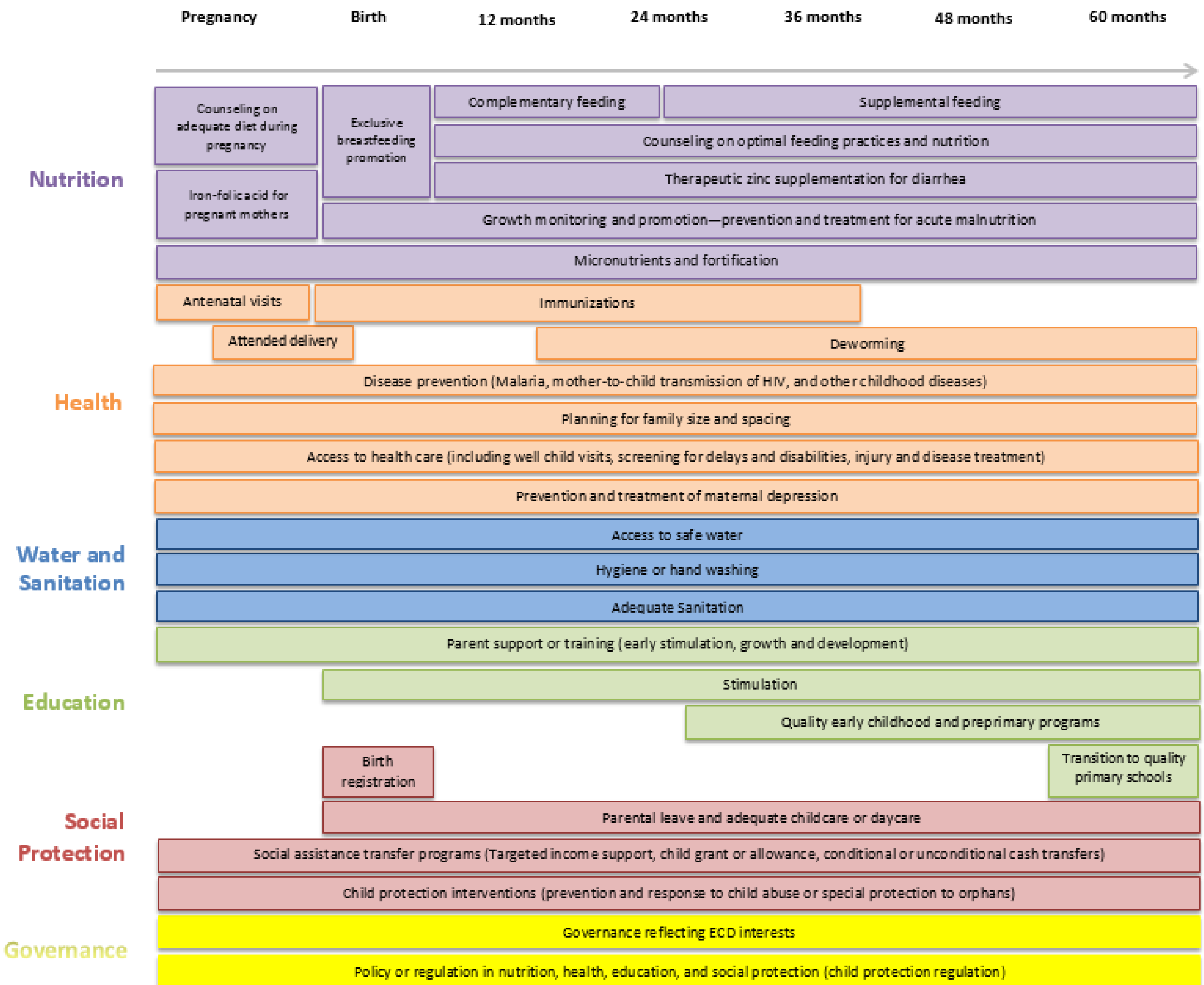
An exception to the "rule of three" was made for outcomes in employment and fine motor skills, both of which have high intrinsic value. Fine motor skills provide a useful measure of the development of an individual's executive function capabilities—an outcome domain in which little post-early childhood evidence exists. Similarly, given the focus on long-term outcomes in this review, employment outcomes are analyzed even though they are infrequently reported.

Criteria for Inclusion and Exclusion of Studies

The inclusion criteria required that the studies (i) employ a quantitative impact evaluation methodology using experimental or quasi-experimental identification strategies with a credible counterfactual[7] to identify causal attribution; (ii) evaluate any intervention occurring in a low or middle-income country; [8](iii) limit the population of analysis to children exposed to the intervention during the early childhood period, defined here as conception to the age of entry into primary school (or age six, when information on primary school age is not available); [9] (iv) assess post-early childhood outcomes occurring from primary school enrollment through adulthood; (v) be published between 1990 and the present (February 2015) and use end-line data occurring no earlier than 1990; and (vi) were subject to some form of peer review. To reduce the risk of publication bias, grey literature was included for consideration if a full-text version is publicly available and the study passes other quality and inclusion criteria. These inclusion criteria were selected to best answer the evaluation questions above. These criteria limited the body of admissible evidence; different evaluation questions may result in the inclusion of otherwise valid studies not incorporated into this review.[10]

A taxonomy of ECD interventions is presented in figure 2, categorized by intervention type and the ideal age at intervention. To be included in this review, a study could be on any type of intervention, whether a typical ECD intervention or not, as long as it was implemented through local capacity channels (for example, governments, nongovernmental organizations, or private sector firms) and would therefore be more easily replicated by local implementers.

Figure 2. Essential Interventions for Young Children and Families



Source: Denboba and others (2014), adapted by IEG.

All included studies had a sample size of 50 or more and an attrition rate lower than 40 percent.[11] The unit of analysis could be at the individual, household, facility (for example, school, center, or clinic), or community level.[12] Studies using aggregated national or regional data, as in cross-country or national interrupted time series analyses with few observations over multiple periods are excluded from this review.

All studies meeting the inclusion criteria were subjected to a double-coded quality review that assessed the strength of the internal, external, and construct validity. The primary criteria for this assessment was the extent to which the identifying assumptions of the identification strategy used were met, whether the data were representative of a definable and policy-relevant population, and whether the key indicators and outcomes in the report were measured in an unbiased and reliable way. (See appendix D for information about the rating process and coding protocol.)

Based on these criteria, studies (or in some cases specific estimation strategies within a study) were assigned an overall quality rating: A, AA, or AAA. As in past IEG systematic reviews, only AA and AAA studies were included in the review. Studies with an AA rating have a credible counterfactual with identification assumptions plausibly fulfilled; those with an AAA rating have a credible counterfactual with identification assumptions clearly fulfilled. Studies graded A leave significant doubts about the validity of the counterfactual and the likelihood that its identifying assumptions have been violated; for completeness these A-quality studies are listed in appendix F.

This process identified 55 sufficiently high quality impact evaluations used for synthesis in this review. Independent of the rigorous standards applied in the search and coding process, challenges inherent to systematic reviews remain (see box 2).

Relatively few impact evaluations can answer the important questions about long-term effects of early childhood interventions posed by this review. Moreover, because the evidence that is available is spread across a broad range of intervention and outcome combinations, the evidence for any particular combination can be quite thin, or completely nonexistent. Thus, although the review contains all causally robust evaluations, results should be viewed as comprehensive but nascent and prone to change as new results are published.

In light of the challenges endemic to reviews (see box 2), rather than assuming that evidence gaps indicate interventions that do not work, this report focuses on interventions shown to be effective. One test of robustness of a review is whether the main messages would change if the cutoff date for inclusion were altered. As explained in appendix C, this review completed an initial search and analysis ending in 2013 and then refreshed that search prior to publication to include material released in 2014. Although there was a substantial uptick in the number of studies from which to draw evidence, the conclusions remained almost entirely unchanged. This finding supports the view that the conclusions included here are likely to be broadly stable as still more evaluative, causal evidence is produced.

Box 2. Challenges and Cautions for Systematic Reviews

Notwithstanding the thoroughness of the search strategy, challenges remain to representativeness and interpretation of results that are common to all systematic reviews. This review meets or exceeds standard practice, such as it exists, in every instance.

Challenges to representativeness of the interventions arise from the fact that the process of selecting interventions to be evaluated by an impact evaluation is purposeful rather than random.

- Some types of projects are less amenable to impact evaluation methods and will be underrepresented.
- Interventions that report on intermediate rather than final outcomes are excluded.
- Importantly, the lack of existing impact evaluations for a family of interventions indicates a need for evaluations in that area, not that the interventions are ineffective.

Challenges to representativeness of the impact evaluations are twofold.
- The review includes only concluded studies; it cannot use impact evaluations that are planned or in process.
- As with all reviews, the sample may suffer from file drawer bias or publication bias wherein studies that yield null results are not completed. Alternately, it has been hypothesized that only studies with experimental designs can be published with statistically insignificant results because of stronger internal validity; this may lead to a false conclusion that randomized trials are more likely to return null results.

Challenges to interpretation of results imply a need for thoughtful application of findings.
- Impact evaluations of projects funded by foreign aid likely underestimate the true effect of the intervention because they measure partial (or local) equilibrium effects rather than the general equilibrium effects resulting from the fungibility of government budgets, which allows countries to reallocate health funding away from of the foreign-funded activities (Wagstaff 2011).
- Null results must be interpreted carefully: they do not necessarily mean there is no effect. They may occur where there is measurement error, insufficient sample size (power) to detect an effect, spillover from treatment to the control group, differential attrition, insufficient behavioral incentives, or implementation challenges. Distinguishing the causes of a null result is often untenable.
- External validity is a persistent challenge. Applicability of results to a different context—time, place, or scale—is likely a function of project complexity (Woolcock 2013), administrative capacity, political supportability, and alignment with the most important barriers of the target environment. The ideas and processes of an intervention may have greater external validity than the intervention itself.

Consensus analysis was carried out across the studies comparing statistical significance and magnitude in forest plots. Where sufficient numbers of studies exist with outcomes in a common construct, meta-analysis was conducted to explore the presence of an overall effect.

The report presents evidence for post-early childhood outcomes of early childhood-age interventions over six important domains: physical, cognitive, language, socioemotional, schooling, and employment and labor market outcomes. Chapter 7 breaks down heterogeneous effects and the differential outcomes observed both between subgroups and over time. Chapter 8 discusses the challenges of evaluating long-term effects of interventions conducted during the early childhood years, and chapter 10 identifies the remaining gaps in knowledge about the post-early childhood effects of ECD interventions.

---

[1] In developing countries, preschool participation contributed to an increase in lifetime earnings by 5–10 percent (Engle and others 2007, 2011). Belfied and others (2000) computed the benefits of the Perry Preschool Program in the United States to be $150,000 (in 2000 dollars) per child through age 40 because of crime reduction.

[2] The post-early childhood period is determined to start on enrollment in primary school or, when that information is unavailable, at age six years.

[3] The effectiveness of a particular intervention is a function of, among other things, the complementary slackness of the constraint in the outcome's production function that that intervention is designed to address. For an example of a production function model of a range of inputs on a range of children's outcomes, see Tanner 2012 and Becker 1993.

[4] This presents a mechanism for the endogeneity of the evidence base by national income. For example, in contexts where a particular input (say nutrition) has a lower shadow price on the production of a particular output, the effect size of an intervention designed to alleviate that (nutritional) constraint will be low. Thus there may be comparatively fewer interventions of that input in that context, and so there may be fewer evaluations of that (nutrition) input.

[5] For a full review of the existing systematic review literature, see IEG (2015).

[6] The World Bank's handbook *Impact Evaluation in Practice* (World Bank 2011) defines efficacy trials as having "heavy technical involvement from the researchers during the implementation of the program" and which do not use "regular implementation channels." For the purposes of the inclusion criteria of this review, this is codified such that studies are excluded which do not rely on local capacity for implementation of the intervention.

[7] Constructing a valid comparison group for use as a counterfactual—the outcome that would have been observed for a participant in the absence of the program—is the defining feature of impact evaluation methods.

[8] Despite being recently classified as high-income countries by the World Bank in July 2013, studies from both Chile and Uruguay are both included in this review due to the fact that end-line data for those studies was collected during the period when they were still categorized as middle-income countries. All countries included in this report are currently World Bank client countries.

[9] The early childhood period is defined as the time between conception and a child's entry to primary school. If age at the time of entry to primary school was not mentioned, 59 months was used as a cut off. Therefore, these impact evaluations must include outcomes measured after

entry to primary school or age six years (post-early childhood) resulting from interventions that occurred before entry to primary school or age six years.

[10] For example, impact evaluations of outcomes measured within the early childhood period are excluded, vacating a fairly substantial literature of nutrition and feeding interventions that demonstrate health benefits, sometimes even a year or more after the end of the intervention but still before primary school.

Similarly, in order to more cleanly demarcate and identify early childhood interventions and post-early childhood effects, evaluations whose intervention or evaluation ages straddle the age cutoff of primary school (or 59 months) are excluded. For example, follow-up studies of a deworming campaign in Uganda that included some seven-year olds and potentially students in the early primary grades are not considered, though they do show robust village-level effects (see Alderman and others 2006; Croke 2014)

[11] In the professional judgment of the team, sample sizes below 50 engender serious doubts pertaining to their external validity; lower sample sizes are also more likely to be under-powered, complicating the interpretation of null results. There was one exception: a single outcome estimate from one of the studies of the Jamaica supplementation and stimulation project had a sample size of 48. Similarly, attrition rates over 40 percent (many would say 20 percent or even less) are judged to be seriously challenged by selection bias. See Chapter 8 for a more detailed discussion on challenges to follow-up impact evaluations, including problems of attrition.

[12] For example, if all students at a given primary school had participated in the intervention, attendance or school performance could be measured at the school level. In practice, almost all of the outcomes were measured on the individual level.

Part I: Effects by Outcome

The effects of interventions on six domains—physical development, cognitive development, language development, socioemotional development, schooling outcomes, and employment and labor market outcomes—are commonly included in evaluations of early interventions targeting poor children because they are negatively affected by early poverty, are believed to benefit from early intervention, and contribute to well-being and adult productivity. Within each domain, outcomes measuring similar constructs were analyzed to simplify analysis and presentation. Evidence comes from 55 impact evaluations of 25 projects of 20 interventions types conducted in 22 countries.

Each chapter in Part I begins with a table of the outcomes reviewed in that chapter. The studies are organized by intervention type and then by age at evaluation, giving readers the opportunity to scan the data on each type of intervention within a domain and underscoring the emphasis of temporal trajectories this review. These tables of outcomes do not include all that are reported in the studies. Instead outcomes are included that have a sufficient number of comparable measures to draw conclusions. In some cases, outcomes with few measures were included because of their unique contributions. For example, the three studies that report on employment outcomes represent the entire body of evidence on post-early childhood employment effects from early childhood interventions.

Throughout the report, boxes are used to highlight well-known interventions or interventions for which there is a series of evaluations at progressing ages on the same subjects. These vignettes give more detailed information on the intervention itself as well as the outcomes across time and outcome domains. All of the interventions analyzed in this report are included in appendix A, which contains a brief description of the intervention, the relevant counterfactual, age at initiation, and age at evaluation. Finally, the numbered references in [brackets] used throughout this report correspond to the numbered list of studies that passed the inclusion criteria for quality, found in the References section.

Chapter 1: Physical Development—*There is no evidence of later-life effects on weight or mid-upper arm circumference, little evidence of an effect on head circumference, and the intervention types that appear to affect height do so inconsistently. The evidence on fine motor skill is too thin to draw firm conclusions.*

Chapter 2: Cognitive Development—*Cognitive development is improved by a range of interventions, and most improvements are seen in measures of general cognition. Nutritional programs had almost no effect.*

Chapter 3: Language Development—*The evidence suggests that early childhood development interventions can have lasting effects on language, although the results were mixed within intervention types and within outcome subcategories.*

Chapter 4: Socioemotional Development—*Improvements in externalizing behavior may multiply as participants age; conditional cash transfers and early stimulation show fairly consistent benefits. It appears harder to create a sustained change in internalizing behavior.*

Chapter 5: Schooling Outcomes—*Early stimulation, preschool, and conditional cash transfers seem most effective in improving schooling outcomes.*

Chapter 6: Employment and Labor Market Outcomes—*Early stimulation can improve labor outcomes, but the evidence base for these outcomes is particularly thin.*

1. Physical Development

There is little evidence of later-life effects on weight, mid-upper-arm circumference, or head circumference, and the intervention types that appear to affect height do so inconsistently. The evidence surrounding fine motor skill is too thin to draw conclusions.

The physical health and nutritional status of young children are important determinants of future well-being and development. Children who do not grow well in the first few years or who experience early illnesses or stressful environments may be at risk for increased likelihood of chronic conditions that can impede normal, healthy functioning (Irwin and others 2007; Shonkoff, Boyce, and McEwen 2009). For malnutrition in particular, and its associated effect on weight and height, the first 1,000 days are key, with some experts positing that stunting occurring during that period is irreversible (Black and others 2008; Lake 2012; Shrimpton and others 2001; UNICEF 2013). Long-term consequences can also extend beyond a person's health. A systematic review of observational studies from low- and middle-income countries found that undernourished children grew into shorter adults and had less schooling and lower economic productivity (Victora and others 2008). Other reviews find that early childhood undernutrition can cause long-term cognitive deficits (Grantham-McGregor and Baker-Henningham 2005; Grantham-McGregor 1995; Mendez and Adair 1999).

Although undernutrition has long commanded the focus of maternal and child health experts, overweight and obesity have increasingly become problems in developing countries (Ebbeling, Pawlak, and Ludwig 2002; Prentice 2006; Bhutta and others 2013; Black and others 2013). The prevalence of both has risen among children under five years old and has been linked to immediate health issues such as high blood pressure and cholesterol as well as type-2 diabetes in adolescence and later-life chronic diseases (Ebbeling, Pawlak, and Ludwig 2002; Black and others 2013). As with growth restrictions, the first 1,000 days is very important for preventing excessive weight gain. Rapid weight gain during that time period is associated with adult lean mass, while weight gain in later childhood is more likely to lead to adult fat mass (Black and others 2013).

Recent multidisciplinary efforts underway that combine the talents and skills of pediatricians, geneticists, and neuroscientists are testing the capacity for early interventions to improve a child's current health status and long-term risk for disease (Shonkoff and others 2012b). There is some evidence of this already. For example, former participants of the Carolina Abecedarian Project in the United States, now in their mid-30s, show significantly lower risk factors for cardiovascular and metabolic diseases (Campbell and others 2014). These health benefits can begin to accrue immediately. There is evidence that Head Start, a U.S. preschool program, had improved the overall health status of three and four year olds when they reached kindergarten (HHS 2010), although physical outcomes were not the primary indicator of interest.

In low- and middle-income countries, nutritional supplementation led to better growth among children under five as did various conditional cash transfer programs (Grantham-McGregor and others 2014; Bhutta and others 2008). However, only a few of the interventions were again examined the post-early childhood period, when some of these benefits disappeared soon after the end of the intervention while others persisted (Grantham-McGregor and others 2014). These mixed results highlight the need for a more systematic review of the evidence concerning the sustained effect of early childhood development (ECD) interventions on physical outcomes.

Table 1.1 maps the 19 studies across 12 countries in this review that investigate physical outcomes. There are 13 different projects and 11 intervention types. Results are measured for six specific indicators: body mass index (BMI), BMI z-score (BMIZ), head circumference, height-for-age z-score (HAZ), mid-upper-arm circumference (MUAC), weight-for-age z-score (WAZ), and weight-for-height z-score (WHZ).

Across the impact evaluations that included physical outcomes, authors consistently focused on commonly used anthropometric measures: height, weight, MUAC, and head circumference. All are proxies of a child's nutritional status, although the final one is less relevant to later-life outcomes as head circumference usually stabilizes by five years old (Figueiras and others 2012) and is often not measured after two to three years old. Furthermore, as these outcomes are indicators of nutritional status and, more broadly, the functional correlates of physical health such as cognition, schooling and employment, their value is more instrumental than intrinsic. For example, decreased height is not in itself necessarily a problem, but it is nonetheless a valuable measure as it can signal an increased risk of negative outcomes such as morbidity, mortality, and impaired cognitive development (Black and others 2013).

Some studies used as outcomes the group differences in the actual value measured (for example, PROBIT in Belarus[24, 33] and maternal supplementation study in The Gambia[1, 20] looked at height differences in centimeters), while others used standardized (z) scores based on a reference population to determine intervention effects. The use of z-scores is often preferable as it allows for easy comparison across ages and populations. Since their publication in 2006, most studies have used the growth standards of the World Health Organization (WHO 2006). Many of these reviewed studies, which gathered data prior to 2006, used the National Center for Health Statistics growth standards (Hamill and others 1979).

Table 1.1. Impact Evaluations Investigating Physical Development

	Study	Country (Project)	Average Age at Intervention (Years)	Average Length of Exposure (Years)[c]	Age at Evaluation (Years)	Evaluated Intervention	Reviewed Outcomes
Nutrition	Kramer and others 2007a [24]	Belarus (Promotion of Breastfeeding Intervention Trial [PROBIT])	0	1	6	breastfeeding promotion	BMI; head circumference*; height; MUAC
	Martin and others 2013 [33]	Belarus (Promotion of Breastfeeding Intervention Trial [PROBIT])	0	1	11	breastfeeding promotion	BMI; head circumference; height; MUAC
	Stewart and others 2009a [45]	Nepal (maternal nutritional supplementation)	in utero	0.75	6–8	vitamins, micronutrients, or fortified food for pregnant women (folic acid, iron, and zinc)	height (folic acid+iron+zinc)*; height (folic acid, folic acid+iron, multiple micronutrient)
	Stewart and others 2009b [46]	Nepal (maternal nutritional supplementation)	in utero	0.75	6–8	vitamins, micronutrients, or fortified food for pregnant women (folic acid, iron, and zinc)	BMI (folic acid, folic acid+iron, folic acid+iron+zinc, multiple micronutrient)
	Devacumar and others 2014 [15]	Nepal (maternal multivitamin supplementation)	in utero	0.17	8.5	vitamins, micronutrients, or fortified food for pregnant women (multivitamin supplement)	BMIZ; HAZ; head circumference; height; MUAC; WAZ; weight (z-scores based on WHO standards)
	Hawkesworth and others 2008 [20]	Gambia, The (maternal supplementation)	in utero	0.5 (DR1)	11–17	vitamins, micronutrients, or fortified food for pregnant women (protein biscuits)	BMI; height***; weight
	Hawkesworth and others 2011 [22]	Gambia, The (maternal supplementation)	in utero	0.5 (DR1)	11–17	vitamins, micronutrients, or fortified food for pregnant women (protein biscuits)	BMI
	Alderman and others 2014 [1]	Gambia, The (maternal supplementation)	in utero	0.5 (DR1)	16–22	vitamins, micronutrients, or fortified food for pregnant women (protein biscuits)	height
	Walker and others 1996 [50][a]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	7–8	supplementary feeding	HAZ; WAZ (z-scores based on NCHS standards)
	Walker and others 2000 [51][a]	Jamaica	1.55	2	11–12	supplementary feeding	BMI; HAZ (z-scores based on NCHS standards)

		(stimulation and supplementation to stunted children)					
	Pongcharoen 2010 [38]	Thailand (micronutrient supplementation to children)	0.5	0.5	9	micronutrients and fortified food for children (iron and/or zinc supplementation)	BMIZ; HAZ*; MUAC; WAZ (z-scores based on WHO standards)
Early Learning/Childcare	Martínez, Naudeau, and Pereira 2012 [35]	Mozambique (preschool)	3.45	1.5	5–9	quality early childhood and preprimary program	fine motor skills*; HAZ; WAZ (no information given on reference population used for calculating the z-score)
	Walker and others 1996 [50][a]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	7–8	stimulation	HAZ; WAZ (z-scores based on NCHS standards)
	Levin and others 2014 [29]	Romania (Bucharest Early Intervention Project)	1.88	2.7	8	stimulation (foster care)	motor skills
	Walker and others 2000 [51][a]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	11–12	stimulation	BMI; HAZ (z-scores based on NCHS standards)
Health	Barham 2012 [5][b]	Bangladesh (Matlab)	NA	continuous	8–14	well child visits	HAZ** (normalized using comparison areas means and standard deviation)
	Ozier 2013 [36]	Kenya (primary school deworming project)	0	1	8–15	deworming	HAZ; height (z-scores based on WHO standards)
	Barham 2012 [5][b]	Bangladesh (Matlab)	NA	continuous	15–19	family planning	HAZ (normalized using comparison areas means and standard deviation)
Social Protection	Behrman and others 2008 [8]	Mexico (Progresa)	1.5	1.5 (DR2)	7–11	CCT	BMIZ; height (z-scores based on WHO standards)
	Manley, Fernald, and Gertler 2012 [32]	Mexico (Progresa)	1	1.5 (DR2)	8–10	CCT—conditionalities	BMIZ; HAZ (z-scores based on WHO standards)
						CCT—cash	BMIZ; HAZ*** (z-scores based on WHO standards)
	Barham and others 2014 [6]	Nicaragua	in utero	3 (DR2)	10 (boys)	CCT	HAZ; WAZ (no information given on reference population

		(Red de Protección Social)					used for calculating the z-score)
	DSD, SASSA, and UNICEF 2012 [14]	South Africa (Child Support Grant)	1	2.5	10	unconditional or targeted income support	HAZ (no information given on reference population used for calculating the z-score)

Note: Bracketed numbers correspond to numbered studies in References. More details for each study are found in appendix A. BMI = body mass index; BMIZ = body mass index z-score; CCT = conditional cash transfer; DR = dose response; DSD = Department of Social Development; HAZ = Height-for-age z-score; MUAC = mid-upper arm circumference; NA = not applicable; NCHS = National Center for Health Statistics; SASSA = South African Social Security Agency; UNICEF = United Nations Children's Fund; WAZ = weight-for-age z-score; WHZ = weight-for-height z-score; WHO = World Health Organization.

a. Jamaica [50, 51] studies have a multiple intervention arm, and each intervention type has a separate row for these studies.

b. Bangladesh Matlab study [5] has "NA" on intervention age and length of exposure because of the nature of the family planning intervention.

c. DR in the length of exposure means the intervention period in terms of the dose response. DR is either randomized rotation (DR1) or phase-in (DR2). In terms of dose response, for instance, DR1 indicates that The Gambia study where treatment group receives protein biscuit only in utero whereas control group receives it only in postpartum, and length of exposure is the length of intervention for treatment. DR2 describes the dose response where early and late treatment effect is compared, and length of exposure is the difference of the intervention period between treatment and control group. "Continuous" indicates that the program effect could continue over time.

* Statistically significant at 10 percent.
** Statistically significant at 5 percent.
*** Statistically significant at 1 percent.

Box 1.1. Breastfeeding Promotion in Belarus Has Few Lasting Physical Benefits

Many ECD programs emphasize breastfeeding. However, because most of the evidence on breastfeeding benefits is based on studies that may have lacked adequate control of biases or confounding factors, some questions remain about the impact of breastfeeding on various medium to long-term child health outcomes (Kramer 2010). For instance, mothers who choose to breastfeed exclusively or for longer may differ in other ways in which they care for their children. One program carried out in Belarus was designed to examine the effects of breastfeeding promotion on children's well-being through cluster randomization. Between 1996 and 1997, 31 maternal hospitals and clinics were chosen as locations for the Promotion of Breastfeeding Intervention Trial (PROBIT), a 12-month intervention promoting breastfeeding among healthy mothers and full-term infants (Kramer and others 2001). The program involved community health workers facilitating breastfeeding and ongoing lactation, and providing postnatal support in the recipient clinics. Hospitals and clinics not assigned to the experimental intervention continued normal postnatal care practices. Unlike many other studies that problematically compare breastfed-only children to formula-fed (only) children, all mothers in both groups intended to breastfeed their children for at least some time (Kramer and others 2001).

The PROBIT program had several immediate impacts on breastfeeding behavior and child health outcomes (Kramer and others 2001). Mothers in the promotional program were more likely to be breastfeeding their infant at 3, 6, 9, and 12 months, compared with mothers attending health facilities following normal practices. These mothers were also more likely to exclusively breastfeed their babies in the first six months of age. Infants of participating mothers experienced less gastrointestinal infections and episodes of atopic eczema, although no significant difference was detected in the prevalence of respiratory tract infections.

Subsequent studies of the program evaluated the longer-term effects on physical health, nutritional status, socioemotional development and cognitive abilities, assessing children when they turned 6.5 years old. Contrary to results suggested from less rigorously designed observational studies, these impact evaluations found virtually no group differences for most outcomes. Children in the breastfeeding-promotion group did not have a significantly different risk of allergic symptoms, such as asthma, hay fever, and itchy rash.[25] The evaluations also did not find a detectable difference in dental health as recorded in routine dental exams conducted by a public health dentist.[26] Similarly, no significant differences were observed between 6.5-year-old children exposed to the breastfeeding promotion program and the control group for measures of height, body mass index, waist or hip circumference, triceps or subscapular skinfold thickness, or systolic or diastolic blood pressure. However, cluster adjusted difference in means of 0.2 centimeters in head circumference was marginally significant and a difference of 0.3 centimeters was significant for females.[24] The lack of physical differences between treatment and control children persisted with age; when measured again at age 11.5 years, no detectable difference was found in BMI, fat and fat-free mass indices, percentage of body fat, waist circumference, triceps and subscapular skinfold thickness, being overweight or obese, and whole blood insulin-like growth factor 1.[33] No significant socioemotional differences were observed on the Strengths and Difficulties Questionnaire ratings of total difficulties, emotional symptoms, conduct problems, hyperactivity, peer problems, or prosocial behavior assessed by either mother or the teacher.[28]

While the benefits of breastfeeding on infants' short-term health and nutritional status are well documented, and physical and cognitive developments have been noted among observational studies, the PROBIT studies provide quantitatively rigorous evidence of nuanced, mixed results of the PROBIT program on child development outcomes across a range of domains between 6.5 and 11.5 years of age. Despite the absence of detectable physical and socioemotional benefits, there were some signs of cognitive improvements and schooling gains. The table below presents the longitudinal findings of evaluations of the Belarus PROBIT program by age and outcome domain. The absence of evaluations measuring the effects beyond age six in many of the outcome domains is indicative of a significant research gap in the understanding of longer-term impacts throughout a child's life.

Outcome Domain by Age at Evaluation in Belarus PROBIT Project

Age	Study	Physical Development	Cognitive Develop-ment	Language Develop-ment	Socio-emotional Development	Schooling Outcome	Employment and Labor Market Outcomes
1	Initial study[a]	5/12	—	—	—	—	—
6	[24, 27, 28]	2/13[24]	1/4[27]	3/3[27]	0/16[28]	1/4[27]	—
11	[33, 34]	5/25	—	—	—	—	—

Note: To provide a more complete scope of these studies, this table includes all reported outcomes and not just those analyzed in the main body of this report (see box 1 in the Introduction for the decision rule for selecting outcomes for analysis). The numerator denotes statistically significant outcome at 10 percent level or better, and the denominator is the number of outcomes in the domain. The [bracketed] superscript number in the "Study" column is indicates study identifier (see References).
a. The "initial study" is the most rigorous evaluation of project effects at the earliest age of beneficiaries available. In this case, the initial study is Kramer and others (2001).

The most common anthropometrics found in the included studies are weight related (that is, weight, BMI, BMIZ, WAZ, and WHZ) and height related (that is, height and HAZ). These outcomes are used to define serious limitations in physical growth. Stunting, an indicator of chronic malnutrition, is defined as a height-for-age z-score that is less than minus two standard deviations from the reference median. Wasting or acute malnutrition is defined as a weight-for-height z-score less than minus two standard deviations from the reference median, while overweight is defined as a weight-for-height z-score at least two standard deviations above the reference median. Weight-for-age z-scores that are less than minus two standard deviations from the reference median indicate a child is underweight, which can reflect both acute and chronic malnutrition.

It is widely believed that nutrition during the early childhood stage has a real and lasting effect on children's weight and height (UNICEF 2007; Leroy and others 2014; Black and others 2013; Bhutta and others 2008; Black and others 2008; Walker and others 2007). It follows that interventions that improve early childhood nutrition would be poised to improve post-early childhood anthropometrics. However, the relationship between nutrition and anthropometrics in the developing world is based primarily on noncausal longitudinal studies, near-term impact evaluations, or theory, none of which can authoritatively speak to the post-early childhood effect of nutrition programs.

For most of the studies that can address this hypothesis, interventions to improve ECD did not have a detectable effect on post-early childhood measures of these outcomes. The lack of significant findings does not necessarily mean that early childhood interventions are unable to affect these indicators long-term, but rather implies that an effective method (for example, intervention type, timing, and duration of intervention) for doing so has not yet been established.

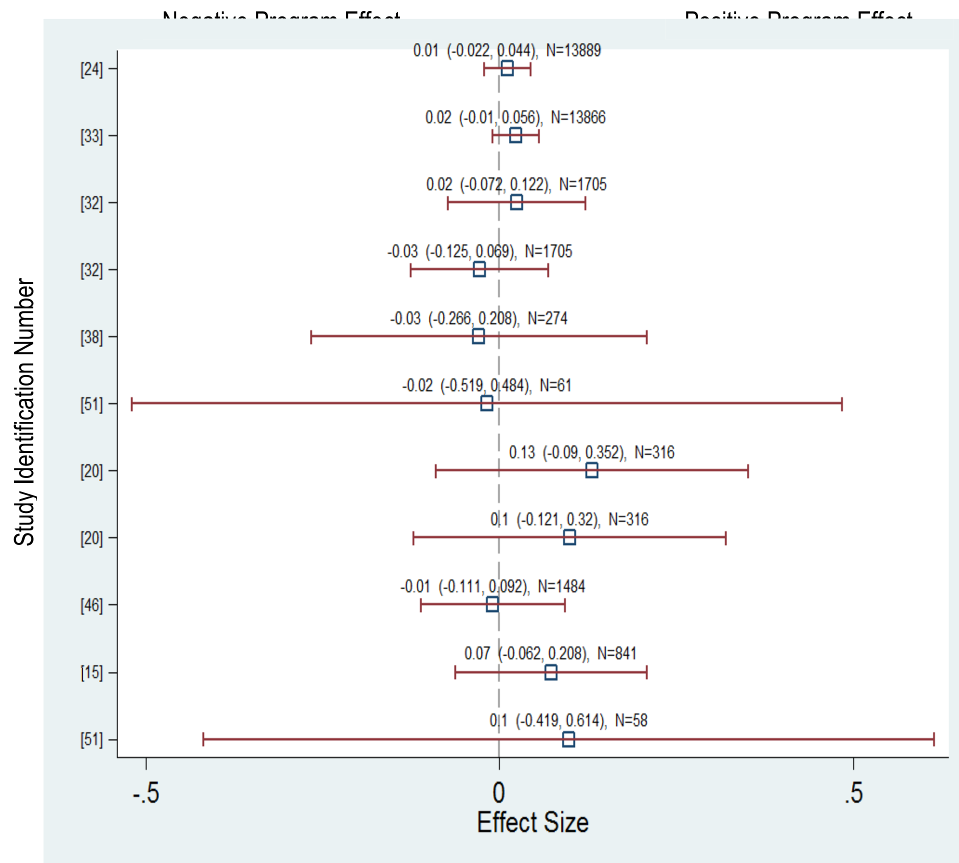
Weight

Early childhood interventions do not appear to have sustained effects on weight.

Evidence from seven different intervention types across five geographical Regions, over a range of ages from six to 17 years old, suggests that early childhood interventions may not have a lasting effect on weight (see figures 1.1 and 1.2). Of the 16 studies that tested for a post-early childhood effect on weight-related measures—WAZ, BMI, BMIZ, or WHZ—none found a significant effect. Meta-analysis on weight-related and BMI outcomes also found no overall effect nor an effect by intervention type (see figure 1.2).

Figure 1.1. Forest Plot for BMI and BMI for Age z-Score

Country	Intervention	Measurement	Average Age at Evaluation (years)	Study and Forest Plot
Belarus	Breastfeeding promotion	BMI	6.5	
Belarus	Breastfeeding promotion	BMI	11.5	
Mexico[a]	CCT cash	BMI for age z-score	9	
Mexico[a]	CCT conditionality	BMI for age z-score	9	
Thailand	Nutrition (child, micronutrient)	BMI for age z-score	9	
Jamaica[b]	Nutrition (stunted children)	BMI	11.5	
Gambia, The[c]	Nutrition (mother)	BMI (girl)	13.8	
Gambia, The[c]	Nutrition (mother)	BMI (boy)	13.8	
Nepal[d]	Nutrition (mother)	BMI	7	
Nepal	Nutrition (mother)	BMI for age z-score	8.5	
Jamaica[b]	Stimulation (stunted children)	BMI	11.5	



Note: The forest plot describes standard mean difference, 95 percent lower and upper bound confidence interval in parentheses, and sample size (N = number). The standard mean difference and confidence interval were calculated by Comprehensive Meta-Analysis software. Study numbers in [brackets] correspond to the numbered studies in References. Studies [32, 38] use 2006 National Center for Health Statistics growth reference to compute standardized z-score. BMI = body mass index; CCT = conditional cash transfer.

a. For Mexico study [32], the CCT effect is disentangled in cash and conditionality, and both estimates are reported in the forest plot. The Mexico [8] study is not included in this forest plot given that it looks at the same average age at evaluation with [32].

b. For The Gambia study [20], each gender estimate is included in the forest plot because the combined total estimate is not available.

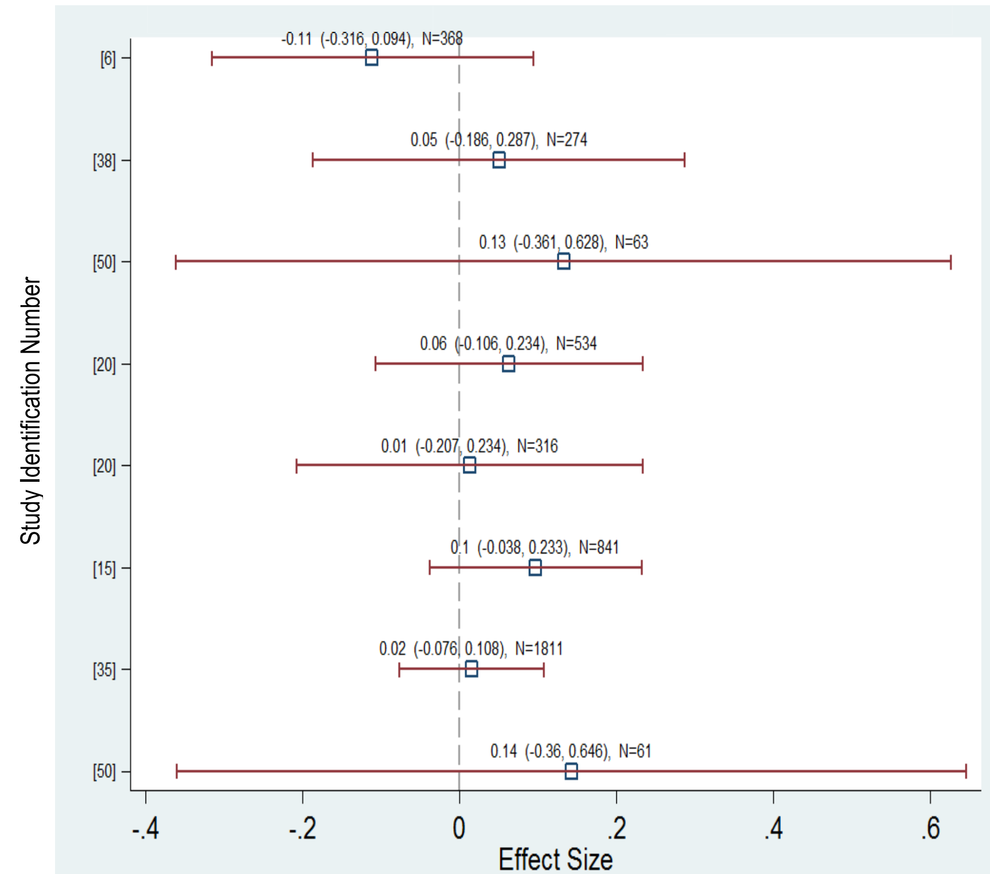
c. For Jamaica study [51], given the combined intervention, information is not available in the article; stimulation intervention compares "stimulation only" and "control group." Similarly, supplementation intervention compares "supplementation only" and "control group."

d. For Nepal study [46], "multiple micronutrient" treatment group compared to the control group is used to compute the effect size.

Figure 1.2. Forest Plot for Weight and Weight-for-Age z-Score

Country	Intervention	Measurement	Average Age at Evaluation (years)
Nicaragua	CCT	WAZ (boy)	10
Thailand	Nutrition (child, micronutrient)	WAZ	9
Jamaica	Nutrition (stunted children)	WAZ	7.7
Gambia, The[a]	Nutrition (mother)	Weight (boy, kg)	13.8
Gambia, The[a]	Nutrition (mother)	Weight (girl, kg)	13.8
Nepal	Nutrition (mother)	WAZ	8.5
Mozambique	Pre-primary program	WAZ	7
Jamaica	Stimulation (stunted children)	WAZ	7.7

Study and Forest Plot



Note: The forest plot shows standard mean difference, 95 percent lower and upper bound confidence interval in parentheses, and sample size (N=number). The standard mean difference and confidence interval were calculated by Comprehensive Meta-Analysis software. Bracketed numbers correspond to the numbered list in References. Study [50] uses National Center for Health Statistics growth reference data to compute standardized z-score. Study [38] uses World Health Organization (WHO) 2006 growth references. Studies [6] and [35] do not; however, because they have been published recently, it is not unlikely these studies use WHO 2006 data as the growth benchmark. CCT = conditional cash transfer. WAZ = Weight for Age z-score.

a. For The Gambia study [20], each gender estimate is included in the forest plot because total estimate is not available in the article.

Height

Although a few interventions resulted in post-early childhood effects on height, the evidence is inconclusive across any specific intervention type.

Evidence shows that it may also be difficult to produce post-early childhood effects on height (see figure 1.3). Although evaluated interventions were more likely to affect height than weight measures, meta-analysis found no overall effect on height-related outcomes. Nor was there a significantly positive meta-effect for the intervention subgroups on height—including for nutrition interventions for which the estimated effect is centered at zero (not shown, available upon request).

When taken individually, only five of 20 impact evaluation (IEs) that examine height outcomes report significant effects. Three of those IEs evaluate nutritional supplementation programs; the other two are a conditional cash transfer program and the Matlab family planning and maternal and child health program. However, other IEs examining nutritional supplementation and cash transfers found no effect, making it difficult to draw firm conclusions based on the available causal evidence.

Despite strong theoretical arguments and IE evidence demonstrating short-term reductions in stunting due to nutrition, the evidence for post-early childhood effects from nutrition programs is thin. Eleven nutrition interventions were evaluated in nine studies, all of which occurred during the first 1,000 days of life when children are thought to be particularly susceptible to stunting (Black and others 2008; Black and others 2013; Bhutta and others 2008). Only two of them had a positive effect on height (the third significant result was negative).

The majority of these interventions (9 of 11) provided micronutrients either to pregnant mothers or to infants.[1, 15, 20, 38, 45, 46] Two of the nine caused a marginally significant increase in height: at nine years old, Thai children who had received iron and zinc as infants experienced a .14 standard deviation increase in HAZ, while six- to eight-year-old Nepalese children whose mothers received folic acid, iron, and zinc while pregnant were 0.64 centimeters taller than those whose mothers did not receive micronutrients.[38, 45]

The third significant micronutrient program, however, led to an initial decrease in height. Protein-rich biscuits given to pregnant mothers in The Gambia caused 11- to 17-year-old girls receiving the supplement in utero to be 1.3 centimeters *shorter* than peers whose mothers received the biscuits for nine months post-partum. There was no corresponding effect for boys, and five years later there was no overall difference in height between children whose mothers had received the supplementation during pregnancy versus postpartum[1] [1, 20] There was no effect on height found from the other six micronutrient interventions.[15, 38, 45, 46]

The remaining two nutrition interventions were the Jamaican supplementary feeding and Belarussian breastfeeding promotion programs.[24, 33, 50, 51] Both were studied soon after early childhood (six to eight years old) and again a few years later (11 to 12 years old), and neither found a detectable effect on children's height.

Taken together, this evidence would seem to indicate that nutrition interventions are generally ineffective at promoting long-term growth, but it is important to note the composition of the available evidence. Only one of the evaluated interventions provided supplementary feeding (though it did not find a lasting effect, possibly due to starting at nine months of age when stunting may have already begun), and there were no causal evaluations of complementary feeding programs or interventions focused on nutrition education. Whereas micronutrient programs typically solely provide essential nutrients foregone because of a limited diet, supplementary and complementary feeding interventions increase caloric intake and have been shown to increase height in the near-term (Bhutta and others 2013). Consequently, it is reasonable to hypothesize that such programs may have a greater and longer-lasting effect on anthropometrics than would micronutrient interventions.[2] At present, the majority of the later-life evidence comes from the latter. So while micronutrient programs seem incapable of producing lasting physical effects, there is almost a complete lack of post-early childhood causal anthropometric evidence regarding feeding interventions; this potential relationship—strongly indicated by theory—should be thoroughly examined before nutrition programs are deemed to be generally ineffective.

Of the three cash transfers—two conditional and one unconditional —only one found a significant overall effect on height, although there is evidence that the unconditional cash transfer (UCT) program significantly helped certain groups. The Mexican conditional cash transfer (CCT) program, Progresa, was examined 10 years after it began by two IEs that looked at the dose-response effect of an additional exposure of 18 months, which occurred before the recipient turned three years old. The first did not find any significant effect on height, but the second used an instrument to isolate the effect of the cash transfers as separate from the program and its conditionalities. It found a highly significant but ultimately small increase in height-for-age because of the cash transfer but no effect stemming from the conditionalities.[3] [8, 32]

The largest effect on height came from the South African Child Support Grant, a UCT to poor households with children, but this effect was isolated to certain groups; there was no effect on children's HAZ overall. The authors theorized the grant would at least be partially spent on improving health and nutrition for the child it was intended to support, but no condition required it.[14] In this particular evaluation of the program, the authors target households with children who were age two years or younger when they received the grant and compares them to children who received the grant starting from two to five years old. There was an increase of approximately 0.2 standard deviations that was marginally significant among girls, and significant among children whose mothers had at least eight years of schooling. The third CCT, which occurred in Nicaragua, found no difference on height-for-age among 10-year-old boys who had started the program in the first 1,000 days rather than as two to five year olds.[6]

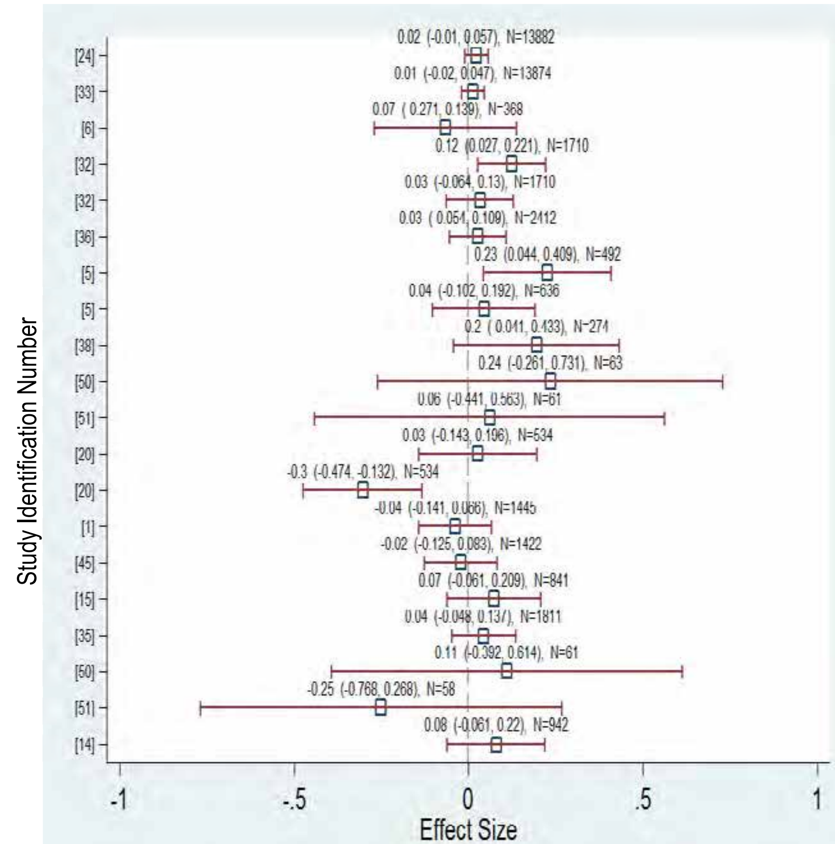
Evidence from two early stimulation programs suggests that this type of intervention may not have a sustained effect on height. Neither the early stimulation program in Jamaica nor community preschool in Mozambique found significant effect on height.[35, 51]

The final two interventions to be evaluated for height were health-related. Starting in 1977, women in Matlab, Bangladesh, were visited in their homes to encourage family planning. Starting in 1982, child health interventions were added in which children under five received immunizations, vitamin A supplementation, and nutrition rehabilitation for children at risk. The impact evaluation that examined this program disaggregated the effects by age group in accordance with when the child health intervention started. The older children—those who were 15–19 years old at evaluation—were not significantly taller than their peers who lived outside the program area, but the younger children—those who were 8–14 years old at evaluation and therefore would have benefited from the child health and nutrition intervention—experienced a significant increase in height-for-age.[5] This evidence suggests that the child health intervention was the driving force behind the change in height, but such a conclusion has not been definitively established. The other health-related IE studied the indirect effect of a community deworming project in western Kenya by focusing on the preprimary younger siblings of primary school children who received the treatment. Ten years later, there was no effect on height or height-for-age.[36]

Figure 1.3. Forest Plot for Height and Height for Age z-Score

Country	Intervention	Measurement	Average Age at Evaluation (years)
Belarus	Breastfeeding promotion	Height (cm)	6.5
Belarus	Breastfeeding promotion	Height (cm)	11.5
Nicaragua[a]	CCT	HAZ (Boy)	10
Mexico[b]	CCT cash	HAZ	9
Mexico[b]	CCT conditionality	HAZ	9
Kenya	Deworming	HAZ	10
Bangladesh	Family planning and MCH	Height z-score	9
Bangladesh	Family planning, maternal health	Height z-score	9
Thailand	Nutrition (child, micronutrient)	HAZ	9
Jamaica	Nutrition (stunted children)	HAZ	7.7
Jamaica[c]	Nutrition (stunted children)	HAZ	11.5
Gambia, The[a]	Nutrition (mother)	Height (girl, cm)	13.8
Gambia, The[a]	Nutrition (mother)	Height (boy, cm)	13.8
Gambia, The[a]	Nutrition (mother)	Height (cm)	19.6
Nepal[d]	Nutrition (mother)	Height (cm)	7
Nepal	Nutrition (mother)	HAZ	8.5
Mozambique	Pre-primary program	HAZ	7
Jamaica	Stimulation (stunted children)	HAZ	7.7
Jamaica	Stimulation (stunted children)	HAZ	11.5
South Africa	Unconditional cash transfer	HAZ	10

Study and Forest Plot



Note: The forest plot describes standard mean difference, 95 percent lower and upper bound confidence interval in parentheses, and sample size (N = number). The standard mean difference and confidence interval were calculated by Comprehensive Meta-Analysis software. Bracketed numbers correspond to the numbered studies in References. Study [50, 51] use National Center for Health Statistics growth reference data to compute standardized z-score, and study [32, 36, 38] use WHO 2006 growth reference. Study [6, 14, 35] is not clear which growth reference data is used for computing the height for age z-score, but given they are published relatively recently, they are likely to use WHO 2006 growth reference. Study [5] height z-score is standardized by subtracting the comparison group mean and dividing by that group's standard deviation for people of the same age and gender. CCT = conditional cash transfer; MCH = maternal and child health; WHO = World Health Organization. HAZ = Height-for-age z-score.

a. For The Gambia study [20], each gender estimate is included in the forest plot because total estimate is not available in the article. Only boy's estimate is available for Nicaragua study [6].

b. For Mexico study [8], it is not included in this forest plot because it looks at same average age at evaluation with [32].

c. For Jamaica study [51], given the combined intervention information is not available in the article, stimulation intervention compares "stimulation only" and "control group." Similarly, supplementation intervention compares "supplementation only" and "control group."

d. For Nepal study [45], "multiple micronutrient" treatment group compared to the control group is used to compute the effect size.

Mid-Upper-Arm Circumference and Head Circumference

There is little evidence that early childhood interventions cause significant differences in MUAC or head circumference in the post-early childhood period.

Three studies examined head circumference or MUAC. Two of the IEs looked at the effect of the PROBIT in Belarus on both measures,[24, 33] while the third study examined the effect of the Thai iron and zinc intervention on MUAC.[38] None of the IEs found a significant effect overall on either outcome, although the head circumference of girls age 6.5 years who participated in PROBIT was 0.3 centimeters larger than girls who did not participate.[24] Gender-disaggregated results were not discussed in the later study of the same program, so it is not possible to determine if this difference persisted. However, the lack of a lasting effect on head circumference, and to a lesser extent on MUAC, is not necessarily concerning as it is really in the first two to three years of life when head circumference is routinely measured and used as an indicator of potential problems in the functional correlates around physical health.

Fine Motor Skills

Fine motor skills may be improved through preschool participation.

The final outcome examined is fine motor skills. Despite little evidence on this outcome, the intrinsic value as an indicator of school readiness warrants its inclusion (Grissmer and others 2010; Cameron and others 2012). In Mozambique, children who had participated in preschool experienced a marginally significant 6.3 percent increase in their fine motor skills scores over the control group.[35] Conversely, a program that promoted early stimulation by placing Romanian orphans with foster care families rather than in institutional homes had no detectable effect on the fine motor skills of 8-year-old foster care children.[29] Given the many differences in the two interventions, it is impossible to determine why one was effective while the other was not.

Physical Development Summary

Despite impact evaluations across various intervention types, Regions, and age groups, very little causal evidence exists of what works to create a sustained effect on a child's anthropometrics. There were no interventions that caused a detectable change in weight and few that affected height.

Given that three of the five effective interventions on height were nutrition programs, these results may seem to reinforce the widely held view that early nutrition is significantly positively linked to anthropometrics, and particularly to height (UNICEF 2007; Leroy and others 2014; Black and others 2013; Bhutta and others 2008; Black and others 2008; Walker and others 2007). However, the other eight nutrition interventions evaluated for their effect on height did not cause a notable change in beneficiaries' stature. Nevertheless, with evidence missing on the later-life effects of major nutrition interventions, additional high-quality IE evidence is necessary to determine whether physical benefits from early nutrition programs do indeed fade over time (as these results imply) or whether interventions that cause later-life effects merely have yet to be evaluated.

The remaining physical outcomes either appear unchanged by interventions during early childhood or else the evidence base is too thin to draw conclusions. Fine motor skills may be improved by preschool, but with two IEs on the outcome, this conclusion is still tentative. Neither MUAC nor head circumference were found to be significantly affected by the evaluated interventions, except for among 6.5-year-old girls in Belarus.[24] There is some evidence that even that effect might disappear four years later.

Almost all of the effective interventions occurred during the first 1,000 days from conception to the child's second birthday. The only intervention that was not specifically limited to that time period—Progresa—was studied by isolating the effect of program participation up to age three. While this reinforces the widely held belief that that time period is especially important, the evidence also suggests that merely intervening during the first 1,000 days is not a panacea. Other interventions, such as the supplementation and early stimulation project in Jamaica[50, 51] and the maternal supplementation project in The Gambia,[1, 20, 22] also occurred during that time period but did not produce significant effects.

---

[1] While the effect on girls was highly statistically significant, there is some question as to the best interpretation of these results. In the original trial, pregnant women in the treatment group received protein-rich biscuits daily starting at 20 weeks' gestation. As the original outcome of interest was birthweight, the control group then received the treatment for 20 weeks after delivery, which in effect rendered subsequent measurements as dose response effects between prenatal and postnatal supplementation. However, there is some evidence from that same region of The Gambia that postnatal supplementation does not affect the quality or quantity of breast milk (Prentice and others 1983). If this is true, then the original intervention design would hold, and the results would represent the effect of prenatal supplementation alone.

[2] A supplementary feeding program in Colombia provides some evidence for this. In this program, participating children and their families received nutritional supplementation for the first three years of the participant's life, which caused a marginally significant increase in participant's height. However, this study did not pass the rigorous quality check employed by

this review because of concerns over high attrition (approximately 45 percent) (Super and others 1990).

[3] There are two versions of this study—the version published in *The Lancet* (Fernald and others 2008) and a subsequent working paper in which the authors address critiques they received on their original paper.[32] This review chose to use the results from the latter version as they were based on an instrumental variable and therefore addressed potential endogeneity problems from using actual transfer amounts as an explanatory variable for cognitive outcomes (children who are successful in school go on to higher grades, in which they earn larger transfers). The results between the two papers are largely similar, with only a change in significance level for two of the outcomes (see table 1.2 of study [32]).

2. Cognitive Development

Cognitive development is improved by a range of interventions, and most improvements are seen in measures of general cognition. Nutritional programs had almost no effect.

Cognition, or cognitive ability, includes problem-solving and analytical skills, short- and long-term memory, math abilities, verbal comprehension, general knowledge, and logic as well as how people respond to new or challenging situations (Johnson 1998). To date, there is no single, universal standardized assessment for measuring cognition across ages and cultures, unlike measures of physical growth, such as weight and height. For this reason, significant heterogeneity is found in the assessments used to capture an individual's capacities.

Explanation of Cognitive Tests

Most intelligence tests are comprised of subscales that measure several aspects of cognitive function. For example, the Wechsler tests for children and adults, widely used in research and as a clinical tool, include subscales assessing verbal abilities (for example, vocabulary and analogies) and nonverbal abilities (for example, perceptual skills and working memory). Most intelligence quotient (IQ) tests include a total, overall score as well as scores on these subscales, such as verbal abilities and performance or nonverbal abilities. Some tests are comprehensive ability tests that typically take at least an hour to complete, while others are abbreviated IQ or screening tests (completed in 15 minutes or so). Generally, the longer tests are considered to be more precise, sensitive, and exhaustive measures of cognition, while the shorter tests provide a more crude estimation of abilities.

Widely used, brief cognitive assessments, like pattern completion tests such as the Raven's matrices, are designed to measure specific aspects of cognition (for example, visual processing). The Raven's matrices in particular have been frequently used throughout the world in part because they are nonverbal assessments that (i) do not require the respondent to be literate or (ii) to have a certain acquired knowledge, and (iii) they are believed to be a highly valid measure of fluid intelligence.[1] Other tests included in this review are measures of executive function abilities[2]—processes such as short-term memory, ability to sustain attention, ability to shift strategies as needed, and inhibition of impulsive responses that reflect how an individual responds to new or challenging situations. These types of tests are often included in lieu of or in addition to more global measures.

While comprehensive measures typically include language subscales, this review separates all verbal and language-related tests from other cognitive assessments to make comparisons of impacts easier to understand across specific outcomes. Language outcomes are reported subsequently.

Evidence from Developed Countries

Several reviews of primarily center-based early child programs in the United States have examined the concurrent and long-term effects on cognition and other related outcomes. A 2003 systematic review of the effectiveness of comprehensive, publicly funded programs revealed sustained positive effects on cognition. Significant increases in IQ were observed in children during their participation in the program, one year later, and between three and 10 years later (Anderson and others 2003). More recent reviews of early education-focused interventions have also consistently found enduring effects on current and later cognitive abilities as well as achievement scores (Camilli and others 2010; Reynolds and Temple 2008).

Despite the existence of evidence showing benefits throughout the lifecycle, several U.S. studies have observed a disappearance of cognitive gains during the early primary years and a subsequent reemergence of these gains later in the child's life (Bitler, Hoynes, and Domina 2014; HHS 2010; Magnuson, Ruhm, and Waldfogel 2007, ultimately translating into improved schooling and labor market outcomes and better scores on tests of externalizing behaviors (Reynolds and others 2001).[3] Interpretations of these findings posit that inputs such as the availability of classroom materials and teaching styles are important for determining whether effects of interventions in preschool are sustained (Barnett 2011). Such findings reaffirm the need for more research on the medium- and long-term effects of early childhood development (ECD) interventions.

Although the majority of the evidence from the United States and developed countries comes from educational programs, parenting support and nonpreschool programs—typically targeting low-income families with children under three years—have also shown the ability to improve cognitive, health, and other developmental outcomes for children in a range of populations and contexts. These programs promote stimulation and teach various parenting skills through home visits, particularly by nurses or other healthcare professionals (Olds, Sadler, and Kitzman 2007). The findings from these programs suggest that similar efforts in low- and middle-income countries could also be successful.

Evidence from Developing Countries

The majority of the ECD literature from the United States and other developed countries focuses on early education and stimulation programs that aim to improve cognitive ability. However, the lower nutritional status and higher mortality rates of children in the developing world have necessitated the prioritization of programs that improve child survival and growth by promoting nutritional supplementation and appropriate feeding practices. Stimulation programs, especially scaled up to the national level, are also typically more expensive and difficult to implement effectively, meaning that very few have been undertaken. While this has contributed to the dearth of longitudinal data on ECD stimulation programs everywhere, the lack of evidence is most pronounced in the developing world.

Evidence from ECD programs in low- and middle-income countries has shown that early childhood interventions can improve children's cognitive, behavioral, health, and schooling outcomes. Among the programs reviewed, interventions that worked directly with children and their families, target more disadvantaged populations, and are integrated with health, nutrition, and educational services provided the largest benefits to children's development (Engle and others 2007, 2011). A review of early childhood interventions in 23 non-U.S. countries looked at both short- and long-term effects[4] of a range of ECD intervention types. Although cash transfers, nutritional supplementation, and educational programs were found to create significant and sustained gains across a range of child development outcomes, interventions including a stimulation component proved to be the most effective (Nores and Barnett 2010; Yousafzai, Rasheed and Bhutta 2013). Despite these encouraging findings, both program coverage and evaluation evidence for interventions from early childhood on later outcomes in developing countries, is low, as evidenced by the relatively few impact evaluations (IEs) and reviews that address this area.

The absence of universal, standardized outcome measures for cognitive and language ability increases the difficulty of designing ECD programs against concrete and well-defined benchmarks. In addition, many countries do not have national policies or guidelines for early childhood education and stimulation programs. However, there is an increasing recognition among developing countries that intervention types other than those focusing on health and nutrition are needed for the poorest children to fully develop their cognitive abilities. There has been a particular interest in combining education or stimulation with other ECD programs to improve nutritional status of young children and prevent the negative effects of stunting (Aguero, Carter, and Woolard 2007; Yousafzai, Rasheed, and Bhutta 2013). Nevertheless, much work remains to translate evidence into policy and ensure the quality and consistency of interventions.

Importance of Measuring Cognition

Cognitive ability plays an important role in enabling an individual to achieve success in areas such as schooling and employment. There is substantial evidence from the United States that low socioeconomic position in early childhood is associated with differential brain development (Hackman and Farah 2009; Raizada and Kishiyama 2010) and with poorer cognitive performance (Bradley and Corwyn 2002; Noble and others 2012; Pechtel and Pizzagalli 2011). Recent research from low- and middle-income countries also suggests that differences in language abilities between socioeconomic groups are apparent at a young age and that the differences persist and even increase once children enter school (Fernald and others 2011; Schady and others 2014). Consequently, one might hope to improve cognition at a young age by using early childhood interventions, and such improvements could have lasting effects not only in a person's cognitive development, but also on their subsequent socioeconomic status.

Table 2.1 maps the 16 unique studies in this review that investigate cognitive outcomes. The studies span 11 projects in 10 countries and include 9 distinct intervention types. Measurements of abbreviated, full-scale, and performance IQ, nonverbal cognition, and executive function are reported. The studies are grouped by the outcome category in which they report measurements; the specific intervention within each category is noted. More detail on each intervention is provided in appendix A. These classifications align with the taxonomy presented in figure 2 in the Introduction.

Table 2.1. Impact Evaluations Investigating Cognitive Development

	Study	Country (Project)	Average Age at Intervention (Years)	Average Length of Exposure (Years)[a]	Age at Evaluation (Years)	Evaluated Intervention	Reviewed Outcomes
Nutrition	Kramer and others 2008a [27]	Belarus (Promotion of Breastfeeding Intervention Trial [PROBIT])	0	1	6	breastfeeding promotion	abbreviated IQ (total score)*; abbreviated performance IQ; nonverbal subscales
Nutrition	Pongcharoen 2010 [38]	Thailand (micronutrient supplementation to children)	0.5	0.5	9	micronutrients and fortified food for children (Iron and/or zinc supplementation)	executive function (processing speed); full-scale IQ; performance IQ; nonverbal cognition (Raven's matrices)
Nutrition	Alderman and others 2014 [1]	Gambia, The (maternal supplementation)	in utero	0.5 (DR1)	16–22	vitamins, micronutrients, or fortified food for pregnant women (protein biscuits)	executive function (backward digit span); nonverbal cognition (Raven's matrices)
Nutrition	Walker and others 2005 [52][b]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	17–18	supplementary feeding	full-scale IQ; performance IQ; nonverbal cognition (Raven's matrices)
Nutrition	Walker and others 2011 [55][b]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	22	supplementary feeding	full-scale IQ; performance IQ
Nutrition	Maluccio and others 2009 [31][c]	Guatemala (INCAP supplementary feeding to children)	0	5.3	25–42	supplementary feeding	nonverbal cognition (Raven's matrices)**
Early Learning/ Childcare	Walker and others 2010 [54]	Jamaica (stimulation to low birthweight infants)	0	2	6	stimulation	executive function (short-term memory)***; full-scale IQ; performance IQ**

	Fox and others 2011 [17]	Romania (Bucharest Early Intervention Project)	1.88	2.7	8	stimulation (foster care)	executive function (processing speed); executive function (working memory); perceptual organization full-scale IQ*
	Bos and others 2009 [11]	Romania (Bucharest Early Intervention Project)	1.88	2.7	8	stimulation (foster care)	executive function (spatial working memory); executive function (stockings of Cambridge)
	Walker and others 2000 [51]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	11–12	stimulation	executive function (processing speed); full-scale IQ**; nonverbal cognition (Raven's matrices)**; performance IQ*
	Walker and others 2005 [52][b]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	17–18	stimulation	full-scale IQ**; nonverbal cognition (Raven's matrices)*; performance IQ*; working memory
	Gertler and others 2013 [18]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	22	stimulation	cognitive factor score***
	Walker and others 2011 [55][b]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	22	stimulation	full-scale IQ***; performance IQ***
Health	Cas 2012 [12]	Indonesia (Safe Motherhood program)	in utero	3	11–14	Access to obstetric and child health care	nonverbal cognition (Raven's matrices)***
Health	Ozier 2013 [36]	Kenya (primary school deworming project)	0	1	8–15	deworming	cognitive factor score**; nonverbal cognition (Raven's matrices)***
Social Protection	Behrman and others 2008 [8]	Mexico (Progresa)	1.5	1.5 (DR2)	7–11	CCT	abbreviated performance IQ

Manley, Fernald, and Gertler 2012 [32]	Mexico	1	1.5 (DR2)	8–10	CCT—conditionalities	abbreviated cognitive performance IQ
	(Progresa)				CCT—cash	abbreviated cognitive performance IQ*
Barham and others 2014 [6]	Nicaragua	in utero	3 (DR2)	10 (boys)	CCT	executive function (processing speed)**; cognitive factor score***; nonverbal cognition (Raven's matrices)**
	(Red de Protección Social)					

Note: Bracketed numbers correspond to numbered studies in References. More details for each study are found in appendix A. CCT = conditional cash transfer; DR = dose response; INCAP = Instituto de Nutrición de Centroamérica y Panamá; IQ = intelligence quotient.
a. DR in the length of exposure means the intervention period in terms of the dose response. DR is either randomized rotation (DR1) or phase-in (DR2). In terms of dose response, for instance, DR1 indicates the study from The Gambia where treatment group receives protein biscuit only in utero whereas control group receives it only in postpartum, and length of exposure is the length of intervention for treatment. DR2 describes the dose response where early and late treatment effect is compared, and length of exposure is the difference of the intervention period between treatment and control group.
b. Jamaica [52, 55] each has a multiple intervention arm, and each intervention type has a separate row for these studies.
c. INCAP provided supplementation to pregnant and lactating women but could not isolate effects as children could also receive the supplement after birth, and the study lacked power to evaluate the intervention by developmental period.

* Statistically significant at 10 percent.
** Statistically significant at 5 percent.
*** Statistically significant at 1 percent.

Box 2.1. The Psychosocial Component of the Jamaica Supplementation and Stimulation ECD Program Produces Lasting Cognitive Benefits

Evaluations of Jamaica's stimulation and nutritional supplementation program tracked participants—stunted children between the ages of nine and 24 months when enrolled in the program—for more than 20 years to assess the long-term effects of the intervention on physical, cognitive, educational, and employment outcomes. Stunting, which is defined as height-for-age less than two standard deviations below the mean, is an indicator of chronic malnutrition that begins early in life (Victora and others 2008). It can be caused by many factors: poor nutrition during pregnancy, early introduction of nutrient-poor liquids and foods, illness, and a diet lacking in fruits, dairy, and protein (Black and others 2013). In addition to poor growth and susceptibility to illness, stunted children are also at risk for impaired cognitive development. Early supplementation has been proposed as a method to facilitate mental and physical development among malnourished children. However, the results of longitudinal studies of a randomized controlled program in Jamaica suggest the benefits of supplementation to stunted children dissipated relatively quickly after the program, whereas benefits of a psychosocial stimulation program provided to stunted children were sustained into adulthood.[18, 19, 50, 51, 52, 53, 55]

Stunted children from poor communities in Kingston participated in a two-year randomized controlled trial during which they were placed in one of four experimental groups: milk-based supplementation, stimulation through weekly home visits from a healthcare worker, supplementation plus stimulation, and a control group of stunted children who were not exposed to either intervention. For both groups receiving stimulation, the health worker facilitated interactions between mother and child, reinforcing positive messaging, engagement with toys, and active play (Grantham-McGregor and others 1991). A group of nonstunted children was later identified through a matching process for comparison purposes.

During the first two years of the program, participants' dietary quality, physical growth, locomotor skills, and hearing and speech were measured every six months. At the end of the 24 months, both the supplementation only and stimulation only groups showed improved hand and eye coordination and locomotor performance. Furthermore, the children receiving the combined interventions performed significantly better than children in the groups receiving either intervention independently. No effect was detected in hearing or speech tests (Grantham-McGregor and others 1991). Supplementation also significantly improved height, weight, and head circumference at 12 months after enrollment, although most of these physical benefits tapered off after the first six months. Stimulation alone did not have a significant effect on physical outcomes during the first 12 months (Walker and others 1991).

Later evaluations collapsed the two stimulation treatment arms (stimulation alone or with supplementation) into one group. At age 7–8 years,[50] 11–12 years,[51] 17–18 years,[52, 53] and 22 years,[18, 19, 55] children who had received stimulation (alone or with supplementation) were compared to children who had not (combining "supplementation only" and the pure control groups). Despite the initial gains observed in mental and physical development, these follow-up studies found that supplementation alone did not cause significant improvements in children's development between the ages of seven and 22 years.[50, 51, 52, 53, 55]

In contrast to the absence of effects observed in the supplementation groups, there were sustained benefits of stimulation across cognitive, language, and schooling outcomes. When assessed between ages 11 and 12, children who received stimulation (either alone or with supplementation) showed better cognitive functioning: they scored higher on IQ tests, verbal scales, and vocabulary exercises; in fact, test scores of the stunted but stimulated nearly caught up with observably comparable nonstunted children, suggesting that early stimulation may be able to mitigate some of the functional consequences of growth restriction in young children.[51] These cognitive advantages were again observed at 17 and 18 years of age.[52] In addition, the stimulation groups had lower dropout rates and higher scores on a battery of educational tests than the combined supplementation only and control group. When evaluated again at age 22, they also had completed more years of schooling.[55]

Children who received stimulation also exhibited better psychosocial functioning and achieved improved employment outcomes. Sixteen years after the intervention, they reported less anxiety and depression, higher self-esteem, and demonstrated better attentional abilities and less oppositional behavior when compared with the individuals who had not received stimulation. However, no difference was detectable in self-reported antisocial behavior between the groups.[53] A 20-year follow up study found that the stimulation intervention had an effect on adult employment. Individuals who received stimulation during their early childhood years reported 25 percent more earnings than nonstimulated individuals, putting them on par with the nonstunted group.[18, 19] These results present a clear picture that psychosocial stimulation produced lasting cognitive gains in stunted children within the context of the Jamaica study, suggesting that it may be able to mitigate some of the functional consequences of growth restriction in young children.

The table below presents the longitudinal findings from the series of studies evaluating the Jamaica early supplementation and stimulation program.

Outcome Domain by Age at Evaluation in the Jamaica Supplementation Project

Age	Intervention Arm	Physical Develop-ment	Cognitive Develop-ment	Language Development	Socioemotional Development	Schooling	Labor Market
≤ 5y[a]	Supplementation	1/2	1/1	0/1	-	-	-
	Stimulation	2/2	1/1	1/1	-	-	-
	Both	2/2	1/1	1/1	-	-	-
7–8 y	Supplementation	0/8[50]	-	-	-	-	-
	Stimulation	0/8[50]	-	-	-	-	-
	Both	0/8[50]	-	-	-	-	-
11–12 y	Supplementation	0/5[51]	0/8[51]	1/4[51]	-	-	-
	Stimulation	0/5[51]	5/16[51]	4/8[51]	-	-	-
	Both	1/5[51]	3/8[51]	2/4[51]	-	-	-
17–18 y	Supplementation	0/2[52]	0/12[52]	1/10[52]	1/8[53]	0/2[52]	-
	Stimulation	0/2[52]	7/12[52]	10/10[52]	6/16[53]	0/2[52]	-

	Both	0/2[52]	2/6[52]	4/5[52]	1/8[53]	0/1[52]	-
22 y	Supplementation	0/1[55]	0/5[18, 55]	0/4[55]	0/9[18, 55]	0/7[18, 55]	-
	Stimulation	0/1[55]	5/6[18, 55]	4/4[55]	4/11[18, 55]	10/16[18, 55]	5/8[18, 55]
	Both	0/1[55]	3/3[18, 55]	1/2[55]	2/3[18, 55]	1/4[18, 55]	-

Note: To provide a more complete scope of these studies, this table includes all reported outcomes and not just those analyzed in the main body of this report (see box 1 in the Introduction for the decision rule for selecting outcomes for analysis). The numerator denotes statistically significant outcome at 10 percent level or better, and the denominator is the number of outcomes in the domain. The [bracketed] superscript number in the "Study" column is indicates study identifier (see References). Results for supplementation include supplementation only versus the control group (unadjusted) and/or supplementation only and both versus the control and early stimulation only (adjusted with covariates). Results for stimulation include early stimulation only versus the control group (unadjusted) and/or early stimulation only and both versus the control and supplementation only (adjusted with covariates). "Both" is for the group that received both early stimulation and supplementation as compared to the control group (unadjusted). "Age" is the age at evaluation.

a. The sources for effects during early childhood are Gardiner and others (2003), Grantham-McGregor and others (1991), and Walker and others (1991, 2004).

General Cognition

General cognition was improved by stimulation, social protection, and sanitation interventions. Nutritional interventions did not have a measurable effect across a range of contexts.

Cognitive outcomes were reported for 10 early childhood programs (figure 2.1). However, only two of these programs, early stimulation and supplementation for stunted children in Jamaica,[51, 52, 55] and a conditional cash transfer in Mexico,[32] report results over multiple stages of the child's life. Evaluations of these programs suggest that ECD interventions have the ability to produce robust and sustained effects on cognitive outcomes. However, the lack of longitudinal data highlights the need for more research to fully understand the effects of the range of ECD interventions on cognitive outcomes throughout a child's lifetime.

Supplementation programs across a range of contexts did not demonstrate significant and sustained effects on multiple measures of cognition. Neither iron nor zinc, given together or individually in infancy, caused a significant difference in full-scale IQ or performance IQ at age nine in Thailand.[38] The PROBIT breastfeeding promotion program in Belarus sustained only a marginally significant effect up to age 6.5 on both abbreviated IQ and abbreviated performance IQ.[27] The supplementation-only component of the Jamaica stunted children interventions did not have a significant effect on any measures of cognitive ability from age six through 22.[18, 51, 52, 55] However, it may be that once stunting occurs—as with this study population—then supplementation alone is not sufficient for producing measurable improvements in cognitive outcomes. This conclusion is supported by evidence suggesting stunting is irremediable after the age of two or three years (Victora and others 2008). These findings support the assertion that the timing and duration of supplementation, in addition to other types of interventions, may be needed to affect the full range of cognitive outcomes beyond the early childhood years, particularly for children who are stunted (as in Jamaica) or are at risk for malnutrition.

The available evidence does suggest that programs incorporating stimulation were much more effective than nutritional programs alone in improving cognitive outcomes beyond the early childhood period. The same study in Jamaica found that already stunted children who were exposed to a stimulation component of the program had significantly higher full-scale IQ scores at ages 17–18 than children who had received only supplementation,[52] and the gains became much larger and highly significant at age 22.[55] Improvements in performance IQ were weakly significant for stunted children at ages 11–12[51] and 17–18 years,[52] but by age 22 the gap between children who had received stimulation either alone or with supplementation and those who had received supplementation only had widened and was highly significant.[55] A different early stimulation program in Jamaica that targeted low-birthweight children did not produce a measurable effect on full-scale IQ at age six years, but it did cause a significant increase in performance IQ.[54]

A study of a foster care program in Romania also found evidence that exposing children to a stimulating environment during early childhood can improve cognitive outcomes. The Bucharest Early Intervention Project (BEIP) randomly assigned Romanian orphans to either remain in institutional care or be placed with foster families. When assessed at age eight, 3.5 years after the program ended, children who had been placed in foster care exhibited marginally significant gains in full-scale IQ scores.[17]

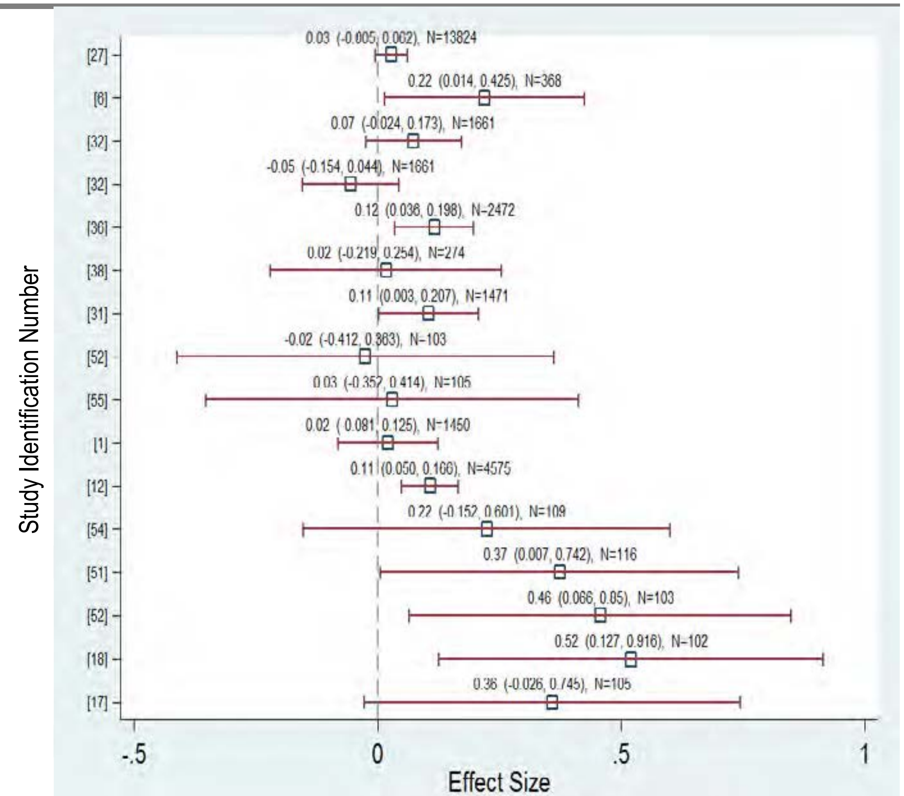
The evidence on the effect of CCTs on general cognition is mixed and particularly thin. An additional 18 months' participation in Progresa, a Mexican CCT that provided nutritional supplements to children under two years old and older if they showed signs of nutritional deficit, and health visits to children under five years old (among other program components), did not produce significant differences in abbreviated performance IQ at ages 7–11.[8] However, a second study of the same program was able to separate the effects of the conditionalities and the transfers and found that the cash component led to a marginally significant improvement in abbreviated performance IQ in 8- to 10-year-olds, while the conditionalities alone caused no detectable difference in cognitive scores.[32] The evaluation of Red de Protección Social, a CCT in Nicaragua[6] that required beneficiaries to receive regular health checkups and growth monitoring for children under five years old, reports some positive effects on cognition based on an index of a broad range of cognitive tests. Scores from Raven's matrices, the Peabody vocabulary test, the Denver screening test, digit spans, and subscales from the Wechsler

Intelligence Scale for Children III were processed into a cognitive factor score. When evaluated at age 10, no significant effects were observed for girls, though boys whose families received the program throughout the first three years of their lives showed a highly significant improvement in the cognitive index measure over boys whose family began the program when they were three to six years old.

Figure 2.1. Forest Plot for General Intelligence

	Intervention	Measurement	Average Age at Evaluation (years)
Belarus	Breastfeeding promotion	WASI	6.5
Nicaragua[a]	CCT	Raven's matrices	10
Mexico[b]	CCT cash	WASI	9
Mexico[b]	CCT conditionality	WASI	9
Kenya	Deworming	Raven's matrices	10
Thailand	Nutrition (child, micronutrient)	WISC-III	9
Guatemala[c]	Nutrition (child, SF)	Raven's matrices	32.3
Jamaica[d]	Nutrition (stunted children)	WAIS	17.5
Jamaica[d]	Nutrition (stunted children)	WAIS	22
Gambia, The	Nutrition (mother)	Raven's matrices	19.6
Indonesia	Safe motherhood program	Raven's matrices	12.5
Jamaica	Stimulation (LBW children)	WPPSI	6
Jamaica[d]	Stimulation (stunted children)	WISC-R	11.5
Jamaica[d]	Stimulation (stunted children)	WAIS	17.5
Jamaica[d, e]	Stimulation (stunted children)	Cognitive factor	22
Romania	Stimulation (foster care)	WISC-IV	8

Study and Forest



Note: The forest plot describes standardized mean difference, 95 percent lower and upper bound confidence interval in parentheses, and sample size (N = number). The standardized mean difference and confidence interval were calculated by Comprehensive Meta-Analysis software. Bracketed numbers correspond to the numbered studies in References. The WASI, WISC, WAIS, and WPPSI reported in this forest plot use total scale scores (i.e., full-scale IQ). CCT = conditional cash transfer; WASI = Wechsler abbreviated scale of intelligence; WISC = Wechsler intelligence scale for children; WAIS = Wechsler adult intelligence scale; WPPSI = Wechsler preschool and primary scale of intelligence. LBW = Low birthweight. SF = supplementary feeding.

a. Nicaragua study [6] includes Cognitive outcome, but they are measured through Denver Development Screening Test, which is not necessarily comparable to other Wechsler scale. Therefore, Nicaragua study uses Raven's Colored Matrices included in the study.

b. For Mexico study [8], it is not included in this forest plot because it looks at same average age at evaluation with [32].

c. Guatemala study [31] uses first three of five scales in Raven's matrices as non-verbal cognitive ability outcome.

d. For Jamaica studies [18, 51, 52, 55] on stunted children, the stimulation compares "stimulation only + stimulation and supplementation" vs "no intervention + supplementation only". Similarly, the supplementation compares "supplementation + supplementation + stimulation" vs "no intervention + stimulation".

e. Jamaica study [18] uses (i) WRAT math, (ii) WRAT reading comprehension, (iii) PPVT, (iv) Verbal analogies, (v) Raven's matrices, (vi) WAIS full-scale IQ tests to compute cognitive factor through factor analysis.

Nonverbal Cognition

Evidence of the effects of early childhood interventions on nonverbal cognition is particularly thin. Although positive outcomes were found for stimulation, deworming, and social protection interventions, these results are only observed for a single program in each intervention type. Only one of four nutritional interventions improved nonverbal cognition.

Nonverbal cognition reflects an individual's ability to reason and recognize relationships between concepts. Since assessments do not rely upon verbalization or existing knowledge, measures of nonverbal cognition can provide insight into the cognitive abilities of a child who might otherwise be limited by poor language abilities. The most commonly applied assessment of nonverbal cognition in the included IEs is Raven's Progressive Matrices, which ask an individual to progress through increasingly difficult pattern recognitions.

Single and multiple micronutrient supplementation interventions did not provide a statistical benefit for nonverbal cognition. No significant improvements in Raven's matrices tasks were observed in nine-year-old children who had participated in the Thailand iron and/or zinc program that compared infants four to six months of age who received supplementation for six months with infants who received a placebo.[38] Offspring of pregnant women in The Gambia who received two high energy, high protein biscuits daily for 20 weeks showed no detectable gains in nonverbal cognition at 16–22 years old when compared to children whose mothers received the biscuits for 20 weeks postnatally.[1] The nutritional supplementation component of the Jamaica also did not find any benefit on tests of nonverbal cognition compared to a control group when program participants were measured between the ages of 17 and 18 years old.[52]

The only nutritional intervention to improve nonverbal cognitive outcomes was the supplementary feeding program of Guatemala's Instituto de Nutrición de Centroamérica y Panamá. A high-energy high-protein supplement was made available to all pregnant and lactating women and children under six. Children who had received the supplement between birth and three years old registered large and lasting advantages in fluid intelligence when tested between the ages of 25 and 42.[31] This gives significant weight to the position that if (and perhaps only if) a nutrition intervention is sustained throughout the first thousand days of a child's life will a nutrition intervention be able to yield results that last beyond early childhood.

The effects of stimulation on nonverbal cognition was measured in two studies, both of which evaluated the Jamaica supplementation and stimulation program for stunted children. The early stimulation component resulted in significant improvements in nonverbal cognition when children were between 11 and 12 years old.[51] Those benefits were still large but marginally significant (perhaps because of the small sample size) when children were again evaluated between 17 and 18 years of age.[52] Non-verbal cognition was not measured at age 22, though general cognition effects become even stronger at that age (see previous section).

Improvements were also observed in nonverbal cognition for the Red de Protección Social for 10-year-old boys in Nicaragua.[6] However, no effects were found for girls. Siblings of children exposed to a deworming program in Kenya also demonstrated significant improvement in both nonverbal cognition and an overall cognitive factor score.[36] The authors posit that this could either be due to a direct nutritional effect from carrying a lower parasitic load, or it may also be that the better health status of children conferred upon the younger siblings of those affected by the intervention could have improved schooling outcomes such as attendance, which may also have contributed to improvements in cognition.

Finally, the Safe Motherhood Program (SMP) in Indonesia employed an army of nurses with one year of training to provide access to obstetric and child health care along with parent training on nutrition and health practices in rural villages.[12] Following children born into this program a decade later revealed a highly significant advantage of 0.11 standard deviations on the Raven cognition test for 10 to 14 year-olds compared to children born into villages that did not live in a beneficiary village or whose village did not receive a nurse until after the child was age four.

Executive Function

There is very little evidence of consistent effects on executive function across the range of intervention domains. Stimulation does not seem to improve executive function as much as other cognitive measures.

Executive function skills enable children to adapt to the changing environment around them, and are assessed by measuring cognitive processes such as working memory and problem solving. Measures of the executive function dimension of cognition were obtained in evaluations of nutrition, stimulation, and social protection interventions, although the evidence base is thin across all intervention types.

Two studies of nutritional programs reported effects on executive function measures, and neither the iron and zinc supplementation program in Thailand[38] nor the maternal supplementation program in The Gambia[1] noticeably improved executive function.

Only one of the three stimulation programs for which executive function was measured was found to improve outcomes, and the only program to measure outcomes at multiple stages of a child's life did not find any significant improvements. Stimulation for low birthweight infants in Jamaica resulted in improved short-term memory at age six. However, no measurable improvements in processing speed were observed among stunted Jamaican children at ages 11–12[51] or working memory at ages 17–18.[52] The Bucharest foster care program also did not measurably improve 8 year olds' executive function, as assessed by perceptual reasoning, working memory, processing speed,[17] and visual and spatial working memory metrics.[2]

Only a single study evaluated the effects of a CCT program on executive function. The assessment of Nicaragua's CCT program[6] found a significant improvement in processing speed among 10-year-old boys whose families were eligible for the cash transfer. No effects were observed among girls.

Cognition Summary

While the evidence base may not be sufficiently robust to be decisive, these findings suggest that stimulation programs may be more likely to be effective than supplementation alone in affecting both full-scale and abbreviated measures of IQ and of nonverbal cognitive ability as measured subsequent to the early childhood period. These findings are in line with the growing body of biological and environment design research that has shown that the relationship established between caregivers and children during stimulating interactions allows infants to tactilely and perceptually explore their environments, facilitating neuronal growth and the development of basic cognitive skills necessary for learning progression throughout adolescence (Ngure and others 2014; Grantham-McGregor and others 2014).

As the positive gains in cognitive development in stunted children in the Jamaica study suggest, stunted children may be poised to particularly benefit from stimulation programs (see box 2.1). Nutrition-deficient children tend to be less likely to interactively engage with their environment, inhibiting their ability to develop cognitive skills through tactile exploration and reciprocal relationships with caregivers. Thus an increasing focus has been given to combining necessary nutritional supplementation with stimulation interventions for stunted children (Grantham-McGregor and others 2014).

The ability of several cash transfer programs included in this review to produce sustained cognitive gains suggests that social assistance interventions, when received during the early childhood period, can result in effective investments in children's development. These investments in areas such as health, nutrition, or education may help to mitigate vulnerability during particularly sensitive periods of cognitive development, thus reducing the lasting negative impacts of toxic stress (Denboba and others 2014; Shonkoff and others 2012a).

Finally, the fact that both Guatemala's INCAP nutritional supplementation and Indonesia's Safe Motherhood healthcare access interventions were most successful when provided to very young children substantiates the view that sustained intervention throughout the first thousand days yields important, lasting cognitive benefits.

---

[1] Fluid intelligence refers to being able to reason and apply logic, often requiring pattern recognition or the understanding of the relationships between things. This is distinct from crystallized intelligence, which refers to the ability to use acquired knowledge and vocabulary.

[2] Executive function abilities have become popular to assess in recent years as they are believed to be related to schooling and a number of life outcomes. See Center on the Developing Child, Harvard University (http://developingchild.harvard.edu/key_concepts/executive_function) for more information and discussion.

[3] For example, although randomized evaluations of the Head Start program detected cognitive gains among program participants between ages three and four (HHS 2005), nearly all of the gains disappeared by the start of first grade (HHS 2010).[3] Similarly, in the Early Childhood Longitudinal Study, Magnuson, Ruhm, and Waldfogel (2007) used nonexperimental methods to tease out the effects of preschool programs, and found that improvements on test scores diminished after kindergarten.

[4] Nores and Barnett (2010) define long-term as beyond age seven years.

3. Language Development

The evidence suggests that early childhood development (ECD) interventions can have lasting effects on language, although the results were mixed within intervention types and outcome subcategories.

The capacity to communicate begins to develop very early in life and is indicated by babbling, responding (nonverbally) to words, gesturing, and speaking. In the United States and in other countries, it has been noted that during the second year of life (12–24 months), children experience a vocabulary explosion during which time the learning of new words increases exponentially. As children enter the preschool years, language capacities become more sophisticated, not only in vocabulary but also in terms of spoken grammar and sentence structure. Children also develop the capacity to identify letters and perhaps words. These skills are important for enabling children to read and do well in school.

Much of the focus on interventions to promote language development come from research highlighting that children growing up in verbally rich households show greater knowledge of words (Hart and Risley 1995), and this is related to future learning (Walker and others 1994). There is substantial evidence that children in poorer homes are talked to far less and hear many fewer complex words than their better-off peers. Thus, the gaps in language abilities by wealth are likely from, in part, less verbal stimulation. As a result, many parenting, stimulation, and education programs emphasize activities involving verbal responsiveness to infants and young children and reading, singing, and telling stories.

The verbal assessments included by studies in this review are measures of receptive (that is, how much is understood) and expressive (that is, how many words are said) vocabulary, such as the Peabody Perceptual Vocabulary Test; scales that test other types of verbal skills (for example, verbal similarities, comprehension); and achievement tests that focus specifically on reading abilities.

Table 3.1 maps the 18 unique studies included in this review that investigate language outcomes. The studies span 15 different projects and 11 intervention types across 13 countries. They are grouped according to the outcome category in which they report measurements; the specific intervention within each category is noted. These classifications align with the taxonomy presented in (figure 2 in the Introduction). While there were a range of outcomes reported, as with other outcome domains, the estimates were separated into three subcategories (verbal abilities, vocabulary, and reading, literacy, and preliteracy) to make the interpretation of results more comparable across studies.

Table 3.1. Impact Evaluations Investigating Language Development

	Study	Country (Project)	Average Age at Intervention (Years)	Average Length of Exposure (Years)[b]	Age at Evaluation (Years)	Evaluated Intervention	Reviewed Outcomes
Nutrition	Kramer and others 2008a [27]	Belarus (Promotion of Breastfeeding Intervention Trial [PROBIT])	0	1	6	breastfeeding promotion	reading ability*; verbal abilities (abbreviated test)**; vocabulary**
Nutrition	Pongcharoen 2010 [38]	Thailand (micronutrient supplementation to children)	0.5	0.5	9	micronutrients and fortified food for children (iron and zinc supplementation)	verbal abilities
Nutrition	Alderman and others 2014 [1]	Gambia, The (maternal supplementation)	in utero	0.5 (DR1)	16–22	vitamins, micronutrients, or fortified food for pregnant women (protein biscuits)	vocabulary (expressive and receptive)
Nutrition	Walker and others 2000 [51][a]	Jamaica (stimulation and supplementation to stunted	1.55	2	11–12	supplementary feeding	receptive vocabulary; verbal abilities; vocabulary
Nutrition	Walker and others 2005 [52][a]	Jamaica (stimulation and supplementation to stunted	1.55	2	17–18	supplementary feeding	reading abilities; receptive vocabulary; verbal abilities; verbal analogies
Nutrition	Walker and others 2011 [55][a]	Jamaica (stimulation and supplementation to stunted	1.55	2	22	supplementary feeding	reading abilities; verbal abilities
Nutrition	Maluccio and others 2009 [31]	Guatemala (INCAP supplementary feeding to children)	0	5.3	25–42	supplementary feeding	reading abilities**
Early Learning/Childcare	Martínez, Naudeau, and Pereira 2012 [35]	Mozambique (preschool)	3.45	1.5	5–9	quality early childhood and preprimary program	receptive vocabulary
Early Learning/Childcare	Walker and others 2010 [54]	Jamaica (stimulation to low birthweight Infants)	0	2	6	stimulation	reading abilities; receptive vocabulary; verbal abilities

	Fox and others 2011 [17]	Romania (Bucharest Early Intervention Project)	1.88	2.7	8	stimulation (foster care)	verbal abilities**
	Walker and others 2000 [51][a]	Jamaica (stimulation and supplementation to stunted	1.55	2	11–12	stimulation	receptive vocabulary; verbal abilities**; vocabulary**
	Walker and others 2005 [52][a]	Jamaica (stimulation and supplementation to stunted	1.55	2	17–18	stimulation	reading***; receptive vocabulary**; verbal abilities**; verbal analogies**
	Walker and others 2011 [55][a]	Jamaica (stimulation and supplementation to stunted	1.55	2	22	stimulation	reading abilities***; verbal abilities***
Health	Ozier 2013 [36]	Kenya (primary school deworming project)	0	1	8–15	deworming	receptive vocabulary*
Social Protection	Secretariat of Social Development 2008 [8]	Mexico (Progresa)	1.5	1.5 (DR2)	7–11	CCT	reading comprehension; verbal abilities (abbreviated test)**
	Manley, Fernald, and Gertler 2012 [32]	Mexico (Progresa)	1	1.5 (DR2)	8–10	CCT—conditionalities	verbal abilities (abbreviated test)
						CCT—cash	verbal abilities (abbreviated test)***
	Barham and others 2014 [6]	Nicaragua (Red de Protección Social)	in utero	3 (DR2)	10 (boys)	CCT	receptive vocabulary**
	Rackstraw 2014 [39]	Honduras (Programa de Asignación Familiar)	1.5	2	13–15	CCT	reading abilities*
	DSD, SASSA and UNICEF 2012 [14]	South Africa (Child Support Grant)	1	2.5	10	unconditional or targeted income support	reading abilities
	Spears and Lamba	India	0	continuous	6–8	adequate sanitation	reading*

Water	2013 [44]	(Total Sanitation Campaign)					
Other	Pathak and Macours 2013 [37][c]	India (women's political reservation)	pre-birth birth	3 (DR1)	8	governance (women's political reservation)	reading abilities (in utero)*; reading abilities (0–5 years); receptive vocabulary

Verbal Ability

The evidence of ECD programs that improve verbal abilities is mixed, and only the stimulation component of a single program produced consistent improvements over an individual's life.

Together with performance IQ, verbal abilities make up full-scale IQ scores. As with other cognitive measurements, they can be measured using an abbreviated or full test. The effects on verbal abilities from nutritional interventions are mixed. Thailand's iron and zinc supplementation program for infants did not produce observable gains at age 9 years,[38] and an evaluation of Jamaica's supplementation programs did not find significant differences at ages 17–18 or 22 years of age.[52, 55] However, children whose mothers participated in Belarus's breastfeeding promotion program (PROBIT) had significantly better scores on an abbreviated measure of verbal abilities at 6 years of age compared to children's whose mothers had not participated (see figure 4.1).[27]

Similar to the trend seen with cognitive outcomes, stimulation interventions were more successful than supplementation interventions in producing improvements in verbal abilities. Of the four studies of stimulation programs that reported verbal ability measurements, all had a significant effect.[1] Three of these studies evaluated the stimulation arms of the Jamaica intervention at different points in the child's life: between the ages of 11–12 years, 17–18 years, and 22 years old.[51, 52, 55] An assessment of the BEIP foster care program in Romania (see box 3.1) also observed improved verbal abilities among 8 year olds who had been placed in foster care as compared to their peers who remained in institutional care.[17]

Box 3.1. Fostering and Stimulation in Romania Improved Language and Other Outcomes

The Bucharest Early Intervention Project was designed to give orphaned and abandoned young children living in institutional care a chance to be exposed to the type of stimulating environment believed to improve many development outcomes. At an average age of 21 months (range of 9–33 months), 136 children were randomly assigned to either remain in the government facility or be placed in a foster care home. The institutional environment was characterized by structured, unresponsive routines and a low caregiver-to-child ratio (Windsor and others 2011). All of the foster homes received ongoing support from program social workers, including training on how to provide a caring, stimulating environment for the children (Nelson and others 2007).

The program officially ended when children were 54 months old. Children could change care environments at any time during or after the project, with the one provision that foster care children would not return to institutional care (though institutionalized children could be fostered or adopted, making the intent-to-treat estimates a lower bound of program effectiveness). Children's development was periodically assessed from entry into the program through the age of 12 years, producing a barrage of studies assessing their development across a wide range of outcomes, including many that are not typically seen in the ECD literature. This review identified 11 studies, reporting a total of 49 outcomes over four domains that measured post-early childhood effects and were of sufficient quality to be included.

Physical Development

During the first 12 months of their placement, foster care children grew significantly more than their institutionalized peers in height and weight, although there was no detectable difference in head circumference or WAZ. After 12 months, they had height and weight scores in normal range, but the institutionalized group did not measurably improve (Johnson and others 2010).

However, physical gains appear to have ended after 12 months. No improvements were observed between 12–18 months (Johnson and others 2010), and none of the 19 physical outcomes in the post-early childhood period reported by studies in this review were significantly better for the foster child group.

Cognition Development

Improvement in cognitive abilities among children placed in foster care became apparent at an early age, and some gains in IQ were still observed in the post-early childhood period. When evaluated at 42 and 54 months, children in foster care performed significantly better on tests of cognitive ability. By age 8 years, these cognitive gains had translated into significantly higher full-scale IQ scores,[17] although there were no significant benefits in multiple measures of executive function.[11, 17]

Language Development

Although both groups learned language over time, the foster care children learned significantly more. By 30 months of age, foster care children showed some early gains in receptive language, but not in expressive language or their overall development quotients (DQs). By 42 months, however, they demonstrated significant improvements in both expressive and receptive language scores as measured by the developmental language scales (RDLS), though their DQs were still not measurably different (Windsor and others 2011).

By 8 years of age, all five of the unique language outcomes reported by studies in this review were significantly better in the foster care (stimulated) group.[17, 56] One of these measures—verbal comprehension—is a traditionally-reported language outcome in the ECD literature with a standardized measurement construct.

Socioemotional Development

Advantages in reduced socioemotional challenges benefits phase in and out through early childhood. Despite the absence of observable differences between the foster care and institutionalized children on multiple measures of socioemotional development at baseline, by 42 months of age foster care children were significantly better at paying attention and exhibited better positive affect (Ghera and others 2009). Twelve months later, foster care children were significantly less likely to have internalizing disorders such as anxiety, but there were no observable differences in externalizing disorders such as ADHD (Zeanah and others, 2009).

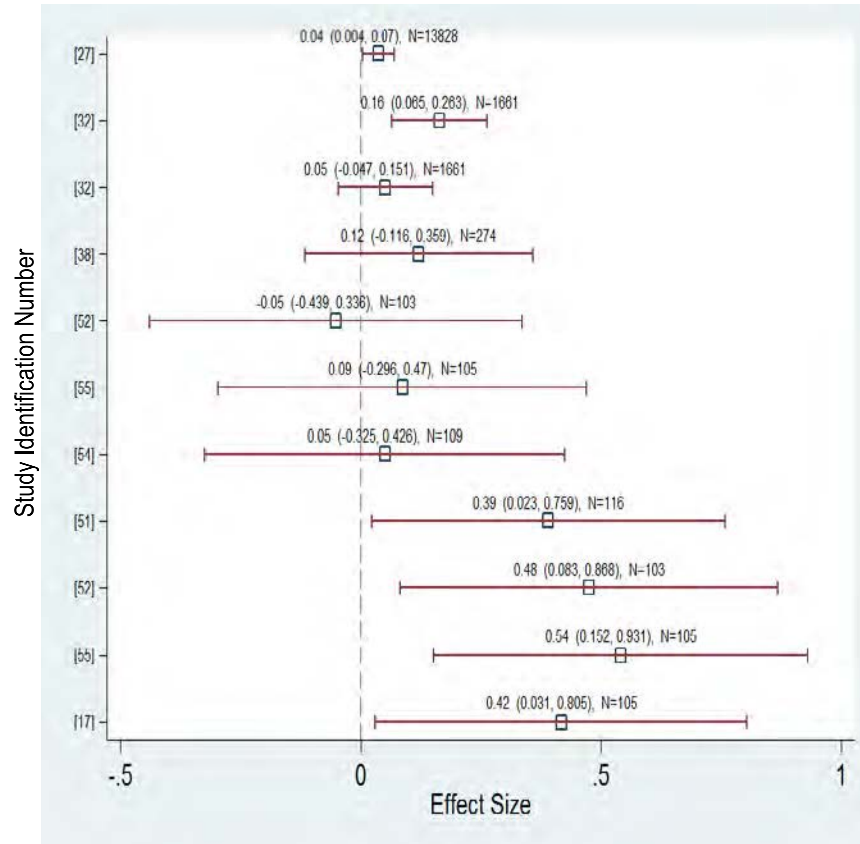
Improvements in socioemotional development solidified during early adolescence when children were 10 and 12 years old. The studies in this review reported 17 unique measurements of socioemotional outcomes at these ages, 11 of which were significant, including higher scores on the Social Communication Questionnaire at 10 years and decreased externalizing behavior[23] at 12 years.[30]

An additional 18 months of exposure to Mexico's Progresa (see box 5.1), which occurred while the children were still eligible for nutritional supplementation and were required to attend additional health checkups, significantly increased abbreviated verbal abilities scores among 7- to 11-year-old children.[8] However, a study that disentangled the effects of abiding by the conditionalities versus receiving the cash transfer found that only the latter component had a significant impact on abbreviated verbal abilities scores among children 8–10 years old.[16, 32]

Figure 3.1. Forest Plot for Verbal Abilities

Country	Intervention	Measurement	Average Age at Evaluation (years)
Belarus	Breastfeeding	WASI	6.5
Mexico[a]	CCT cash	WASI	9
Mexico[a]	CCT conditionality	WASI	9
Thailand	Nutrition (child, micronutrient)	WISC-III	9
Jamaica[a]	Nutrition (stunted children, SF)	WAIS	17.5
Jamaica[a]	Nutrition (stunted children, SF)	WAIS	22
Jamaica	Stimulation (LBW children)	WPPSI	6
Jamaica[b]	Stimulation (stunted children)	WISC-R	11.5
Jamaica[b]	Stimulation (stunted children)	WAIS	17.5
Jamaica[a]	Stimulation (stunted children)	WAIS	22
Romania	Stimulation (foster care)	WISC-IV	8

Study and Forest Plot



Note: The forest plot shows standard mean difference, 95 percent lower and upper bound confidence interval in parentheses, and sample size (N=number). The standard mean difference and confidence interval were calculated by Comprehensive Meta-Analysis software. Bracketed numbers correspond to numbered studies in References. All the results come from the verbal abilities score calculated within each test (WASI, WISC, WAIS, WPPSI). CCT = conditional cash transfer; WASI = Wechsler abbreviated scale of intelligence; WISC = Wechsler intelligence scale for children; WAIS = Wechsler adult intelligence scale; WPPSI = Wechsler preschool and primary scale of intelligence. LBW = Low birthweight. SF = Supplementary feeding.

a. Mexico study [8] is not included in this forest plot given that it looks at same average age at evaluation with [32].

b. For Jamaica studies [51, 52, 55] on stunted children, the stimulation compares “stimulation only + stimulation and supplementation” versus “no intervention + supplementation only.” Similarly, the supplementation compares “supplementation + supplementation + stimulation” versus “no intervention + stimulation.”

Reading and Literacy

Reading and literacy outcomes were improved by stimulation, sanitation, and governance programs, although there are only single studies of each of these intervention types. Nutritional interventions had mixed results, and social protection programs did not have a measurable effect on reading and literacy

Tests of reading abilities (for example, word reading, sentence completion, and context comprehension) are also likely to play a large role in determining educational achievements. However, supplementation did not consistently improve reading outcomes. Two studies that measured the effects of nutritional interventions on reading outcomes evaluated the Jamaica supplementation program and found that between the ages of 17 and 18 years old,[52] and again at age 22,[55] children who received supplementation did not have noticeably improved reading abilities compared to those who had not been given supplementation.

However, an assessment of the breastfeeding promotion program in Belarus found a marginally significant improvement in reading among 6.5-year-old children whose mothers had been encouraged to breastfeed (see figure 3.2).[27] In Guatemala, children who had participated in the supplementary feeding program of the Instituto de Nutrición de Centroamérica y Panamá demonstrated higher levels of reading comprehension between the ages of 25 and 42 years old.[31] It is difficult to determine, however, whether the positive effects found in Belarus and Guatemala are due to the specific nature of the intervention or the ages at which children received the intervention.

Stimulation also improved reading and literacy. The same two studies that reported the effects of the supplementation only arm in Jamaica also evaluated the stimulation component and found highly significant advantages in reading levels at both 17–18 years of age and 22 years of age.[52, 55] Interestingly, these studies allow for a direct comparison of the effects of supplementation versus stimulation on reading outcomes for the same participant group at the same ages, and suggest that stimulation is better able to improve post-early childhood reading outcomes (see figure 3.2). However, it must be noted that these findings are only applicable to a single program that was not implemented at scale and cannot be reliably generalized across contexts until further research has been done.

Three studies evaluated the effects of conditional cash transfers on reading comprehension. An assessment of the effects of an additional 18 months of exposure to Progresa in Mexico did not observe significant improvements when children were tested between the ages of 7 and 11 years.[8] Similarly, neither an unconditional cash transfer and income support program in South African nor a CCT in Honduras improved reading comprehension[14] or literacy[39] among 10-year-olds.

A single study was also found that assessed the effects of a sanitation intervention on language outcomes. In India, children living in areas where the Total Sanitation Campaign was implemented during the first year of their lives were exposed to more latrines than children born either in the same district in a different year or in a different district in the same year. At ages 6, 7, and 8 years, these children had marginally significantly higher literacy scores.[44]

One study assessed the impact of a political intervention in India on children's development outcomes. In Andhra Pradesh, local political seats were randomly reserved for women over several election cycles. The study measured learning outcomes among children who were exposed to the political reservation cycles during different periods of early childhood. The first cycle of reservations occurred five years prior to the birth of the first cohort of children and ended soon after their birth. Before being evaluated at age 8 years, the first cohort had thus received five years of the exposure to the treatment period and 8 years of exposure to post-treatment period in which improvements in norms and attitudes towards female political leadership and the issues they espoused may have persisted. A second cohort consisted of children who were born at the beginning of the next reservation cycle. These children lived in a treatment period for the first five years of their lives, but only had three years of exposure to the post-treatment period before being assessed at age eight years.

The two treatment cohorts were compared to a cohort of children who were six years old before seats were allocated for women in their districts. Although they experienced two years of exposure to the treatment period, this occurred after their entry to primary school, and they did not have any exposure to the post-treatment environment. Although no improvements in reading were observed among the second cohort, the first cohort of children did have significantly higher scores on an early grade reading assessment at age 8 years.[37] These results suggest that the long exposure to the post-treatment environment experienced by the first cohort, during which time some of the policies and norms established during the period when seats were allocated for women may have continued and so may have played a crucial role in affecting language outcomes beyond the early childhood period.

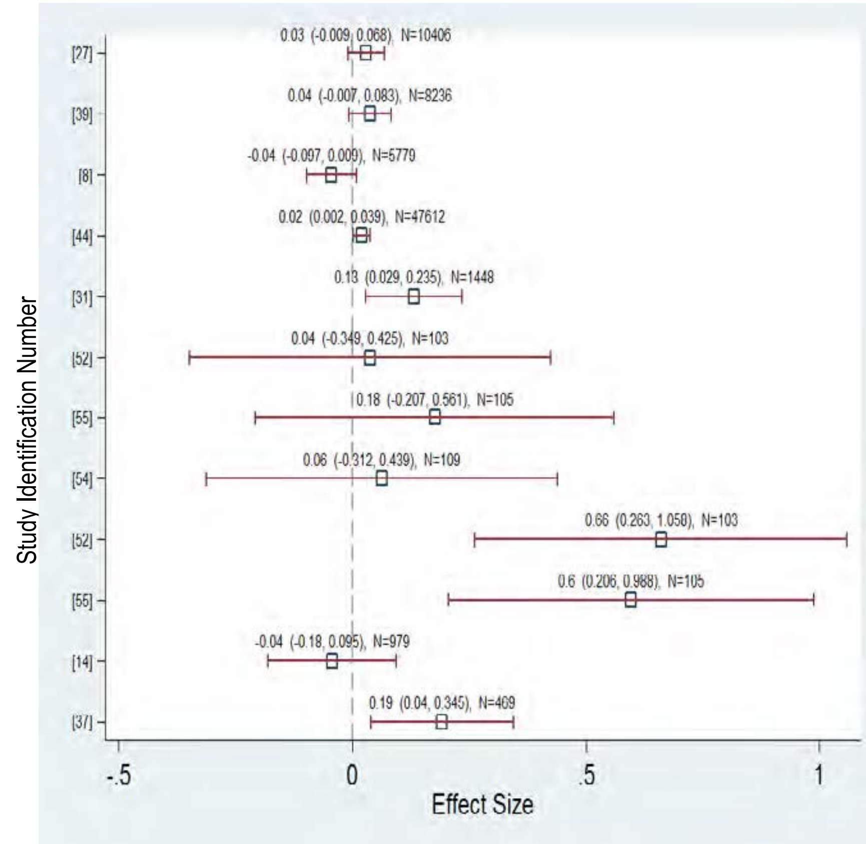
---

[1] Study [54], which provided stimulation to low birthweight Jamaican children, reported results for "verbal abilities," for six year olds, but no significant effects were observed.

Figure 3.2. Forest Plot for Reading and Literacy

Country	Intervention	Measurement	Average Age at Evaluation (years)
Belarus	Breastfeeding	Reading ability (TRAP)	9
Honduras	CCT	Literacy	14
Mexico	CCT	Ability to read	9
India	Sanitation	Recognizing letters	6
Guatemala	Nutrition (child, SF)	IARCT	32.3
Jamaica[a]	Nutrition (stunted children, SF)	Group reading test (RCC)	17.5
Jamaica[a]	Nutrition (stunted children, SF)	WRAT (Reading (log))	22
Jamaica	Stimulation (LBW children)	ERA (sqrt)	6
Jamaica[a]	Stimulation (stunted children)	Group reading test (RCC)	17.5
Jamaica[a]	Stimulation (stunted children)	WRAT (Reading (log))	22
South Africa[b]	UCT	Ability to read a story	10
India	Women's political reservation	EGRA	8

Study and Forest Plot



Note: The forest plot describes standard mean difference, 95 percent lower and upper bound confidence interval in parentheses, and sample size (N = number). The standard mean difference and confidence interval were calculated by Comprehensive Meta-Analysis software. Bracketed numbers correspond to the numbered studies in References. CCT = conditional cash transfer; UCT = unconditional cash transfer. TRAP = Teacher Ratings of Academic Performance. SF = Supplementary feeding. LBW = Low birthweight. IARCT = Inter-American reading and comprehension test. RCC = Reading context comprehension. WRAT = Wide Range Achievement Test. ERA = Early reading assessment. EGRA = Early grade reading assessment.

a. For Jamaica studies [52, 55] on stunted children, the stimulation compares "stimulation only + stimulation and supplementation" versus "no intervention + supplementation only." Similarly, the supplementation compares "supplementation + supplementation + stimulation" versus "no intervention + stimulation."

b. For South Africa study [14], the outcome is measured through the Early Grade Reading Assessment in which a child has to do a timed reading of letters and familiar words.

Vocabulary

The evidence of ECD interventions' effects on vocabulary capabilities is too thin and mixed to inform a general conclusion.

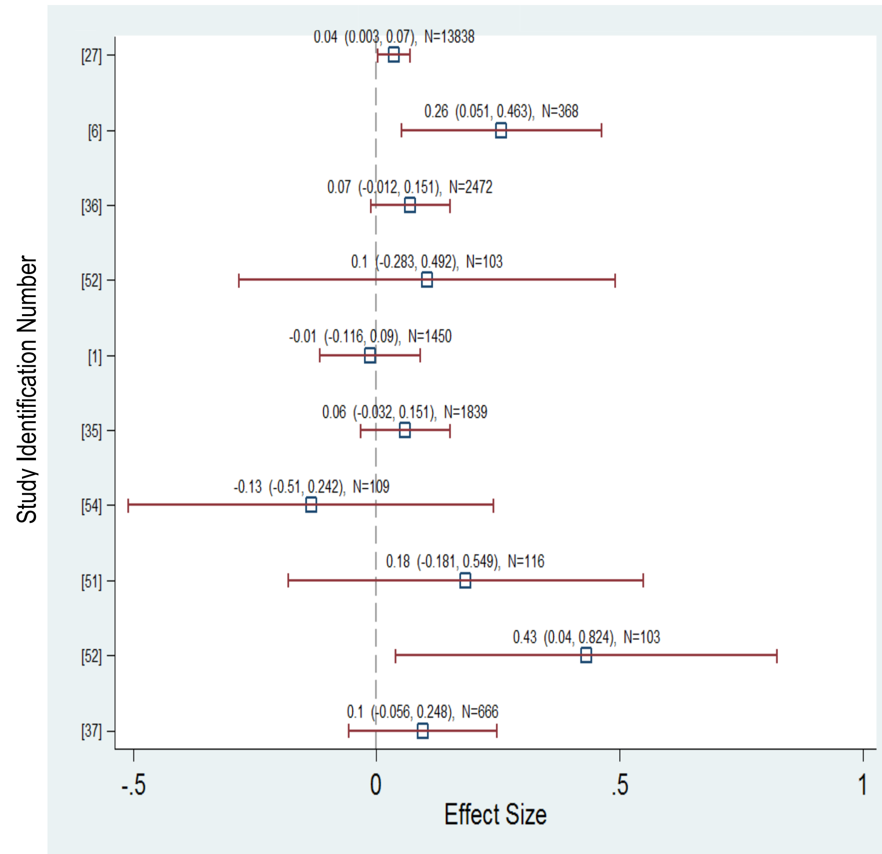
The evidence is inconsistent for measurements of vocabulary. Among nutritional interventions, the breastfeeding promotion program in Belarus resulted in improved reading among 6.5 year olds,[27] but The Gambia nutritional intervention did not produce detectable improvements in vocabulary between the ages of 16 and 22 years (see figure 3.3).[1]

Programs promoting stimulation also had mixed effects. The preschool and mother training program in Mozambique did not noticeably improve vocabulary scores among participating children between the ages of five and nine years old,[35] while the Jamaica program that provided stimulation to stunted children did significantly improve vocabulary abilities at ages 11–12 years and age 17–18 years.[51, 52] The lack of density of evidence for these intervention types makes it difficult to form broad conclusions about what type of programs can successfully and consistently produce long-term improvements in vocabulary abilities.

Three intervention types' effect on vocabulary outcomes are represented by a single study each: social protection programs, disease prevention treatments, and governance interventions. An evaluation of Nicaragua's CCT program found significant improvements in vocabulary among 10-year-old boys whose families had been eligible for the program.[6] Siblings of children who had been part of a school deworming program had only a marginally significant improvement in vocabulary abilities at age 15 years.[36] In India, the random allocation of political seats to women did not produce measurable improvements in vocabulary for a cohort of children for which seat reservations occurred for five years prior to birth, or a cohort for which seats were allocated during the first five years of their lives, when compared to children who were not exposed to a reservation period until age six years.[37] Again, the absence of a dense evidence base for each of these intervention types limits the ability to draw reliable conclusions from these findings.

Figure 3.3. Forest Plot for Vocabulary

Country	Intervention	Measurement	Average Age at Evaluation (years)	Study and Forest Plot
Belarus	Breastfeeding promotion	WASI	6.5	
Nicaragua[a]	CCT	TVIP[b] (boy)	10	
Kenya	Deworming	PPVT	10	
Jamaica[c]	Nutrition (stunted children, SF)	PPVT	17.5	
Gambia, The	Nutrition (mother)	MHVT	19.6	
Mozambique	QEC and PPP	TVIP[b]	7	
Jamaica	Stimulation (LBW children)	PPVT	6	
Jamaica[c]	Stimulation (stunted children)	PPVT	11.5	
Jamaica[c]	Stimulation (stunted children)	PPVT	17.5	
India	Women's political reservation	RV	8	



Note: The forest plot describes standard mean difference, 95 percent lower and upper bound confidence interval in parentheses, and sample size (N = number). The standard mean difference and confidence interval were calculated by Comprehensive Meta-Analysis software. Study numbers in brackets correspond to the numbered studies in References. CCT = conditional cash transfer; PPVT = Peabody Picture Vocabulary Test; TVIP = Test de Vocabulario en Imagenes Peabody; WASI = Wechsler abbreviated scale of intelligence. SF = supplementary feeding. LBW = Low birthweight. MHVT = Mill Hill Vocabulary Test. RV = Receptive vocabulary. QEC = Quality early childhood. PPP = Pre-primary program.

a. Nicaragua study [6] only reports boy's outcome.

b. TVIP is a Spanish version of PPVT.

c. For Jamaica studies on stunted children, the stimulation compares "stimulation only + stimulation and supplementation" versus "no intervention + supplementation only." Similarly, the supplementation compares "supplementation + supplementation + stimulation" versus "no intervention + stimulation."

Language Summary

As with cognition, the evidence suggests that stimulation is better able to improve post-early childhood language outcomes than are micronutrient or macronutrient supplementation programs. This is in line with research that shows the quality of parent-child interactions (for example, how parents speak to and respond to infants and young children) is an important predictor of language development. The human brain develops in time-sensitive periods during which interaction with an individual's environment facilitates neuronal connections that are important for the development of cultural skills. Stimulation programs play a particularly important role in facilitating this process by encouraging caregivers to interact with children through a reciprocal relationship that establishes verbal and nonverbal communication skills (Black and Dewey 2014; Grantham-McGregor and others 2014; Wachs and others 2014).

There is solid rationale for expecting that various types of nutritional interventions to children at-risk for or suffering from deficiencies such as chronic malnutrition and iron deficiency could positively impact cognitive and language outcomes. However, the available evidence suggests that supplementation alone may not be enough to produce sustained effects on these abilities.

4. Socioemotional Development

Improvements in externalizing behavior do seem to manifest as participants age, and both conditional cash transfers (CCTs) and early stimulation may improve abbreviated measures of children's behavioral problems, but it appears harder to create a sustained change in internalizing behavior.

Socioemotional functioning refers to a broad range of intra- and interpersonal competencies. Intrapersonal competencies include, for example, how individuals view themselves, how they manage their feelings, how they approach and respond to problems or difficulties, and what their capacities are for self-motivation, perseverance, and attention. Interpersonal skills are concerned with how individuals establish and maintain relationships and how they interact with others. In the early years, the primary socioemotional tasks include forming loving and trusting relationships with adults, becoming independent, recognizing and learning how to control emotions and impulses, learning to read others' emotions, and developing empathy. Many of these skills are important for starting school ready to learn, and are believed to be necessary for succeeding in school, and subsequently, other aspects of life. Good early socioemotional development is also important in that it contributes to a person's ability to cope with anxiety and depression, maintain good self-esteem, and become self-sufficient. For this reason, socioemotional outcomes are important objectives of many early childhood development (ECD) interventions.

Children growing up in poor or adverse environments may be at increased risk for developing socioemotional responses or behaviors that can negatively impact not only their current functioning (Evans and others 2005; Kalil, Yoshikawa, and Ziol-Guest 2014), but also their future mental and physical well-being (Pechtel and Pizzagalli 2011; Shonkoff and others 2012a; Schilling, Aseltine and Gore 2007; Slopen, Koenen, and Kubzansky 2014). Prolonged exposure to abuse, neglect, violence, or other adversity can trigger a "toxic stress response" in a child, which can interfere with brain development and health organ functions, leading to stress-related diseases and cognitive impairment that can persist into adulthood (Shonkoff and others 2012a).

Longitudinal studies in developed countries have found low-income status associated with increases in externalizing[1] and internalizing[2] behaviors across childhood and adolescence (Bradley and Corwyn 2002; Brooks-Gunn and Duncan 1997; Evans, Li, and Whipple 2013), and there is evidence that these persist into adulthood (for example, Schilling, Aseltine and Gore 2007). There is also strong indication from developed nations that the effects of adversity early in life are apparent by the preschool years. For example, two large recent studies in the United States and United Kingdom found four-year-old children from the poorest homes showed more hyperactivity and problems with attention than did children from the wealthiest homes (Waldfogel and Washbrook 2011). In another national cohort study, English children who were raised in homes with high levels of cumulative adversity—including financial stress, maternal psychopathology, and neighborhood disadvantage—had more behavioral problems than children living in homes facing relatively little adversity (Slopen, Koenen, and Kubzansky 2014).

There is consensus in the literature from high-income countries that ECD programs can result in long-term improvements in an individual's social competence and psychological functioning (Ramey and Ramey 1998; Schweinhart, Barnes, and Weikart 1993; Schweinhart and others 2005; Gorey 2001; Camilli and others 2010; Yoshikawa 1994; Zigler and Styfco 1994; Zigler, Taussig, and Black 1992). For example, both in the short-term and up to 27 years after participating in preschool programs in the United States that were designed to improve cognitive and socioemotional functioning, participants experienced increases in social competence, defined as how well a child is able to interact in social situations (Anderson and others 2003).

Another robust effect found in the developed world is a decrease in behaviors such as crime. Evaluation of the Perry Preschool Project indicates fewer arrests among beneficiaries when followed up at age 40 years (Scweinhart and others 2005). Reynolds and others (2001) found that participants in Chicago's Child-Parent Centers, a program to help engage parents from low-income households in their children's preschool education, had lower arrest rates as young adults, while Donohue and Siegelman (1998) estimated that ECD programs can pay for themselves through reduced crime rates alone.

In low- and middle-income countries, one finds similar trends in regard to the association between adverse environments and poor socioemotional development (Kessler and others 2010). Reviews suggest that young children exposed to poverty and associated risks (that is, undernutrition, infectious disease, and insensitive parenting) exhibit more problematic behaviors (Walker and others 2007, 2011), and that these may continue into adulthood if there is no intervention to change their trajectory (Grantham-McGregor and others 2007).

As in high-income countries, however, evidence shows that a change in socioemotional outcomes is possible (Nores and Barnett 2010; Baker-Henningham and Lopez Boo 2010). For example, trials of parenting interventions to promote early stimulation through parent-child interactions have shown improvements in children's socioemotional outcomes (Engle and others 2011; Walker and others 2007; Walker and others 2011). Others have theorized that adequate early nutrition can also lead to socioemotional benefits (Black and others 2013; Engle and others 2007; Walker and others 2007).

A range of measures is used to estimate the effects of interventions on socioemotional outcomes, which are indicated in table 4.1. For younger children, many of these results are reported by parents or teachers, while measurements for older children and adults are often self-reported. The table details the 10 studies[3] in six countries that investigate socioemotional outcomes. There are seven different projects and six intervention types, identifying the specific indicators of both internalizing and externalizing behavior and the age of the individual at the time of evaluation.

Evaluations of programs in four different countries included estimates of externalizing behavior and distractibility during post-early childhood years. The effects from four unique interventions are shown in (figure 4.1).

Table 4.1. Impact Evaluations Investigating Socioemotional Development

	Study	Country (Project)	Average Age at Intervention (Years)	Average Length of Exposure (Years)[b]	Age at Evaluation (Years)	Evaluated Intervention	Reviewed Outcomes
Nutrition	Kramer and others 2008b [28]	Belarus (Promotion of Breastfeeding Intervention Trial [PROBIT])	0	1	6	breastfeeding promotion	strength and difficulties questionnaire (total difficulties)
Nutrition	Pongcharoen 2010 [38]	Thailand (micronutrient supplementation to children)	0.5	0.5	9	micronutrients and fortified food for children	Freedom from Distractibility Index
Nutrition	Walker and others 2011 [55][a]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	22	supplementary feeding	anxiety; depression; involved in a physical fight; involved in a violent crime; social inhibition; weapon use
Early Learning/Childcare	Walker and others 2010 [54]	Jamaica (stimulation and supplementation to stunted children)	0	2	6	stimulation	attention (map search, opposite-same (switching)); strength and difficulties questionnaire (total difficulties)**
Early Learning/Childcare	Humphreys and others 2015 [23]	Romania (Bucharest Early Intervention Project)	1.88	2.7	12	stimulation (foster care)	externalizing behavior**; hyperactivity; internalizing behavior
Early Learning/Childcare	Walker and others 2006 [53]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	17–18	stimulation	anxiety***; attention deficit**; depression**; hyperactivity; oppositional behavior*; self-esteem**
Early Learning/Childcare	Raine and others 2003 [41]	Mauritius (Child Health Project)	3	2	17	quality early childhood and preprimary program	anxiety; attention problem; hyperactivity**
Early Learning/Childcare	Gertler and others 2013 [18]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	22	stimulation	externalizing behavior factor; internalizing behavior factor**
Early Learning/Childcare	Walker and others 2011 [55][a]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	22	stimulation	anxiety; depression**; involved in a physical fight*; involved in a violent crime**; social inhibition*; weapon use

	Raine and others 2003 [41]	Mauritius (Child Health Project)	3	2	23	quality early childhood and preprimary program	court-reported criminal offenders*; self-reported criminal offenders**
Social Protection	Behrman and others 2008 [8]	Mexico (Progresa)	1.5	1.5 (DR)	7–11	CCT	strength and difficulties questionnaire**
Social Protection	Manley, Fernald, and Gertler 2012 [32]	Mexico (Progresa)	1	1.5 (DR)	8–10	CCT—conditionalities	strength and difficulties questionnaire**
						CCT—cash	strength and difficulties questionnaire

Note: Bracketed numbers correspond to numbered studies found in References. More details for each study are found in appendix A. CCT = conditional cash transfer; DR = dose response; INCAP = Instituto de Nutrición de Centroamérica y Panamá.

a. Jamaica [55] has a multiple intervention arm, and each intervention type has a separate row for this study.

b. DR in the length of exposure means intervention period in terms of the Dose Response. Specifically, DR here describes the dose response where early and late treatment effect is compared, and length of exposure is the difference of the intervention period between treatment and control group.
* Statistically significant at 10 percent.
** Statistically significant at 5 percent.
*** Statistically significant at 1 percent.

Box 4.1. Early Education and Health Enrichment in Mauritius Results in Improved Socioemotional Outcomes through Early Adulthood

The Mauritius Child Health Project, a two-year nursery school program for children between the ages of three to five, was started in 1972 based on recommendations by the World Health Organization (WHO) that countries focus on interventions that addressed malnutrition and poverty during the early childhood years. Although there are no estimates of effects during early childhood, beneficiaries were evaluated at three later points in time. One hundred three-year-old children, chosen from a population of 1,795, who were experimentally matched to a group of 100 comparison children based on similarity in psychophysiological baseline measures, gender, and ethnicity were evaluated at age 11.[40] Evaluations of the participants at ages 17 and 22 followed-up on 83 out of 100 treatment participants, and 355 individuals from the original population were matched on 10 variables (ethnicity, gender, age, nutritional status, cognitive ability, temperament, autonomic reactivity, parental social class, social adversity, and mother's age at birth) to construct a comparison group.

The intervention included preschool education, nutritional meals, educational programs, physical exercise, health assessments, and remediation for behavioral problems and learning disorders. Parental involvement was encouraged and home visits were also conducted. The children in the comparison groups also attended traditional community "petite écoles," but the teacher-pupil ratio was lower for students in the program (1:5.5 in the treatment group compared to 1:30 in the comparison group). Additionally, the food was of better quality and the health and educational curriculum better organized in the treatment group.[41]

Evaluations of the Child Health Project were the first to examine the effect of early education and health enrichment on psychophysiological functioning. The studies focused on measuring biological indicators of stress such as the level of a child's skin conductance, known as electrodermal activity, which reflect an individual's level of emotional stimulation and ability to pay attention and process information. These psychophysiological indicators were studied to understand the development of behavioral problems and antisocial behavior during early childhood. When children who were enrolled in the nursery program were tested between six and eight years after the intervention, at age 11, they showed signs of increased psychophysiological functioning—they were better able to process information and were more cognitively aroused than children of the same age who had not been exposed to the environmental enrichment program.[40]

Other studies were conducted to assess the effects of the program on adult outcomes of schizophrenia, conduct disorder, and criminal behavior. Results obtained through self-reported questionnaires and the Revised Behavior Problem Checklist suggested that at age 17 years participants showed fewer of schizophrenia and antisocial behavior, and at age 23 years were less likely to engage in criminal behavior. Interestingly, these positive outcomes were more pronounced among children who had been malnourished at age three years,[41] suggesting a catch-up effect.

The table below presents all of the outcomes reported by evaluations of the Mauritius Child Health Project. The lack of evaluations in five of the six outcome domains suggest that a significant research gap exists in the full understanding of the effects of early education and health enrichment on children's holistic development.

Outcome Domain by Age at Evaluation in the Mauritius Child Health Project

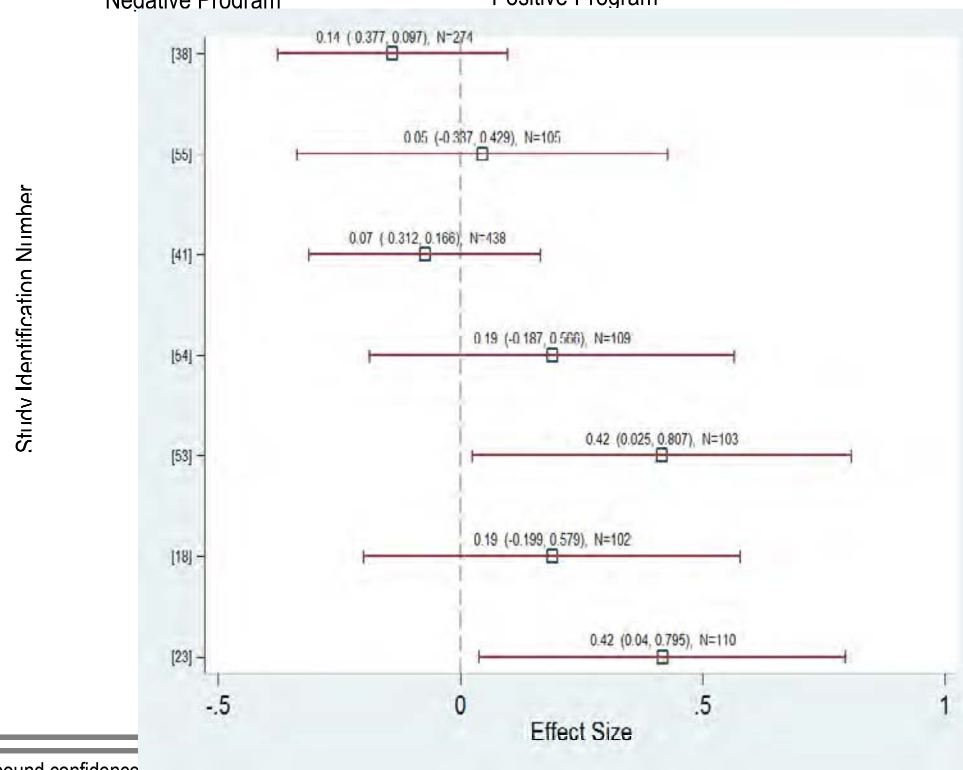
Age	Study	Physical Development	Cognitive Development	Language Development	Socioemotional Development	Schooling Outcomes	Employment and Labor Market Outcomes
11	[40]	4/7	-	-	-	-	-
17	[41]	-	-	-	5/8	-	-
23	[41]	-	-	-	1/6	-	-

Notes: To provide a more complete scope of these studies, this table includes all reported outcomes and not just those analyzed in the main body of this report (see box 1 in the Introduction for the decision rule for selecting outcomes for analysis). The numerator denotes statistically significant outcome at 10 percent level or better, and the denominator is the number of outcomes in the domain. The [bracketed] superscript number in the "Study" column is indicates study identifier (see References). No studies were found that produced causal estimates of the program during early childhood.

Figure 4.1. Forest Plot for Socioemotional Outcomes

Countr	Intervention	Measurement	Average Age at Evaluation (years)	Study and Forest
Thailand	Nutrition (child micronutrient)	Freedom from DI[a]	9	
Jamaica	Nutrition (stunted children, SF)	Social inhibition[b]	22	
Mauritius	QEC and PPP	Attention	17	
Jamaica	Stimulation (LBW	Attention (map search)[d]	6	
Jamaica[e]	Stimulation (stunted children)	Attention deficit[f]	17.5	
Jamaica[e]	Stimulation (stunted children)	Externalizing BF[g]	22	
Romania	Stimulation (foster care)	Externalizing symptom	12	

Negative Program | Positive Program



Note: The forest plot describes standard mean difference and 95 percent lower and upper bound confidence [...] size (N=number). Bracketed numbers correspond to numbered studies in References. Externalizing behavior scores in [28] were measured through supplemental behavior questions from the Canadian National Longitudinal Survey of Children and Youth, assessed by teachers but not included because details are lacking. Argentina's quality early childhood and preprimary program [9] is excluded because it measures attention using teacher assessment questionnaires ("How many of your students pay attention in class?") in administrative records of the National Education Ministry (i.e., National Education Assessment Operation). DI = Distractibility index. SF = supplementary feeding. LBW = Low birthweight. QEC = Quality early childhood. PPP = Pre-primary program. BF = Behavior factor.

a. Freedom from distractibility index in [38] is derived from arithmetic and digit span subtests.
b. Social inhibition in [55] is measured through 3 subscales from the inventory on Interpersonal Problems.
c. Attention problems in [41] is measured through Revised Behavior Problem Checklist, which is not comparable with other socioemotional outcomes.
d. Attention (Map Search) in [54] is measured through Test of Everyday Attention for Children.
e. For Jamaica studies [18, 53] on stunted children, the stimulation compares "stimulation only + stimulation and supplementation" versus "no intervention + supplementation only." Similarly, the supplementation compares "supplementation + supplementation + stimulation" versus "no intervention + stimulation."
f. Attention deficit in [53] is measured through Conners' Parent Rating Scale (short form).
g. Externalizing behavior factor in [18] is from the factor analysis of the WRAT math, WRAT reading comprehension, PPVT, verbal analogies, Raven matrices, and WAIS full scale intelligence tests.

Externalizing Behavior

Early childhood interventions appear to have a delayed effect on externalizing behavior, with no detectable effect on children but some improvements in teenagers and young adults.

Generally it appears that early childhood interventions have not demonstrated an effect on school-aged children's externalizing behavior (see figure 4.1). For example, the early stimulation intervention given to low birthweight infants in Jamaica had no detectable effect on a six-year-old's ability to pay attention.[1, 54] (Although similar to the psychosocial stimulation and nutritional supplementation interventions in Jamaica, this intervention was separate and started about a decade after the other.) In Thailand, a Freedom from Distractibility Index[2] was calculated for nine year olds who received iron or zinc supplements or both when they were infants. There was no discernable difference between the freedom from distractibility index scores of children who received the micronutrients and those who did not.[38]

Some evidence shows externalizing behavior could change in the teenage years and beyond as a result of interventions that occurred during the early childhood period. At 12 years old, Romanian children who were randomly assigned to foster care had significantly better externalizing behavior (comprised of oppositional defiant and conduct disorders) than did those who were raised in institutional homes. The different living environment did not influence hyperactivity rates, however.[23] At age 17, participants in the Child Health Project in Mauritius had significantly lower rates of motor excess, a measure of hyperactivity; there was no significant effect on attention problems.[41] Participants in the psychosocial stimulation intervention in Jamaica scored on average 3.44 points lower (out of a possible 36 points) at 17–18 years of age on the attention deficit questionnaire than did nonparticipants,[53] indicating better attention abilities. No significant difference was found in oppositional behavior or hyperactivity between the two groups.

A separate study looked at these same three external behaviors (hyperactivity, attention deficit, and oppositional behavior) when the participants were 22 years old, combining them into a factor score. Higher values indicate better functioning. The study found that stimulation had a positive, but ultimately nonsignificant, effect on reducing externalizing behavior for participants as a whole. Among women, however, it led to a significant 0.58 standard deviation improvement in their externalizing behavior factor score.[3, 18]

These same two programs—psychosocial stimulation in Jamaica and environmental enrichment in Mauritius—were also evaluated for their effect on violent behavior. In both cases there is evidence of benefits stemming from the interventions. At 23 years old, the number of self-reported criminal offenders in Mauritius was significantly lower among participants—23.6 percent versus 36.1 percent among the control group. There was also a marginally significant difference in court-reported criminal offenders, with only 3.6 percent of the enrichment group being reported as opposed to 9.9 percent of the control group.[41] In Jamaica, 22 year olds who participated in the psychosocial stimulation program were 36 percent as likely to be involved in or start a physical fight—although these results were only marginally significant—and 33 percent as likely to be involved in a violent crime.[4], [55] There was no effect, however, on weapon use.[5]

Internalizing Behavior

The evidence on post-early childhood internalizing behavior is too thin and inconsistent to draw conclusions.

Early stimulation in Jamaica had a sustained effect on internalizing behavior, but the Jamaican supplementation, Mauritian early enrichment, and Romanian foster care programs did not cause any detectable changes.[18, 23, 41, 53, 55] Jamaican 17 and 18 year olds who received early stimulation exhibited lower levels of anxiety and depression and greater self-esteem. Their anxiety score decreased by 2.81 points (out of a possible 28 points), their depression score decreased by 0.43 points (out of a possible 26 points), and their self-esteem score rose by 1.55 points (out of a possible 32 points).[53]

These improvements continued into adulthood. At age 20, a factor score[6] comprised of these three components showed a strong, positive effect overall—a 0.39 standard deviation increase—driven by improvements among women: the women's score improved by a highly significant 0.76 standard deviations, while the effect on men's scores was not significant.[18] Two of the same outcomes were reexamined two years later when participants were 22 years old, and while the effect on anxiety had disappeared, there was still a significant 0.35-point decrease on the depression scale.[55]

There was no detectable effect on anxiety among 17 year olds who participated in the Mauritius early enrichment program, nor among 22 year olds who had received the nutrition portion of the Jamaican intervention.[41, 55]

Strengths and Difficulties Questionnaire

When combining measures of externalizing and internalizing behavior, it appears that conditional cash transfers and early stimulation may be effective.

Four studies used the Strength and Difficulties Questionnaire (SDQ), a brief screening test of both prosocial (positive) and problematic behaviors for use with children three to 16 years of age, allowing direct comparison across findings. These impact evaluations (IEs) report on the total behavioral difficulties score from the SDQ, a composite score built from reports on both internalizing and externalizing behaviors.

Three of the IEs found that early childhood interventions can improve subsequent behavioral problems, while the fourth found no effect. In rural children in the early treatment group, who started Progresa 18 months before those in the later treatment group, had significantly better scores (that is, fewer behavioral problems) than those who started later.[8] There is evidence that this benefit may arise from the conditionalities imposed by the program rather than from the cash transfer itself. One study isolates the effect of the conditionalities and the cash separately for Progresa and finds that an additional 18 months of program participation during the early childhood period leads to significantly fewer behavioral problems, but there was no noticeable effect from the additional 18 months of cash transfers.[32]

In Jamaica, a cohort of low-birthweight babies was chosen to receive early stimulation through weekly home visits in which the mother was taught to converse and play with her child. This interaction led to a significant improvement in behavior at six years old, reducing the total difficulties score by 0.4 standard deviations.[54]

The only intervention that did not significantly affect children's behavior problem score was the breastfeeding promotion program in Belarus (see box 1.1). Both parents and teachers were asked to complete the SDQ, and neither reported significantly different behavioral difficulties for children who had benefited from breastfeeding promotion.[28]

Socioemotional Summary

While the evidence overall for the effect of early childhood interventions on later-life socioemotional outcomes is thin and at times inconsistent, the general trend did coincide in part with previous theoretical or early childhood IE work in the developing world. In particular, the programs that seem to be most effective are those that promote early stimulation and early learning (Engle and others 2011; Walker and others 2007, 2011). This is not always true, however, as shown by the Mauritian project's lack of effect on internalizing behavior. Furthermore, the theorized link between early nutrition and socioemotional development cannot be validated for later-life development as none of the three nutrition interventions found a significant post-early childhood effect (Black and others 2013; Engle and others 2007; Walker and others 2007). Interestingly, as in the

United States, these programs can have a lasting effect on violent behavior, reducing a person's likelihood to commit a crime.

---

1 Externalizing behaviors refer to behaviors such as hyperactivity, disruptive behavior, aggression, fighting, delinquency, and other unruly behavior.

2 Internalizing behaviors refer to behaviors such as social withdrawal, inhibition, depression, anxiety, eating disorders.

3 Study [41] contains two separate evaluations at ages 17 and 23 and so is listed twice.

1 A preschool intervention in Argentina did have a significant and relatively large effect on third graders' attention, but this measure is not included in the main analysis because it is based on teachers' responses to the question, "How many of your students pays a lot of attention in class?" with little other information on the measure's validity.[9]

2 This index is derived from the Arithmetic and Digit Span subscales of the Wechsler Intelligence Scale for Children III (Wechsler 1991) and reflect working memory, attention, and concentration abilities.

3 A study of the breastfeeding promotion program, PROBIT, in Belarus also calculated externalizing and internalizing behavior factor scores, but the results are not included here because too little information was provided on the psychometrics of the outcomes to be able to determine their comparability to the other outcome measures. Neither score was significantly different between the treatment and control groups.[28]

4 A violent crime was defined as a fight with a weapon, hurting someone with a weapon, carrying a gun in the past month, threatening someone with a gun, shooting someone with a gun, or being a gang member.

5 Weapon use was defined as carrying a weapon within the past month or ever threatening someone with a weapon. The authors do not discuss how prevalent weapon use was among the respondents, so it is possible that the null effect is due to a power issue rather than a true lack of effect.

6 In combining these three measures, the authors transformed the factor scores so that higher levels are more desirable.

5. Schooling Outcomes

Early stimulation, preschool, and conditional cash transfers (CCTs) seem most effective in improving schooling outcomes, while nutrition programs on average did not have an effect.

Schooling outcomes, which range from the number of years completed to academic performance, were some of the most widely studied in the evaluations of interventions from early childhood on later outcomes. There are a number of possible pathways through which early childhood interventions could affect schooling. For instance, improved cognitive development could result in increased scholastic achievement, while healthier children are better able to attend classes.

A 2003 systematic review of the effectiveness of comprehensive early childhood development (ECD) programs in the United States revealed significant improvements in school readiness, achievement, and retention rates among enrolled children (Anderson and others 2003). Additional long-term studies demonstrated that children who attend these programs were less likely to repeat a grade and more likely to graduate high school when compared with their peers who had not been enrolled (Schweinhart 2007). Two decades after children participated in Chicago's Child-Parent Centers, they demonstrated lower rates of grade retention, a lower likelihood of being enrolled in special education programs, and a higher percentage of high school completion (Reynolds and others 2001). At age 28, their schooling achievements translated to labor market gains and reduced arrest rates.

Early childhood development studies have also used theory, longitudinal data, and at times impact evaluations to try to establish which interventions work best in the developing world. As found in the United States, preschool is believed to produce later-life effects on schooling outcomes (Engle and others 2007, 2011) as are nutrition programs, presumably through better health and cognition (Black and others 2013; Bhutta and others 2013; Walker and others 2007; Engle and others 2007). Empirical evidence supporting the latter is mixed.

Table 5.1 maps the 21 studies that investigate schooling outcomes in 16 countries. There are 17 different projects and 13 intervention types, identifying the four specific indicators most commonly assessed and the age of the individual at the time of evaluation.

Table 5.1. Impact Evaluations Investigating Schooling Outcomes

		Study	Country (Project)	Average Age at Intervention (Years)	Average Length of Exposure (Years)[b]	Age at Evaluation (Years)	Evaluated Intervention	Reviewed Outcomes
Nutrition		Kramer and others 2008a [27]	Belarus (Promotion of Breastfeeding Intervention Trial [PROBIT])	0	1	6	breastfeeding promotion	math achievement; other subjects; reading**; writing**
		Pongcharoen 2010 [38]	Thailand (micronutrient supplementation to children)	0.5	0.5	9	micronutrients and fortified food for children (Iron and/or zinc supplementation)	English; math achievement; on-time primary school enrollment; science; Thai
		Alderman and others 2014 [1]	Gambia, The (Maternal Supplementation)	in utero	0.5 (DR1)	16–22	vitamins, micronutrients, or fortified food for pregnant women (protein biscuits)	school years completed
		Walker and others 2005 [52][a]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	17–18	supplementary feeding	math assessment
		Walker and others 2011 [55][a]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	22	supplementary feeding	general exams; math assessment; school years completed
		Maluccio and others 2009 [31]	Guatemala (INCAP supplementary feeding to children)	0	5.3	25–42	supplementary feeding	school years completed (men, women**)
Health		Cas 2012 [12]	Indonesia (Safe Motherhood program)	In utero	3	11–14	Access to obstetric and child health care	math assessment***; on-time primary school enrollment; school years completed**
Early Learning/Childcare		Martínez, Naudeau, and Pereira 2012 [35]	Mozambique (preschool)	3.45	1.5	5–9	quality early childhood and preprimary program	on-time primary school enrollment**
		Berlinski, Galiani, and Gertler 2009 [9]	Argentina (preprimary education)	4	1	8	quality early childhood and preprimary program	math achievement**; Spanish**

		Valdes 2011 [13]	Chile (Early Childhood Care and Education)	2.9	1.8	10	quality early childhood and preprimary program	math achievement***
		Berlinski, Galiani and Manacorda 2008 [11]	Uruguay (preschool)	4	1.5	7–15	quality early childhood and preprimary program	school years completed***
		Walker and others 2005 [52][a]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	17–18	stimulation	math assessment
		Gertler and others 2013 [18]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	22	stimulation	general exams*; probability of attending post-secondary school*; school years completed*
		Walker and others 2011 [55][a]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	22	stimulation	general exams*; math assessment**; school years completed**
Social Protection		Todd and Winters 2011 [47]	Mexico (Progresa)	2	2.8 (DR1)	6–9	CCT	on-time primary school enrollment*
		Secretariat of Social Development 2008 [8]	Mexico (Progresa)	1.5	1.5 (DR2)	7–11	CCT	school years completed
		Behrman, Parker, and Todd 2009 [7]	Mexico (Progresa)	1.5	1.5 (DR2)	6–14	CCT	on-time primary school enrollment (boys); on-time primary school enrollment (girls)*; school years completed***
		Rackstraw 2014 [39]	Honduras (Programa de Asignación Familiar)	1.5	2	13–15	CCT	school years completed***
		DSD, SASSA and UNICEF 2012 [14]	South Africa (Child Support Grant)	1	2.5	10	unconditional/targeted income support	numeracy; on-time primary school enrollment (boys); on-time primary school enrollment (girls)**; school years completed**
		Attanasio and Vera-	Colombia	3	1.2	8–17	childcare/daycare	probability of attending secondary

Child care		Hernández 2004 [4]	(Hogares Communitarios)					school**
Water and Sanitation		Spears and Lamba 2013 [44]	India (Total Sanitation Campaign)	0	continuous	6–8	adequate sanitation	numeracy (6 years old)**; numeracy (7–8 years old)
		Xu and Zhang 2014 [57]	China (rural drinking water program)	1	1	18–25	access to safe water	school years completed (exposed to 0–2 years)***; school years completed (exposed to 3–5 years)
Other		Pathak and Macours 2013 [37][c]	India (women's political reservation)	pre-birth birth	3 (DR1)	8	governance (women's political reservation)	numeracy

Note: Numbers in [brackets] correspond to the numbered studies in References. More details for each study are found in appendix A. CCT = conditional cash transfer; DR = dose response; INCAP = Instituto de Nutrición de Centroamérica y Panamá.

a. Jamaica [52, 55] each has a multiple intervention arm, and each intervention type has a separate row for these studies.

b. DR in the length of exposure means intervention period in terms of the Dose Response. DR is either randomized rotation (DR1) or phase-in (DR2). In terms of dose response, DR1 indicates the treatment and control group received the intervention for the same period of time but at different ages. DR2 describes the dose response where early and late treatment effect is compared, and length of exposure is the difference of the intervention period between treatment and control group.
c. India study [37] has multiple experimental arms. One treatment group is in utero when the political seats were randomized for women, and the other treatment group is between the ages of newborn and five years old during the reservation. The control group was children who were not exposed to reserved seats until six years and beyond.
* Statistically significant at 10 percent.
** Statistically significant at 5 percent.
*** Statistically significant at 1 percent.

On-Time Primary School Enrollment

Preschool programs and cash transfers appear effective in promoting on-time primary school enrollment.

Given this report's focus on children's development starting at entry into primary school, on-time primary school enrollment is one of the first possible educational outcomes that can be assessed. Evidence exists that early childhood interventions can successfully promote on-time enrollment. Of the five interventions evaluated, two of them had a significant beneficial effect, while the third had no overall effect but did decrease the probability of delayed enrollment among certain groups. The final two had no effect. In Mozambique, a preschool program significantly increased the probability that a child would enroll in primary school at age six by about 10.2 percentage points,[35] while the Mexican Progresa, led to a marginally significant 5 percent increase in the probability of enrolling at age six among children who were younger than three when they started the program.[47]

A second evaluation of Progresa examined the dose response of an additional 18 months of program participation and found that it significantly lowered the primary school starting age for 7- to 8-year-old girls but not for boys or for 9–11 year olds. The effect was fairly small (0.05 years), and the authors hypothesize that it might be due to better nutrition: the 7–8 year olds had been eligible for the nutritional supplementation, whereas the majority of the 9–11 year olds had not been.[7] The South African unconditional cash transfer also significantly improved on-time primary school enrollment for girls: there was a 26.5 percent reduction in the likelihood of delayed primary school enrollment for girls who starting receiving the Child Support Grants as birth rather than at 6 years old.[14] A similar effect was found for children whose mothers had less than 8 years of schooling, but there was no detectable effect for boys. No difference was found in the age at entry to primary school among the groups of Thai children who received iron, zinc, or both as compared to the placebo group or among Indonesian children whose villages had participated in a Safe Motherhood program.[38, 12]

Box 5.1. Mexico's Progresa Conditional Cash Transfer Program Improves Post-Early Childhood Schooling Outcomes

Progresa, a large-scale social assistance program in Mexico, was created to improve the lives of poor families through interventions in health, nutrition, and education. The program used conditional cash transfers to encourage healthy behavior and found significant effects on schooling outcomes throughout a child's life. Every two months, participating families received a cash transfer worth 20–30 percent of their household income. To qualify for the transfer, family members of all ages had to follow prescribed health interventions such as regular doctor visits and nutritional education programs. Several of the components aimed to improve early childhood development, including the requirement that children under two years old receive immunizations and take nutritional supplements. Mothers of children two to five years old had to attend health and hygiene information sessions, and children were required to have their nutrition monitored and take a nutritional supplement if they presented with signs of undernutrition (Gertler 2004).

Evaluations of Progresa revealed that the program had significant effects on children's anthropometric and physical outcomes during their early childhood years. Infants below the age of six months whose families lived in urban areas and were enrolled in the program grew significantly more in both height and weight than infants of the same age in nonparticipating families (Leroy and others 2008), and children up to three years old who were exposed to 24 months of the interventions were significantly taller and were less likely to be ill and anemic (Gertler 2004). When isolating the effect of the cash transfer from the program conditionalities, there were still significant, although relatively small, improvements in physical outcomes among 24- to 68-month-old children who had been enrolled in Progresa since birth (Fernald, Gertler, and Neufeld 2008).

Although these physical gains were not sustained in the primary school period, subsequent studies did identify longer-term schooling effects. Children who started Progresa when they were still young enough to receive the nutritional supplementation (less than five years old) were more likely to enroll in primary school at age six and less likely to miss school[47] than were children who enrolled in Progresa at age five or older. Seven to eight-year-old girls who were two young to receive educational scholarships but did benefit from the health interventions, entered primary school at a younger age than girls who had been exposed to the program for 18 fewer months. This trend that was also observed when comparing seven- to eight-year-old girls who had been enrolled in Progresa for five years to a matched sample of girls who had never participated in the program. In addition to entering primary school at a younger age, children who were old enough to qualify for the educational grants also increased the number of grades they completed by almost 9 percent.[7] Although it appears that data may be available to isolate the ECD effect by netting the combined effect from the school-only effect, this analysis has not been done to date.

In 2002, Progresa became known as Oportunidades (and was rebranded as Prospera in 2014), and school attendance of children age nine years and older in the household was introduced as a condition for cash transfers. The timing of this new component has made it difficult to separate the effects of the early childhood health interventions from the effects of increased school attendance due to the cash transfer, as some children may have received both during the time of evaluation. Since the transfer was given to families, it is also possible that children who were too young to qualify for the grant still benefited from cash the family received for an older sibling's school attendance. Despite the difficulty in disentangling causal pathways, the combined effect of both nutritional inputs and schooling indicates and prolonged, age-appropriate programs transitioning from early childhood into the schooling period can be effective in increasing schooling attendance. Several studies have also tried to assess the program's lasting impact on additional developmental indicators. The table below presents all of the outcomes reported by evaluations of Progresa included in this review and shows that lasting effects were also observed in the physical, cognitive, language, and socioemotional domains, although the evidence is not consistent.

Outcome Domain by Age at Evaluation in Mexico Progresa Study

Age	Study	Physical Development	Cognitive Development	Language Development	Socioemotional Development	Schooling Outcomes	Labor Market Outcomes
Below 5	Initial Studies	18/35	3/3	1/1			
6–9	[47]	-	-	-	-	2/2	-
7–11	[8]	0/4	0/1	1/2	1/1	0/2	-
8–10[2]	[32]	1/2	1/1	1/1	0/1	-	-
8–10[3]	[32]	0/2	0/1	0/1	1/1	-	-
6–14	[7]	-	-	-	-	8/10	-

Note: To provide a more complete scope of these studies, this table includes all reported outcomes and not just those analyzed in the main body of this report (see box 1 in the Introduction for the decision rule for selecting outcomes for analysis). The numerator denotes statistically significant outcome at 10 percent level or better, and the denominator is the number of outcomes in the domain. The [bracketed] superscript number in the "Study" column is indicates study identifier (see References).

a. Five "Initial Studies" were found that give estimates of program effects during early childhood: Barham (2011); Fernald, Gertler, and Neufeld (2008); Gertler (2004); Leroy and others (2008); and Rivera and others (2004).

Years of Schooling Completed

Given time, preschool, early stimulation, and cash transfers appear to increase schooling, but nutrition programs may not be effective.

The most widely studied educational outcome was years of schooling completed, and the evidences suggests that over time, various intervention types could positively affect this indicator. In a retrospective study in Uruguay, evaluators examined the effect of preschool on subsequent schooling and found that benefits grew with time. At seven years old, there was a significant decrease of 0.34 years of schooling among children who attended preschool, which the authors posit is from a delay in preschoolers starting primary school, but they give no evidence or explanation for this view. One explanation may be that preschool displaced early enrollment in primary school, as seen in a World Bank project in Cambodia (Bouguen and others 2013).

Despite the inauspicious beginning, former preschoolers eventually overtake their peers in schooling attainment. By age 11, those who attended preschool had 0.25 more years of schooling than those who did not, and the number grew to a highly significant 0.79 years by the time the participants were 15 years old, suggesting that preschool attendance, even if it delayed entry to primary school, prevented dropout or grade retention.[10] The increase was more pronounced among children whose mothers had lower levels of education, and those who lived outside of the capital city, Montevideo.

A Safe Motherhood program in Indonesia also led to a significant increase in years of school completed. In this program, the government deployed 54,000 newly graduated nurses to rural villages throughout the country in order to increase women's access to healthcare and safe delivery services. The nurses were expected to remain in the same village for three to six years, where they would provide public healthcare during regular business hours and could establish a private practice on the side. Eleven to 14-year-olds who were exposed to this program before they turned four years old had completed 0.17 more years of schooling than those who either received the program after they turned four years old or who lived in villages without the program. Interestingly, the effect is much larger and highly significant for those who are born in villages where the program was already established before they were born—ranging from 0.48 to 0.52 years, depending on whether the SMP nurse arrived one, two, or three years prior to the child's birth. The effect is still large if the nurse arrived at the year of birth (0.23 years) or within the first year after the child's birth (0.29 years), but by two years old, the effect has disappeared and remains insignificant if the nurse does not arrive until the child is three years of age.[12] This may indicate that for health programs to have an effect, they must be in place early in a child's life and have had sufficient time to develop, perhaps thereby creating a more generally healthy population in which the child can grow.

After five to six years of exposure, a dose-response assessment compared the early treatment group that received benefits when Progresa began in 1998 with the late treatment group incorporated 18 months later. There was no significant difference in grades of school completed between the two groups.[7, 8] However, one study also compared the early treatment group to a matched group of nonparticipants. While there was no effect on years of schooling completed for six to eight year olds, there was a highly significant increase in schooling among nine- to eleven-year-olds, with a particularly large return for both boys (0.4 years) and girls (0.27 years).[7]

Two other cash transfer programs also significantly increased schooling among participants. In Honduras, ten years later, children whose families had been eligible to receive the *Programa de Asignación Familiar* (PRAF) transfer from the birth of the child to age three had completed 0.11 more school years than their noneligible peers.[39] A similar effect was found in South Africa: at age ten, children whose families had enrolled in the Child Support Grant when the child was born had 0.14 years more of schooling than those whose families enrolled when the child was six.[14] While both of these results are significant, the effect size is fairly small and could indicate a difficulty in using years of schooling as an outcome at a young age. Differences in schooling accumulation may not be large enough to be detectable.

Overall, these findings are generally encouraging, but as shown in both Uruguay and Mexico, examining longer-term impacts for schooling outcomes is particularly helpful since it allows time for differences to grow. Fortunately, two studies of the same intervention conducted their follow-up analysis about 20 years after the conclusion of the intervention, when much of schooling would be completed. In Jamaica, 20-year follow-up studies to a psychosocial stimulation program for stunted children showed that beneficiaries had completed between 0.36 and 0.61 more years of education than those who did not.[18, 55] This improvement was sufficient to catch up to the nonstunted comparison group recruited at baseline, which while not randomly selected was subsequently shown to be similar in characteristics to the larger urban poor population in Jamaica.[18]

While it may be expected that cash transfers (many of which specifically promote education) and preschool programs would increase schooling among participants, the largest effect on schooling actually came from a clean water program in China.[57] Children whose villages had treated water by the time they were 0–2 years old had 1.7 more years of schooling than their nonexposed peers by the time they were 18–25 years old. There was no effect on children who were exposed when they were 3–5 years old, however, which may indicate that such interventions, which are likely to work through improved health, are more effective when started at very young ages.

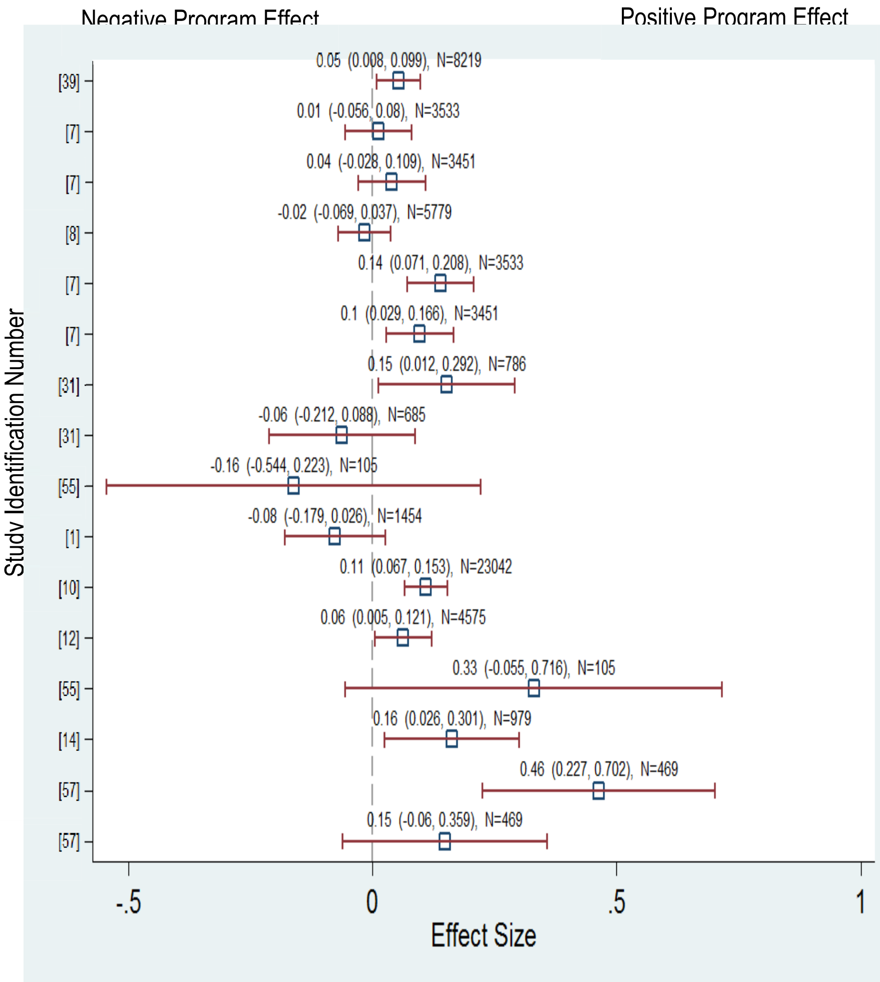
The final type of intervention evaluated for its effect on years of schooling is nutrition. Despite some positive findings in Guatemala, it appears that nutrition programs may not be as effective in promoting increased schooling. Twenty-five years after Guatemala's INCAP program ended, women who had received atole, a protein-rich supplemental beverage, before they were 36 months old had completed 1.17 years more schooling than women who did not.[31] There was no corresponding effect among male participants. However, these results are likely capturing a catch-up effect: after starting from a lower average of accumulated schooling, females increased their years of school completed by a larger margin than males, resulting in greater equity across genders than in the previous generation.

This conclusion is supported by the lack of an effect in the other two nutrition interventions examined. In The Gambia, which examined the difference in timing between giving mothers protein-rich biscuits during gestation or postpartum, there was no effect on the highest grade achieved at ages 16–22 years by the children born during the intervention.[1] In Jamaica, 22 year olds who had received supplementation until they were two years old had not completed significantly more school than those who had not received the supplementation.[55]

Figure 5.1. Forest Plot for School Years Completed

Country	Intervention	Measurement	Average Age at Evaluation (years)
Honduras	CCT	Highest grade attained	14
Mexico	CCT	Highest grade attained (girl)	7
Mexico	CCT	Highest grade attained (boy)	7
Mexico	CCT	School years completed	9
Mexico	CCT	Highest grade attained (girl)	10
Mexico	CCT	Highest grade attained (boy)	10
Guatemala	Nutrition (child, SF)	Highest grade completed (boy)	32.3
Guatemala	Nutrition (child, SF)	Highest grade completed (girl)	32.3
Jamaica[a]	Nutrition (stunted children)	Highest grade level	22
Gambia, The	Nutrition (mother)	Highest grade attained	19.6
Uruguay	Pre-primary program	Years of schooling	15
Indonesia	Safe motherhood program	School level attained	12.5
Jamaica[a]	Stimulation (Stunted children)	Highest grade level	22
South Africa	Unconditional cash transfer	Highest grade attained	10
China	Water quality	Highest grade attained	19
China	Water quality	Highest grade attained	22

Study and Forest Plot



Note: The forest plot describes standardized mean difference, 95 percent lower and upper bound confidence interval in parentheses, and sample size (N = number). The standard mean difference and confidence interval were calculated by Comprehensive Meta-Analysis software. Bracketed numbers correspond to numbered studies in References. SF = Supplementary feeding.

a. For Jamaica study [55], the stimulation compares “stimulation only + stimulation and supplementation” and “no intervention + supplementation only.”

Post-Primary Attendance

Participants of early stimulation programs appear more likely to attend post-primary school.

Two early childhood interventions were studied for their effect on another common schooling outcome, post-primary attendance, and both had a positive effect. In Colombia, Hogares Comunitarios established childcare centers where a madre comunitaria cared for children up to age six, providing nutritional meals and stimulation programs.[4] During high school, from ages 13 to 17, children who participated were 19.8 percent more likely to be in school than those who did not participate.[1] Additionally, those who received psychosocial stimulation as part of the Jamaican intervention were three times more likely than those who had not received early stimulation to have some college education, again catching up to the nonstunted comparison group.[18] However, this effect was only marginally significant.

School Performance

Interventions that provided early stimulation had a sustained effect on school performance, while those that focused on nutrition alone tended not to have an effect.

Attendance and completed schooling are important measures, but they do not necessarily indicate if students are learning. To assess real changes in human capital, a better metric may be performance at school. Of the six impact evaluations that assess achievement indicators, interventions providing early stimulation had positive effects in later periods of a child's life, and those that focused on nutrition alone tended not to have an effect. PROBIT in Belarus (box 1.1) did not significantly change primary students' performance in math and other subjects, but there was a significant, but economically small, positive effect on reading and writing.[27] In Thailand, primary-aged students who had received iron and zinc supplements as infants did not perform significantly better in math, science, or Thai or English language than those who did not.[38] Lastly, supplementation of children from 9–24 months old in Jamaica had no detectable effect on the number of secondary-level exams passed.[55] Although there are reports of implementation problems with the supplementation portion of the Jamaica intervention, the lack of effects in Belarus and Thailand reinforce doubts about the effectiveness of using nutrition to affect school performance.

Stimulation, whether through home visits or preschool, can positively affect children's academic performance. When assessing the early psychosocial stimulation program provided to children in Jamaica, researchers found a marginally significant increase in the percentage of participants who passed at least one Caribbean Examinations Council exam (a standardized exam taken at the end of 11th grade) but no effect on the percent of students who passed four or more.[18] There was also a significant, large increase in the number of students who passed at least one Caribbean Advanced Proficiency Examination, which is taken at the end of the 13th grade to prepare for college entry. No one in the either the control group or the supplementation-only group had passed, compared with 9 percent in the treatment group, comprised of those who had received early stimulation as well as stimulation and supplementation together.[2] In a second study of the same intervention, stimulation led to a marginally significant increase in the number of secondary-level exams passed.[55] In Argentina, a preschool program led to a significant increase of about 8 percent of the mean in both math and Spanish scores for eight-year-old children.[9]

In Chile, stimulation and nutrition were combined in the Early Childhood Care and Education program, in which children attended preschool, received school meals, and parents were encouraged to attend classroom activities and parenting workshops.[13] Using the Sistema de Medición de la Calidad de la Educación, a national standardized test, to evaluate fourth graders, researchers found that children who participated in preschool improved about 0.2 standard deviations in math, reading, and social science test scores. This effect was fairly constant whether the child started the program at two, three, or four years old, but it was largely driven by boys. For instance, in math, girls saw only a 4.9-point rise (.08 standard deviations) compared with the 11.5-point increase (0.2 standard deviations) among boys. The effects were largest for children in the second quintile, followed by those in the third and finally by those in the first.

Box 5.2. What Helps Numeracy and Maths?

Math skills are a particularly important piece of human capital accumulation. STEM education (Science, Technology, Engineering and Mathematics) skills are receiving increased attention in the curricula of high-, middle-, and low-income economies. When combined with results for numeracy (or early mathematics), the achievement scores seen in this section begin to tell an interesting story.

Interventions that stimulate the brain improve numeracy and mathematics. Twelve impact evaluations estimate effects on numeracy or math achievement. Of those, four estimates were of three interventions that were designed to challenge the brain prior to primary school, and three yielded significant results. The Argentine preschool program and the Chilean stimulation/nutrition early education program both made notable inroads towards improving math scores for eight- and ten-year olds, respectively.[9, 13] While the positive math assessment results for the Jamaica early stimulation program were not significant at age 17–18, they were by age 22.[52, 55]

Nutrition Interventions have yet to demonstrate lasting effects on math. None of the estimates coming from four studies of nutrition-related interventions produced improvements in math or numeracy scores: breastfeeding in Belarus at age 6.5,[27] micronutrients in Thailand at age nine,[38] or supplementation in Jamaica at ages 17–18 or 22.[52, 55]

Results from nonstandard ECD interventions are a thin, if mixed, bag. A governance intervention of reserving quotas of elected representatives' seats for women produced no differences in number recognition among eight-year-olds.[37] On the other hand, the IE of the Child Support Grant CCT in South Africa found that 10-year-olds who had received the cash transfer at the beginning of primary school at age six had a marginally significant advantage in an arithmetic assessment over those who had enjoyed the transfer from birth and throughout early childhood, even though it induced significant improvements in years of schooling completed—perhaps indicating an induced selection bias wherein the transfer was incentivizing on the capability margin.[14] Indonesia's Safe Motherhood program had a similar outcome, significantly increasing the years of schooling 11–14 year olds had completed while also making them highly significantly better at math than children who did not live in Safe Motherhood villages or whose village did not receive a nurse until the child was at least four years old.[12] Last, India's Total Sanitation Campaign produced a significant improvement in recognizing double-digit numbers for six-year-olds (but not for seven- or eight-year-olds, perhaps because the test was not sufficiently discriminatory).[44]

Schooling Summary

As shown, many different intervention types can positively affect schooling outcomes in low- and middle-income countries. In general, preschool, early stimulation, and cash transfers were the most successful programs, although the strong effect produced by the clean water program in China highlights the potential of nonstandard interventions The preschool finding is unsurprising given earlier work (Engle and others 2007, 2011), but the latter two interventions suggest a promising focus area. However, despite the widely held belief of an indirect effect of health and nutrition interventions on schooling outcomes (Black and others 2013; Bhutta and others 2013; Walker and others 2007; Engle and others 2007), little evidence was found. Only two of the five nutrition interventions produced significant results and then only in certain outcomes or for specific groups.

---

[1] No significant effect was seen for primary school-aged children (8 to 12 years old).

[2] There were only 94 participants in the sample.

6. Employment and Labor Market Outcomes

Promising if yet thin evidence suggests early stimulation improves labor market outcomes.

Poor health during childhood, which can be prevented with prenatal and early childhood development (ECD) interventions, often translates into poor health and low occupational status during adulthood (Case and Paxson 2006). Studies from the United States show that early childhood services devoted to enriching an individual's environment by increasing inputs in education, health, and nutrition can determine future labor market outcomes (Smith 2009; Cunha and Heckman 2009). Long-term follow-up studies of participants in the Perry Preschool Program and Chicago's Child Parent Center program at ages 40 and 28, respectively, found significant labor market returns.[1] These gains in employment benefit both the individual in higher earnings and society through reduced welfare dependence (Barnett 1996).

Theoretical and observational literature suggests various pathways by which early childhood interventions may lead to improved labor market outcomes. Some have modeled the link from preschool and center-based programs to participants' wages, finding a positive relationship (Engle and others 2011), while others have theorized that health and nutrition interventions can improve health and cognition, which would in turn increase a person's economic productivity (Black and others 2013; Bhutta and others 2013; Black and others 2008). Finally, some authors posit that the path from early stimulation to better employment runs through improved cognitive and socioemotional outcomes (Engle and others 2007).

As yet there is still sparse experimental/quasi-experimental establishing the causality of these proposed links. Table 6.1 describes employment and labor market effects for the only early childhood intervention yet to be credibly impact evaluated on such outcomes. Despite the outcomes from this single study being well-identified, more high-quality evidence from developing economies is needed in order to establish whether the well-identified outcomes from this study and the observed developed country outcomes and hypothesized pathways hold for labor market outcomes for developing economies generally.

Table 6.1. Impact Evaluations Investigating Employment and Labor market Outcomes

	Study	Country (Project)	Average Age at Intervention (years)	Average Length of Exposure (years)	Age at Evaluation (years)	Evaluated Intervention	Reviewed Outcomes
Early Learning	Gertler and others 2013 [18] Gertler and others 2014 [19]	Jamaica (stimulation and supplementation to stunted children)	1.55	2	22	stimulation	earnings***; migration*; employment

Note: Bracketed numbers correspond to numbered studies in References. More details for each study are found in appendix a. CCT = conditional cash transfer; DR = dose response; INCAP = Instituto de Nutrición de Centroamérica y Panamá.
* Statistically significant at 10 percent.
** Statistically significant at 5 percent.
*** Statistically significant at 1 percent.

A recent, well-regarded study of the early stimulation program for stunted children in Jamaica found positive long-term impacts on earnings and employment. The authors examined monthly earnings for first, last, and current jobs as well as average earnings overall, and found that 20 years after the intervention ended, children who received psychosocial stimulation had significantly higher earnings over their entire time in the labor market. Average monthly earnings were 30 percent higher when considering all jobs and 39 percent higher for full-time permanent jobs.[19]

This effect was more pronounced among women: average monthly earnings for full-time permanent jobs were 49 percent higher among female participants and 37 percent higher among male participants. The trend was also reflected at different periods in time, whether for the participant's first, last, or current job. For example, female participants earned 66 percent more per month in their current job than female nonparticipants, and there was no detectable effect for males.[19] These gains were sufficient for the intervention group to catch up to the earnings of a nonstunted comparison group identified at baseline. While this comparison group was chosen nonrandomly, it was similar in composition to the larger urban poor population in Jamaica.

Increased opportunities stemming from migration is one possible contributing factor to the higher salaries observed among participants. The authors found marginally significant evidence that psychosocial stimulation during childhood caused participants to be 10 percentage points more likely to have migrated to the United States or the United Kingdom by the time they were 22 years old.[18] As the authors noted, migration is an interesting outcome for a number of reasons, including its possible effect on human capital and earnings. They theorized that the early stimulation could have improved skills to a point that beneficiaries faced higher incentives to move abroad to take advantage of better education or labor market opportunities.

The authors acknowledge that their results could be due to censored data, since information on earnings is available only for those who are employed. To address this issue, they tested the relationship between early stimulation and employment and found that overall there was no detectable effect on employment. Of the four measures tested—employed, employed full time, being employed in a non-temporary job, or on looking for work—there was a marginally significant 12 percent increase in being employed, but there was no effect on the other three employment outcomes. .

Given extremely limited evidence available on employment, this chapter also considers studies that would have passed the quality screening but for high attrition rates. Accordingly, one additional piece of evidence is added. In a 19-year follow-up of the Turkish Early Enrichment Program (see box 9.1), children of mothers who participated in a training program to improve cognitive stimulation at home were no more likely to be well-employed than other children (Kagitcibasi and others 2009). This conclusion was based on the Occupational Status Index, which classifies jobs based on income level and prestige. However, a marginally significant increase was observed in beneficiaries' average age at gainful employment; the authors presented this as a positive indicator, citing a cost-benefit analysis from Turkey that suggests that an earlier starting age usually corresponds to less-qualified jobs and lower lifetime earnings. If true, this suggests that TEEP may induce better employment opportunities.

Employment and Labor Market Summary

Of all of the outcome domains examined in this report, employment has the least evidence from which to draw conclusions. Only one ECD project has reliable estimates of labor market effects, though a second, less robust, impact evaluation also estimated related outcomes. Both report positive findings. Also promising is that these interventions fall within the range of interventions thought to affect employment. Existing theories have linked early stimulation to improved economic outcomes (Engle and others 2007). Bolstered by the positive findings from the United States, these initial findings should spur evaluators to continue to add to this field to better determine the early childhood interventions and pathways that can promote employment and consequently economic growth. In the meantime, these findings are encouraging for policy makers—early stimulation holds promise to be able to improve labor markets and perhaps subsequent economic growth.

---

[1] Schweinhart and others (2005) found that Perry Preschool participants at age 40 were more likely to be employed, had higher earnings, and relied less on external economic support. Reynolds and others (2001) observed higher earnings among former participants in Chicago's CPC program at age 28.

Part II: Factors Influencing Outcomes

Part I inspects the overall effects of early childhood interventions. Part II examines specific characteristics of beneficiaries and projects that can influence the effectiveness of the interventions. These characteristics are important not only to target interventions correctly, but also to determine the appropriate timing of interventions to avoid wasting finite resources.

Chapter 7 covers heterogeneous effects—whether interventions have different effects based on individual characteristics such as gender and socioeconomic status. Chapter 8 is motivated by a deeper look at one of the defining features of this review: time. Temporal effects are examined within the impact evaluation evidence for insights into the best age at intervention for sustained impacts—with particular attention on the first 1,000 days from conception to age two, benefits of longer program exposure, and whether program effects persist across time.

Chapter 7: Heterogeneous Effects—Girls, Boys, and the Poor— *While overall there appears to be gender neutrality in outcomes, girls and the poor (and children of better-educated parents) are much more likely than not to see improvements in schooling outcomes.*

Chapter 8: Time Effects—Longer, Later, Lasting?—*The literature is underdeveloped to answer important time questions. Evidence suggests that the persistence of effects over time varies by outcome domain, and longer exposure to some programs can be beneficial. More research is needed to determine whether interventions at a very early age achieve abiding benefits. The impact evaluation evidence to date suggests otherwise.*

7. Heterogeneous Effects—Girls, Boys, and the Poor

The second of the World Bank's "twin goals" is to improve shared prosperity. Beneficiary analysis of distributional impacts and heterogeneous effects can shed light on how to achieve this goal through the vehicle of early childhood development.

Important heterogeneous effects have been found in both developed and developing countries, and an analysis of these distributional differences can be used to inform more effective targeting of future programs. In the United States, there is evidence that women enjoy long-term significant benefits from preschool programs across a range of outcomes, while the long-term benefits for males are limited (Anderson 2008). When disaggregating by socioeconomic status, recent research has shown that the cognitive and noncognitive benefits of the Head Start program were largest for those individuals at the bottom of the achievement distribution. The authors hypothesize that the large effects of the Perry Preschool Program may also have been due to the low baseline abilities of the participating students (Bitler, Hoynes, and Domina 2014). Recently, Schady and others (2014) complemented this research, finding that a similar pattern exists in the developing world. In fact, even in relatively better-off areas, differences in cognitive development between socioeconomic groups are observed at a young age can persist past early childhood (Schady and others 2014). Therefore it is important that policy makers know which groups benefit the most from any given intervention and be able to identify the heterogeneous makeup of targeted intervention areas.

The impact evaluation (IE) evidence on heterogeneous effects used in this report covers a broad range of outcomes and interventions. Of the 55 studies in this review, 24 reported heterogeneous effects across all six outcome domain categories.[1] The interventions span combinations of 11 different intervention types: conditional cash transfers (CCTs), unconditional cash transfers (UCTs), parenting education, general disease prevention, childcare centers, micronutrients, early stimulation, supplementary feeding, breastfeeding promotion, preschool, and water and sanitation.

Using that evidence, this report focuses on heterogeneous effects by gender and socioeconomic status (SES). Although effects were reported across many different groups, these two were the domains with the largest evidence base. Furthermore, an evaluation of them can also contribute to meeting the World Bank's goal of shared prosperity as well as its renewed focus on gender equality.

This report discusses two different methods of reporting heterogeneous effects. In the first, the authors tested the two groups against each other. For example, they analyzed whether girls were significantly more likely to benefit than were boys by comparing the difference between beneficiary and nonbeneficiary boys to the difference between beneficiary and nonbeneficiary girls.

In the second, the authors used subgroup analysis to ascertain whether one group benefited or not, regardless of the effect size relative to the other group. That is, they tested whether girls in the treatment group were significantly different from girls in the control group, regardless of the boys' outcomes. The results from both of these methods are included in table 7.1 and are discussed separately.

Notable is that various contextual factors could also be influencing the results presented here. For example, there are likely Regional differences in gender roles and norms that affect who participates in certain interventions as well as differences in measuring income or education. The time period in which the intervention took place could also affect a parent's perceptions of gender roles.

Effects by Gender

Twenty-five impact evaluations disaggregated the interventions' effects by gender. As show in table 7.1, the majority of the outcomes measured were gender-neutral especially in the physical and socioemotional domains and for nutrition and early learning/childcare interventions. This indicates that there was not a statistically significant difference in the outcomes between genders, not that neither gender benefited, as can be seen in the Girls and Boys columns.

In examining the subgroup analyses, it can be determined whether a gender is more likely than not to benefit—for example, whether girl beneficiaries have better outcomes than girl nonbeneficiaries. Girls are much more likely than not to benefit from interventions that affect schooling outcomes and slightly more likely than not to improve in socioemotional measures. It appears, however, that interventions affecting cognitive development are less likely to be effective for girls, and both girls and boys are much more likely not to see lasting benefits in physical outcomes than they are to see them. Furthermore, no interventions demonstrated improvements for boys in either linguistic or socioemotional development. In fact, the only domain in which boys were more likely than not to benefit was in cognition.

When broken down by intervention type, it once again appears that most interventions were gender-neutral. The only exception was for water which found that boys' cognitive and linguistic development was significantly more likely to benefit from clean water than were girls, but the India Total Sanitation Campaign was the only water intervention to enter the review.[44] Both girls and boys were more likely than not to see *no* effect from social protection programs and more likely than not to see an effect from early learning and childcare programs. Girls benefited in a slight majority of the measures evaluated for nutrition programs, while there was not a single reported later-life benefit for boys from nutrition.

Latin America and the Caribbean and Europe and Central Asia exhibit a fairly high degree of gender parity, a single measure from Sub-Saharan Africa indicates that girls benefited significantly more than did boys, and there is some evidence of a greater effect for boys than girls in South Asia. Neither gender was more likely than not to benefit in Latin America and the Caribbean, Sub-Saharan Africa, or South Asia, but girls in Europe and Central Asia did tend to see long-lasting effects.

Table 7.1. Heterogeneous Effects by Gender

GENDER	Girls > Boys	Boys > Girls	Girls = Boys	Girls*	Boys[a]	Number of Unique IEs
Domains						
Physical		2	25	5/14	1/14	10
Cognition		2	2	1/4	3/4	6
Language		1	2	2/4	0/4	5
Socioemotional	2	1	12	4/7	0/5	5
Schooling	3	1	8	17/21	10/22	10
Employment			1	0/1	1/1	1
Interventions						
Nutrition	2	2	29	7/13	0/11	8
Early Learning/ Childcare	1	2	14	14/20	11/20	9
Health				1/1	1/1	1
Social Protection	2		7	7/17	3/18	6
Water and Sanitation		3				1
Region						
LAC	2	2	27	2/5	0/3	11
ECA	2	2	18	17/26	12/27	5
SSA	1			5/11	3/11	3
SAR		3	5	4/9	0/9	6

Note: Studies with heterogeneous effects by gender: [4, 6, 7, 8, 9, 10, 13, 14, 15, 18, 23, 28, 31, 33, 35, 36, 39, 44, 45, 47, 49]. ECA = Europe and Central Asia; IEs = impact evaluations; LAC = Latin America and the Caribbean; SAR = South Asia Region; SSA = Sub-Saharan Africa.
a. The fraction represents the number of measures in that domain, intervention, or Region that showed a significant effect for girls (boys) over the total number of measures that were evaluated specifically for girls (boys).

Effects by Socioeconomic Status

The two most-reported dimensions of socioeconomic status for which the reviewed studies provided heterogeneous effects are household wealth and parents' education level. Even so, these important subgroup analyses are found in only eight studies—five for wealth and three for education. Results are given in table 7.2.

Table 7.2. Heterogeneous Effects by Socioeconomic Status

INCOME	Poor > Rich	Rich > Poor	Poor = Rich	Poor	Rich	Number of Unique IEs
DOMAINS						
Physical			3			1
Cognition			1			1
Language	1		1	1/1	1/1	1
Socioemotional			1			1
Schooling	5		1	5/7	1/7	4
Employment						0
INTERVENTIONS						
Early Learning/ Childcare	3		1	3/5	0/5	3
Social Protection	3		6	3/3	2/3	2
REGION						
LAC	4		7	4/6	2/6	4
SSA	2			2/2	0/2	1
EDUCATION	Less > More	More > Less	Less = More	Less	More	Number of Unique IEs
DOMAINS						
Physical				0/4	3/4	1
Cognition						0
Language				0/2	0/2	1
Socioemotional						0
Schooling		2		1/5	4/7	3
Employment						0
INTERVENTIONS						
Early Learning/ Childcare		2		0/2	2/2	1
Social Protection				1/9	5/11	2
REGION						
LAC					2/2	1
SSA		2		1/11	5/11	2

Note: Studies with heterogeneous effects by socioeconomic status: [8, 9, 13, 14, 16, 35, 39, 47]. IEs = impact evaluations; LAC = Latin America and the Caribbean; SSA = Sub-Saharan Africa. Fractions represent the number of measures in a domain, Region, or intervention that showed a significant effect for a poorer/less educated (richer/better educated) group over the total number of measures evaluated specifically for such groups.

Physical outcomes appear to benefit the rich and the poor equally, while the poor and those with better educated parents benefit significantly more from interventions that improve schooling than do children from richer families or those with better educated parents. In fact, the poor and those with better educated parents are more likely than not to see benefits from those interventions. Children with better educated parents were also more likely than not to see later-life benefits in physical development, but parents' education was not an important factor for contributing to language benefits from ECD interventions.

In approaching the analysis by intervention type, social protection programs did appear to benefit participants differentially by either dimension of socioeconomic status, perhaps because there is likely less heterogeneity within these already-targeted interventions.

It appears that when there is a differential effect by SES, the poor and children of better educated parents benefit significantly more than the rich or children of less educated parents. While the effect for the poor is encouraging, the effect for the better educated is perhaps unsurprising: ECD programs can perpetuate schooling advantages. The interaction effect with parents' education indicates that children from better-educated homes able to take better advantage of ECD interventions in realizing gains to their own schooling.

By Region, children from poorer families were more likely than not to benefit from early learning and childcare interventions and from interventions in Latin America and the Caribbean and Sub-Saharan Africa, while children from richer families were not. In Sub-Saharan Africa, neither children from more nor less educated parents were more likely than not to see later-life effects. It bears repeating, though that thus far there are only three studies that can credibly address distributional effects by education, and extrapolation from these to an entire Region (or the developing world generally) should be done with appropriate caution.

Heterogeneous Effects Summary

Very few studies reported heterogeneous effects by either gender or socioeconomic status, making it difficult to draw strong conclusions as to which groups benefit more from early childhood interventions.

Based on the available evidence, however, the later-life effects of early childhood intervention appear to be mostly gender-neutral, especially in the physical and socioemotional domains and for nutrition and early learning/childcare interventions;

there is usually no significant difference in the benefits accrued to girls versus boys. However, there are specific domains, intervention types, and Regions in which groups are more likely than not to see benefits. For examples, girls are much more likely than not to benefit from interventions that affect schooling outcomes, but both genders are much more likely *not* to enjoy lasting physical benefits outcomes than they are to see them.

Conversely, for socioeconomic status, there are some groups that are significantly more likely to benefit than others. Interventions affecting physical outcomes appear to benefit the rich and the poor equally, but the poor and those with better educated parents benefit significantly more from interventions that improve schooling than do children from richer families or those with less educated parents. These findings, although preliminary, are useful in helping to target future interventions to the intended beneficiaries.

---

[1] Studies that gave heterogeneous effects by gender: [4, 6, 7, 8, 9, 10, 13, 14, 15, 18, 23, 28, 31, 33, 35, 36, 39, 44, 45, 47, 49]. Studies that gave heterogeneous effects by SES: [8, 9, 13, 14, 16, 35, 39, 47]. Studies [3, 42] also present heterogeneous effects by gender or SES, but the outcomes are not relevant for this analysis.

8. Time Effects—Longer, Later, and Lasting?

Because time has a central place in the report's scope and inclusion criteria, this chapter treats three additional dimensions of time as it relates to early childhood programs. The first is the temporal trajectory of benefits—that is, outcomes within a given project are traced over time to understand whether and how they change as a child ages. The second dimension discussed here is the child's age at the initiation of the intervention—in particular whether an intervention occurred during the first 1,000 days of a child's life, from conception to two years of age. Finally, the chapter explores the effects of the length of time that a child was exposed to the intervention.

Temporal Trajectories

One straightforward argument for why early intervention can be more cost effective than later interventions is that the benefits are realized for a longer period of time (Carneiro and Heckman 2003). However, implicit in this argument is the assumption that these benefits exist and persist over time. The hope is that early intervention changes the trajectory of a child's development so that benefits continue to accrue as a child grows.

Evidence from developing countries is mixed on this point. Some studies on preschool programs in the United States points to a "fade-out" effect wherein the benefits to cognition and schooling from early childhood development (ECD) interventions shrink over time (Magnuson, Ruhm, and Waldfogel 2007; Deming 2009; Barnett 1992; Berrueta-Clement and others 1984; Currie and Thomas 1998). Other reviews show the effects of early interventions persisting over time, suggesting the intervention itself may be important in determining long-term impacts (Camilli and others 2010; Reynolds and Temple 2008).

The time element of the inclusion criteria of this review allows interventions in lower- and middle-income countries to be assessed to determine whether program effects persist or fade out in developing countries.

Thirty-six of the 55 studies included in this report examine just six projects[1] by evaluating their effects at multiple points in time. Unfortunately, few of these studies evaluated the same outcome indicator (or even the same outcome domain) over time for the intervention. This omission is a lost opportunity to understand how the influence of an early childhood intervention across a child's life. Consequently, the evidence on temporal trajectories is still thin and is limited to the cognitive, linguistic, and socioemotional domains.

Still, there are a few interventions that are evaluated for effects on the same outcome over time. This section on temporal trajectories is divided into two parts. It begins by exploring whether benefits from early childhood continue into the post-early childhood period before shifting focus to programs for which there are estimates from the same outcome domain at different times in the post-early childhood period.

FROM EARLY CHILDHOOD TO LATER OUTCOMES

The first approach to identifying the temporal trajectories of early childhood programs was to determine whether benefits found in early childhood persisted into later periods. A search was conducted for impact evaluations that estimated effects for outcomes occurring during early childhood for the 24 interventions cited in this report.[2] Six projects were identified that had impact evaluation estimates of the same (or very similar) physical and cognitive development outcome in both early childhood and after age 6.[3]

The six programs analyzed in this section are Mexico's Progresa; the Gambian maternal supplementation program; the Bucharest Early Intervention Program (BEIP); the Nepali maternal supplementation program with iron, zinc and folic acid; and the Jamaican early supplementation and stimulation programs. Five of the six interventions (all but BEIP) reported early measurements of height and weight, which can be compared to subsequent estimates.[4] The majority of physical estimates found consistent null results over time, but some did fade as the children grew, and in a single case, an initial null result became significant.

Once disaggregated by intervention type, the evidence is either too thin or too heterogeneous to draw firm, general conclusions. The three nutrition programs provided temporal trajectories for physical outcomes. Two of them found an early childhood effect on weight that disappeared by follow-up at 6–8 years old in Jamaica and by 11 years old in The Gambia. Similarly, supplemented children in Jamaica experienced an initial increase in height that disappeared by 6–8 years old. There was no difference in birth length among the Gambian children, and although a significant difference in height was detected at 11 years old, that result disappeared by the final evaluation at 16 years old (Ceesay and others 1997; Walker and others 1991).[1, 20, 50, 51] The third program, which encompassed four interventions of various micronutrient combinations in Nepal, found no detectable effect on five of eight weight or height metrics at either birth or six to eight years old. However, there were changes over time in the other three anthropometric measures. An initially significant benefit in birthweight among the folic acid plus iron and multiple micronutrient groups faded by the time they were school age, but despite there being no differences at birth, children whose mothers received folic acid plus iron and zinc were taller at age 6–8 years (Christian and others 2003).[45]

For the only conditional cash transfer program—Progresa—its effect on both height and weight faded after early childhood (Fernald and others 2008),[7, 8] while the single early simulation intervention—Jamaica—found no effect either within the first 12 months of the intervention nor at 7–8 or 11–12 years old (Walker and others 1991).[50, 51]

Cognition outcomes within early childhood were only found for BEIP and for the Jamaica supplementation and stimulation project and both showed a consistent effect across time.

In Jamaica, Grantham-McGregor and others (1991) report increasing estimates of the composite development quotient (DQ) for the group receiving both supplementation and stimulation over the 24 months of project exposure (beginning at nine months of age, on average). For the stimulated-only and supplemented-only groups, there was an initial dip in the DQ at six months' exposure, followed by a consistent upward trend, with the stimulated group always outperforming the supplemented one. Estimates for the performance and the hearing and speech subcomponents of the DQ, roughly comparable to the verbal and performance elements of cognition discussed in this report, were both large and highly statistically significant at completion of the program for the stimulated group. This upward trajectory is consistent with the trend seen in verbal and performance IQ from ages 12 to 17 to 22 for the stimulation-related outcomes.

In Romania, 42- and 54-month-old children who were randomly assigned to foster care had a significantly higher DQ and full-scale IQ score than those who had randomly been chosen to stay in institutionalized care (Nelson and others 2007). When they were assessed again at 8 years old, there was still a marginally significant improvement in foster children's full-scale IQ. However, there is some evidence that the children from institutional care may catch up, as a separate analysis by the authors found that the institutional care children's IQ score had modestly increased between 54 months and 8 years old, while the foster children's IQ scores showed little change.[17]

BEYOND EARLY CHILDHOOD

Despite the valuable evidence that comes from tracing an outcome's trajectory between the early childhood and later-life periods, this analysis does not necessarily capture the full picture. Some outcomes, such as socioemotional metrics, are difficult to measure in early childhood, and consequently, many of the valid comparisons are between later-life studies. Therefore additional insights can be gained by focusing on programs that evaluate similar measures across time within the post-early childhood period.

Cognitive, linguistic, or socioemotional development can be traced for three programs: early stimulation and supplementation in Jamaica, BEIP, and preschool in Mauritius. An additional intervention, the Guatemalan supplementation program INCAP, is also analyzed in this section using supplemental material from less robust evaluations. The results from the INCAP analysis are suggestive rather than conclusive because of concerns regarding the risk of bias of those studies, but given project's longevity and importance in the literature, this analysis uses the best available evidence to illustrate the potential long-term benefits from early childhood interventions.

Socioemotional Outcomes

The evidence indicates that some socioemotional benefits of ECD interventions can persist through childhood and beyond. Researchers in Jamaica focused on internalizing and externalizing behavior. Some socioemotional outcomes of this project were measured both at adolescence and early adulthood (for example, anxiety, depression, and information on antisocial behavior including fighting) while others (for example, parent ratings of attention, hyperactivity and oppositional behavior) were dropped, and new measures were added to be more age appropriate. There is evidence of persistence for at least some internalizing behavior outcomes. Initially, at 17–18 years old, all three measures of internalizing behavior they included had improved.[53] Two of the three outcomes were re-examined at 22 years old, and one of them was still significant, although both had decreased in magnitude. A fourth measure of internalizing behavior was also measured at age 22 and showed a positive benefit.[55]

It is harder to determine a pattern for externalizing behavior. In Bucharest, improvements in externalizing behavior appeared later but were still present during the most recent evaluation. At eight years old, foster children were not significantly more socially engaged than their peers who had grown up in institutional care. However, by 10 years old they had marginally significantly better reciprocal social interactions, and at 12 years old, they were significantly less likely to exhibit oppositional defiant disorder or conduct disorder.[3, 23, 30]

In Jamaica, only a single relevant outcome was included in the study at 22 years old, which was different than the five outcomes reported in the previous study. In the earlier study, only one of the five measures show a significant improvement at ages 17–18,[53] and the single outcome reported in the later study found that stimulation significantly improved the externalizing behavior of 22 year olds.[55] In Mauritius, the goal of the preschool intervention was to reduce the prevalence of certain types of mental illness, and therefore the outcomes reported were for externalizing behavior and psychiatric disorders. At age 17, preschool helped improve all three of the psychiatric disorder outcomes as well as two of the four external behavior outcomes. By 23, however, neither of the psychiatric disorder outcomes was still significant, while one of the two external behavior measures remained significant.[41]

The same study also examined whether nutritional status at the time of intervention interacted with the intervention to affect subsequent socioemotional outcomes, and the results were similarly nonconclusive. At age 17, three of the seven measures of externalizing behavior were significantly influenced by nutritional status—leading to larger improvements among the children who had been malnourished at baseline—while at 23 years old, only one of the six measures used varied significantly by nourishment status. Taken with the previous results as well as those from Jamaica, it is difficult to conclude whether there is a persistent effect on externalizing behavior.

Cognitive and Language Development

Evaluations of the early psychosocial stimulation in Jamaica also looked at cognitive and language outcomes, which show no evidence of fade-out effects. Full-scale, performance and verbal abilities were already significantly improved by 11–12 years old, and remained significant at 17–18 and 22 years old.[51, 52, 55] In fact, the magnitude of the effect increased between 17–18 and 22 years old. Similarly, two other language outcomes—measures of verbal abilities and vocabulary—actually improved. No significant difference was found at 11–12 years old, but there were significant improvements by 17–18 years old.

One of the most intensely studied interventions across time is the Guatemalan nutritional supplementation program, INCAP. Unfortunately, only one of the 13 studies examining the post-early childhood effects of INCAP passed the quality check for risk of bias used in this review (see box 9.1). Nevertheless, given the program's importance in the early childhood development literature, it is worthwhile to review the progression of outcomes from the 13 studies. All studies are based on three data collection rounds that followed the initial endline: when the children who received the supplementation were 11–27, 21–29, and 25–42 years old. In the majority of studies, the population was restricted to those who participated before they were three years old.

The results of these less robust studies suggest that INCAP participation is associated with a lasting effect on cognitive, linguistic and schooling outcomes, but the implications for physical outcomes are less clear because—with the exception of blood pressure—identical measures of health were not followed over time. By the time of the first followup, when participants were 11–27 years old, INCAP was associated with significantly better performance on various linguistic and cognitive development indicators such as reading, vocabulary, and information processing; these benefits persisted in 25–42 year olds. Furthermore, improved schooling outcomes are observed for beneficiaries at ages 21–29 and 25–42 years old with exposure to the program. Although there was no difference in blood pressure for beneficiaries versus nonbeneficiaries at 21–29 years old, by 25–42 years old, there were some signs that the intervention may protect against high blood pressure.

TEMPORAL TRAJECTORIES SUMMARY

The evidence on temporal trajectories is thin because of a lack of repeated measures over time. There are at least two potential reasons for this. First, relevant measures of well-being change as an individual's age. For example, in Jamaica attention and oppositional behavior were rated by parents when children were 17, but parent ratings may not have been appropriate for the follow-up study of 22 year olds. However, while the exact measure may change, there may still be elements of comparability for the underlying construct (for example, internalizing or externalizing behavior), in which case more work needs to be done to map comparable age-appropriate measures to encourage careful longitudinal work. Second, there may be an implicit assumption among researchers or policy makers that outcomes are temporally static. However, the evidence reviewed here certainly refutes such a notion of time-invariant returns.

Rather than universal time-invariance, it appears that the consistency of the effects depends on the domain. The current body of evidence indicates that socioemotional benefits may fade over time, while physical outcomes consistently remain difficult to achieve; the majority of studies found no significant difference in anthropometrics either in early childhood or later in life. Cognitive benefits, however, appear to persist and may even strengthen over time—in contrast to the fade-out effects found in the Head Start evidence in the United States. Clearly more work is needed to track outcomes over time to determine the temporal trajectories of impacts from early childhood interventions whether early benefits are indeed sustained.

Stemming from the 2008 *Lancet* series on nutrition, much has been written about the importance of the first 1,000 days of a child's life from conception to two years old for long-term development. Although a global consensus has formed in particular around the belief that nutrition during this time period can have a lasting effect on a person's growth and development, very little causal evidence exists on the later-life effects of intervening during this time period. What is known may indicate that a different approach than those evaluated may be necessary to take advantage of this important developmental stage.

Age at Intervention: The First 1,000 Days

Of the programs included in this review, seven nutrition programs (encompassing 12 interventions) began during the first 1,000 days. Six suggest that neither early exposure nor relatively prolonged exposure is by itself sufficient to produce lasting effects. The seventh provides insights into a potentially effective approach.

The first four interventions—a breastfeeding promotion program in Belarus and three Thai micronutrients interventions administered during the first six months of life—were of relatively short duration. With the exception of possible language benefits and a reduction in problematic eating attitudes from breastfeeding, these nutrition interventions demonstrate very little impact on later-life outcomes.[24, 27, 28, 33, 34, 38, 42]

Supplemental feeding in Jamaica lasted much longer. It began when the children were nine months old and lasted 1.25 years until they were two years old. Although initial gains were seen in physical health and cognition as a result of this supplemental feeding (Walker and others 1991), subsequent evaluations revealed that virtually all benefits had faded as the children aged.[18, 50, 51, 52, 55]

Recent evidence indicates that stunting (and the effect on its functional correlates of cognition) may actually begin as early as conception (Prendergast and Humphrey 2014). It may be that the food supplementation intervention in Jamaica simply began too late. Indeed, of the three programs (encompassing six interventions) to provide micronutrients to pregnant women in The Gambia and Nepal, there was a marginally significant decrease in stunting resulting from two of the interventions. Overall, however, there was virtually no detectable effect on later-life outcomes in any of the domains studied in this report.[1, 15, 20, 21, 22, 45, 46]

While the results of the previous six programs may seem to indicate that early nutrition has little lasting effect, the final early nutrition program had a significant effect on outcomes across three different domains. Examining the difference between this intervention and the others may provide insight into a more effective approach to early nutrition. Guatemala's INCAP supplementation program encouraged pregnant mothers as well as women and their children under age six to go to feeding centers. By comparing the point estimates and patterns of statistical significance for those exposed to the intervention from conception to age three versus age three to age six, the impact evaluation concludes that improved nutrition during the first two to three years of life is more important than later years.[31] Younger children had larger and significant results for highest grade completed, reading comprehension, and nonverbal cognition.

The results in Guatemala, while hardly conclusive, could indicate that effective nutrition interventions must not only start early but also last longer. An intervention in Colombia reinforces this idea. (High attrition precluded this IE from passing the initial quality check, but there are some indications that the groups are indeed balanced.) Families with pregnant mothers were chosen based on whether their unborn child was at risk for stunting (as determined by whether a majority of the older siblings were stunted). The families were then randomly assigned to nutrition or early stimulation interventions that lasted from pregnancy until the child was three years old. This led to a significant improvement in weight-for-age and height-for-age at six years old, but not in weight-for-height or two other physical measures (Super and others 1990).

Taken together, we observe that only when they last throughout pregnancy and up to and beyond age two do nutrition interventions demonstrate effects that last beyond early childhood. Such interventions have shown remarkably large and long-lasting benefits on cognition, reading and education, even 25 years later.

Near-term impact evaluations and other evaluative evidence support the importance of the first 1,000 days for improving early childhood outcomes. To affect outcomes beyond the early childhood period, however, these nutrition interventions seem to indicate that not only is early intervention key, but sustaining the program throughout the entire first 1,000 days may also be necessary.

Beyond nutrition, three impact evaluations—that of South Africa's unconditional cash transfer, the clean water program in China, and the Safe Motherhood program in Indonesia—specifically isolated the effect of treatment during this period as compared to later periods. In South Africa, researchers compared children who began the program before the age of two with those who began between the ages of two and five years. At 10 years of age, there was no significant difference in the height-for-age or prevalence of stunting between these two groups.[14]

In China, however, the effect of having access to clean water when a child is zero to two years old on years of schooling completed is almost three times as high as schooling for children who did not have access to clean water until they were six to 25 years old. There is no differential effect for exposure from three to five years old versus 6 to 25.[57] And in Indonesia, children who were born into SMP villages, which provided better access to obstetric and child health care together with parent training for nutrition and health practices, were significantly better off than children for whom these benefits came after age four.[12] These children had improved cognition, math scores, and completed years of education (though there was no difference for on-time school enrollment). Interestingly, these SMP effects grow stronger the earlier in the child's life the program was put into place, with the strongest effects for those for whom the SMP nurse arrived two years prior to birth and falling off significantly for children for whom the SMP nurse arrived more than two years after birth. This suggests both that it may take time for the benefits of a health intervention aimed at behavioral change to be fully realized, and that there may be a critical window for health interventions within the first 1,000 days.

One last piece of evidence regarding the importance of the first 1,000 days comes from the Bucharest Early Intervention Program. Children were placed with foster families at any time between 9–33 months old, and four of the evaluations tried to determine whether age at placement significantly affected subsequent outcomes. The results are inconclusive: while the foster care children had better electroencephalograms if the child was placed before 24 months and better social skills if placed before 20 months, time of placement was not a significant correlate for executive function or memory, and it was placement before 26 months rather than 24 that was correlated with a significant improvement in a child's processing speed—suggesting that the critical window for cognitive stimulation may extend beyond age two.[2, 11, 17, 49]

With so few and so diverse non-nutrition intervention, it is difficult to draw strong conclusions on what the post-early childhood effect of intervening during this critical time period may be compared with later interventions.[5] Though the evidence and rationale prioritizing certain interventions during this time period is fairly new, it is hoped that ongoing and future studies can provide more information on the optimal age at intervention to achieve lasting impacts. The first 1,000 days is held to be particularly important for health and nutrition interventions, while exposure to stimulating environments and interactions throughout the early childhood years are important for cognitive and socioemotional development—this is broadly supported by the impact evaluation literature on post early childhood effects.

Nevertheless, further research into the additional benefits of stimulation during the first 1,000 days is still warranted. A coherent research agenda that investigates the relative benefits of intervening at different periods for the standard portfolio of early childhood interventions would be tremendously valuable to developing effective ECD protocols.

Length of Exposure

Regardless of when an intervention begins, it is important to determine how long an intervention should run in order to be most effective. It may be that a longer exposure period is needed in order to change participants' habits or fundamentally alter their health, in which case ending the intervention too soon would weaken the desired effects. Conversely, continuing an intervention past its optimal point could diminish its reach if scarce resources are spent over-exposing participants rather than expanding the scale of the intervention.

This concept is similar to dose-response but with a time dimension. Phased interventions, including randomized rollouts, can complicate the general interpretation of findings, but they are ideal for understanding the incremental value of additional exposure to an intervention. Unfortunately, few of the included interventions were implemented or evaluated in a manner conducive to determining the effect of varying intervention lengths.[6]

The two main types of interventions that were analyzed by length of exposure for differential impacts lasting beyond the early childhood period are cash transfers and preschool programs. The two cash transfer programs included are Mexico's Progresa and South Africa's Child Support Grant (CSG). For Progresa, the studies all examined the effect of an additional 18 months of program participation, usually when the child was younger than three, on outcomes when the children were 8–10 years old. The study on CSG looked at the differential effect on 10 year olds who enrolled at birth versus at age six. A nutritional supplementation program and a childcare program were also evaluated for length of exposure, but the evidence on these programs comes from a single study each.

The cash transfer programs examined a range of outcomes, while the preschool and childcare interventions were evaluated solely for their effect on schooling. It appears that additional exposure to cash transfers does not tend to improve children's cognitive, language, physical, or schooling outcomes, but it can have a positive effect on subsequent socioemotional development. The evidence from preschool and childcare programs, however, was mixed, with no clear indication as to why additional time in some programs was more beneficial than in others. It should be noted that, as with other areas of this report, the evidence is thin, and conclusions are suggestive rather than definitive.

Estimates from four preschool programs in four different countries—Chile, Colombia, Mozambique, and Uruguay—indicate no clear conclusion as to whether children's subsequent schooling outcomes benefit from longer enrollment in preschool. In all four of these studies there is a potential endogeneity problem as a parent's decision to enroll their child in more years of preschool is likely related to other characteristics that would affect subsequent schooling outcomes.[7]

SCHOOLING OUTCOMES

In Chile, children could attend preschool for one to three years, and while children's achievement test scores increased gradually with each additional year of preschool, the difference in the scores was not statistically significant.[13] Similarly, while preschool attendance in Uruguay overall led to an increasing advantage over nonbeneficiaries in terms of years of schooling over time—a half year by age 12 and 0.8 years more schooling by age 15—there was not a statistically significant difference in years of school attended for those who had an additional year of preschool versus those who were in preschool one less year.[10]

While attendance may not be affected by longer exposure to preschool, there is some evidence that enrollment is. For 5–9 year olds in Mozambique, each additional month that a child had been enrolled in preschool increased the probability of enrolling in primary school as well as the probability of on-time enrollment by about 1 percentage point. It did not, however, have a detectable effect on the probability of dropping out of school.[35] Tenure in childcare also had a positive effect on school enrollment. In Colombia, how much longer a child participated in Hogares Comunitarios may not have affected outcomes for 8–12 year olds, but 13–17 year olds who spent longer in the program were more likely to enroll in school and to pass a grade. When these estimates were disaggregated by gender, it appears that the effect on enrollment was largely driven by boys, while the effect on grade progression was stronger for girls.[4]

Conversely, additional exposure to cash transfers does not appear to affect a child's subsequent schooling outcomes. The number of years of school completed was not significantly different for children who had spent an additional 18 months in Progresa, nor was their probability of progressing on time through school.[7, 8] One study found a strong effect of additional time in Progresa on absenteeism of 0.12 fewer days missed or 53.6 percent fewer absences, and it also reported a marginally significant increase in the probability of primary school enrollment at age six.[47] However, a second study[47] found no detectable effect on age at enrollment for seven- to eight-year-old boys and a marginally significant decrease in age at enrollment for seven- to eight-year-old girls.[7] Despite the lack of significant findings, the effects on the education variables were all positive and of fair magnitude; the authors postulate that the analysis may have lacked statistical power to detect significant effects.

Finally, an evaluation of INCAP supplementary feeding program in Guatemala also tried to determine the added benefit of longer exposure periods. The results were not presented in full within the study, but the findings suggest that exposure for more than 12 months is important in order to change highest grade completed, reading comprehension and nonverbal cognition.[31]

COGNITIVE AND LANGUAGE DEVELOPMENT

While the evidences is mixed, it appears that overall, additional exposure to cash transfers does not affect post-childhood cognitive and language abilities. Two Progresa studies looked at the effect of an additional 18 months of program participation before the child turned three years old on cognitive and language abilities when participants were 8–10 years old. Overall, the effects were not significant for nonverbal abilities and ability to read, but there was a significant increase in participant's language development.[8] However, the null results may be masking an effect from the cash transfer itself. When the program effect was disaggregated to isolate the effect of the cash versus the conditionalities, the conditionalities continued to have a nonsignificant effect, but the cash caused a marginally significant increase in the cognitive assessment score and a highly significant increase in the verbal assessment score.[32] The authors theorize that this difference could stem from various sources, including the parents' ability to provide a better learning environment and more nutritious food from the additional income. More evidence is needed, however, as an additional six years receiving the unconditional cash transfer—the Child Support Grant (CSG)—did not lead to any additional benefit to either a child's ability to read or their reading.[14]

SOCIOEMOTIONAL DEVELOPMENT

Both Progresa studies that included socioemotional outcomes measured the change in a child's score on the Strengths and Difficulties Questionnaire, and unlike for cognition and language, it appears that additional time in the program does benefit the children's socioemotional development. Overall, there was a significant improvement in children's behavior problem score, which seems to be driven by the program's conditionalities rather than the cash transfers.[8, 32]

PHYSICAL DEVELOPMENT

Similar to the null trend found in the main analysis of physical outcomes, additional exposure to cash transfers on average did not affect physical outcomes. Both the overall program effect as well as the effect of the cash and conditionalities led to small and nonsignificant changes in body mass index.[8, 32] The same was true overall for height, although there was a highly significant 0.05 standard deviation increase in children's height-for-age caused by an additional 18 months of cash transfers.[8, 32] CSG was only analyzed for its effect on the probability of illness in the last 15 days and how long that illness lasted, neither of which was affected by an additional six years in the program.[14]

The only intervention to address length of exposure that was not a cash transfer or preschool program was a national policy in India to increase female political participation. A 1992 law required that one-third of seats in local councils as well as one-third of village-level leadership positions be reserved for women. In Andhra Pradesh, this was accomplished by randomly reserving one-third of the seats in each election cycle and rotating through by each five-year cycle. The authors exploit this random allocation by identifying children who were born at the beginning of the second election and dividing them into three groups based on location: the first group of children were born into areas where women had been in power for the previous five years and there were eight years for any changes to continue to develop, the second group lived in areas where women were assigned seats in the second election cycle and so would be in power for the first five years of a child's life and then there would be three years for any changes to continue to develop, and the third group (the control group) was born into areas where women would not benefit from the reservations until the children were six years old.

At eight years old, the first group of children had significantly better reading and word recognition scores than the third group (although their math, reading and receptive vocabulary scores were not significantly different), while there was no detectable difference between the second and the third groups.[37] This suggests that the longer period enjoyed by the first group, during which time policies could not only be influenced by female representation but also have time to create changes in the community, may be necessary to reap the benefits of female political participation.

In summary, given how important length of exposure is to determining the optimal timing of an intervention, little evidence is available at this time to guide policy makers on the effect of longer participation in any given intervention. What evidence does exist, however, highlights two important areas in which longer exposure times may be helpful in producing benefits. For preschool or childcare programs as well as supplemental feeding, it appears that longer exposure can lead to higher school enrollment and completion rates, while additional involvement in a cash transfer program during the early childhood period could help reduce behavioral problems.

Summarizing Time Effects

In examining three elements of time—whether benefits persist over time, age at intervention, and length of exposure to an intervention—this section contributes to the ongoing discussion of how to structure early childhood interventions and the potential for lasting benefits. For the latter question, it appears that the persistence of effects varies by outcome domain; socioemotional benefits may fade over time, and physical outcomes remain consistently hard to achieve, but cognitive benefits appear to persist and may even strengthen over time. The former—how to structure an effective intervention—is perhaps more difficult. Despite much being written about the importance of the first 1,000 days, the three non-nutrition interventions that begin during this time period were too varied to draw conclusions, and six of the seven nutrition interventions tended to show no post-early childhood effect. Based on the single nutrition intervention with significant later-life outcomes, it may be that not only is early intervention key, but sustaining the program throughout and beyond the entire first 1,000 days may also be necessary. The evidence surrounding length of exposure is also limited, but it appears that schooling and socioemotional benefits may accrue from longer exposure to certain interventions.

---

[1] The seven interventions for which multiple evaluations in the post-early childhood period passed the quality check are: Mexico's Progresa, PROBIT in Belarus, the Jamaican early stimulation and supplementation program for stunted children, maternal supplementation in The Gambia, the Bucharest Early Intervention Program, and a preschool program in Mauritius.

[2] Although there were diligent attempts to find studies of early outcomes for the interventions studied in this review, that search process was not as exhaustive as was the core search for post-early childhood studies. There may be additional studies that are not captured here that estimate comparable outcomes during early childhood for the 24 projects included in this review. Nevertheless, this analysis provides useful preliminary evidence on the temporal trajectories of cognitive and physical outcomes that could be affected by early childhood interventions.

[3] Schooling and employment outcomes are not relevant for children under six and socioemotional measures are still somewhat underdeveloped and infrequently applied to this young population.

[4] While some of the earlier estimates are unstandardized, and so are imperfect comparisons to the latter standardized measures, most of the unstandardized estimates were done for groups with small enough variation in age that one would not expect the standardized measures to differ significantly.

[5] Multiple iterations of meta-regression were run to try to understand the effect of beginning the intervention during the first 1,000 days and age at intervention as moderators, but none returned robust significant results. However, it is quite possible that this result is an artifact of the small number of studies available for inclusion, resulting in the regressions being underpowered.

[6] A simple analysis was done to determine the minimum and average duration of projects that produced at least a marginally significant result in four of the outcome domains (physical development and employment and labor market domains were not included due to the overall lack of results in physical development, and the fact that there was only a single impact evaluation with employment outcomes that passed the quality check). For cognition, language, and schooling, the minimum project exposure required for at least a marginally significant result was one year (the Promotion of Breastfeeding Intervention Trial in Belarus, deworming in Kenya, and preschool in Argentina), and the minimum for socioemotional was 1.5 years (Progresa dose-response). However, it should be noted that the average exposure time to produce at least a marginally significant change in any of the four domains was two years.

[7] In Chile, the author tries to address this problem by using propensity score matching, but for the other studies, the results should be taken as conditional correlations rather than causal impacts. Furthermore, the studies in Chile and Uruguay compare across treated individuals, while the children in Colombia and Mozambique are compared to children who did not receive

the intervention. Consequently, the estimates from the former two studies may be viewed as the difference in longer versus shorter preschool enrollment (infra-marginal effects), while the latter two studies should be interpreted as the difference between enrolling in preschool and not enrolling (extra-marginal effects).

Part III: Improving Study Quality and Coverage

The previous chapters analyze the current body of evidence for the post-early childhood effects of early childhood interventions. Part I examines the results by outcome, while Part II tries to determine how effects varied by subgroup and by time.

Part III takes a different approach. Rather than analyzing what is available, its objective is to aid in strengthening and broadening the knowledge of the later-life effects of early childhood development (ECD) interventions.

Despite impressive and pioneering work done to begin the study of the later-life effects of early childhood development, much is left to do to create a comprehensive evidence base for future policy makers.

Chapter 9 is intended for evaluators of early childhood interventions. It engages in a more technical discussion of common challenges for conducting impact evaluations (IEs) that straddle a long time frame between the intervention and a follow-up evaluation. Standard estimation challenges of confounding characteristics, attrition, program design, and external validity are often compounded in IEs of the post-early childhood effects of early childhood interventions that cover so much time and several developmental phases as a child moves from early childhood through primary age to adolescence and finally into adulthood.

Chapter 10 highlights notable evidence gaps and provides guidance for future evaluation to move toward a more complete understanding of impacts from early childhood interventions on later outcomes.

Chapter 9: Improving Study Quality—*Given the longer evaluation period, IEs of later-life effects of early childhood interventions are particularly susceptible to various evaluation challenges that can compromise causal inference. Particular challenges include confounding characteristics, high attrition, initial program design, and poor external validity.*

Chapter 10: Knowledge Gaps to Fill—*Evaluations must expand to cover a broader range of intervention types both inside and out of the traditional ECD sector and should include key evaluation components such as analysis of complementary methods, cost efficiency, and heterogeneous effects. More evaluations are needed everywhere, but particularly in the Middle East and North Africa and East Asia and the Pacific Regions.*

9. Improving Study Quality

This chapter includes a more technical discussion of challenges to causal inference commonly found in impact evaluations that span phases of human development from early childhood through elementary age into adolescence and adulthood. Its intended audience is those working to generate empirical evidence on early childhood development (ECD).

Early childhood development evaluations aiming to estimate impacts after a prolonged period face four challenges: confounding, attrition, design allowing follow-up, and external validity.[1] These challenges are not unique to ECD evaluations or longer-term evaluations, but they may be compounded here. In particular, issues of attrition and confoundedness are primarily responsible for excluding more than half of the impact evaluations otherwise eligible for this review. These are much higher rates of exclusion on the grounds of quality than found in previous systematic reviews by the Independent Evaluation Group (IEG), indicating these issues may be a particular challenge for ECD impact evaluations with a longitudinal element. Only three impact evaluations of the 55 used in this review were rated AAA and are believed to be confidently free of these challenges and have a low risk of bias—two from Jamaica[18, 19] and one from Uruguay.[10] The other 52 AA evaluations included were deemed plausibly free of these challenges with a moderate risk of bias.

This is not to denigrate the impact evaluation work on longitudinal ECD interventions that were necessarily designed a decade ago or more, and often without inclination to become a vehicle to study later-life effects. Indeed, many of the studies considered for review were groundbreaking in their examination of various interventions on immediate and long-term outcomes, and current perspectives on ECD have undoubtedly been influenced by this work. It is hoped that highlighting these challenges can inform and improve designs of future evaluative work.

Confounding Characteristics

Confounding factors, such as baseline differences between treatment and control groups or changing circumstances over time, can affect the developmental trajectories of the recipient population and therefore influence later outcomes. This risk grows with time as more opportunities for intervening events occur as well as longer periods in which baseline differences can cause nonnegligible changes in the groups. Before-after, interrupted time series, or simple ex post comparisons, are unlikely to be free of confounding factors other than the intervention that may explain observed changes or differences for beneficiaries. Many of the impact evaluations discovered during the search did not meet the minimum standard criteria for inclusion principally on "risk of bias" grounds because they lacked a credible counterfactual.

For example, the search uncovered 15 potential impact evaluation studies on the supplementary feeding program of Guatemala's Instituto de Nutrición de Centroamérica y Panamá (INCAP), yet the identification of effects for most of evaluations is based on a matched-randomized strategy that is flawed and unreliable (see box 9.1). Consequently, there were systematic differences between the treatment and comparison groups at baseline. This is of particular concern in long-term follow-up studies since the passage of time allows for confounding factors to interact with and magnify these initial differences. Because 14 of the INCAP evaluations do nothing to correct for these differences, the studies' risk of bias exceeds the standards set by this review, and they are not included.

Box 9.1. Identifying Challenges of Follow-Up Evaluations of ECD Interventions: Confounding Characteristics in the Instituto de Nutrición de Centroamérica y Panamá Program in Guatemala

In Guatemala, the Instituto de Nutrición de Centroamérica y Panamá (INCAP) began a series of studies to assess the impact of nutritional deficiencies on children's ability to learn. Between 1969 and 1997, INCAP administered two supplements to preschool-aged children in four Guatemalan villages: atole, a high-protein energy drink, and fresco, a fruity drink that lacked substantive nutrition.[31]

The benefits from this intervention manifested early and persisted over time, suggesting that early childhood nutritional interventions can influence human capital accumulation throughout adulthood. The first study of the effects of the program was carried out on children who were between infancy and three years old when they received the supplements. The evaluation found that the greater nutritional intake from atole resulted in better physical development (Martorell, Klein, and Delgado 1980). Twenty-five years later, children who received the atole supplement between the ages of zero and seven years had higher scores in nonverbal cognitive skills and reading comprehension, and girls had completed an average of 1.2 more years of schooling, though this advantage was not found for boys.[31]

Despite these promising results, much of the potential impact of this program remains unknown because of challenges in the initial evaluation design that were not corrected for in subsequent studies. From a pair of large villages and a pair of small villages, the design randomly selected one village from each pair to receive the intervention, resulting in one large and one small village being in the treatment group with the other large and small village being assigned to the control.[2] Although the sample size is large (more than 1,000 children), there are too few units of randomization (the four villages in this case) to rely on the Law of Large Numbers to generate treatment and comparison groups that are statistically equivalent. The randomization process allowed only four possible permutations of groupings.

As a result, the validity of the counterfactual relies on the strength of the pre-randomization matching exercise, which used only size and geography to match the large and small villages.

There again, though, the small number of matching variables is not sufficiently large for matching to be credible over all relevant observed and unobserved characteristics. Early authors indicate systematic baseline differences between the treatment and control group over other important characteristics (for example, Pollitt and others 1993).[3]

Only a single evaluation of the later-life effects of this program corrected for these problems (Maluccio and others 2009). However, given the initial positive findings and the importance of nutrition to early childhood development, it may be a worthwhile exercise to re-estimate program effects from previously generated INCAP data using these methods.

Similarly, the series of studies carried out on the Matlab family planning intervention in Bangladesh are frequently cited for their effects on fertility and related maternal and child livelihood outcomes. Despite the fact that these outcomes are often interpreted as resulting from the random assignment of treatment and control groups, the villages were not chosen randomly. The two sample populations differed along several baseline characteristics. Of the five Matlab studies identified as relevant in the search phase, only one[5] adequately addressed these differences and passed the quality rating criteria, the others being assigned an A rating.

Concerns for potentially biased effect estimates are not limited to these well-known interventions; similar violations of the identifying assumptions of impact evaluation methodologies afflict nearly all of the A-rated studies, resulting in their exclusion.

Attrition

If the participants who leave the study are systematically different from those who stay, or if assignment to the treatment group is correlated with attrition, then the remaining sample will produce biased estimates. The time lag between intervention and evaluation for long-term studies also results in the risk of high attrition rates. The likelihood of attrition increases as individuals transition into adulthood and establish their own households. Studies that were not originally designed as longitudinal evaluations often lack sufficient protocols for maintaining contact information for participants. The difficulty in tracking down individuals may be exacerbated in low- and middle-income countries, where movement and migration may be responses to urbanization, poverty, civil unrest, or the effects of natural disasters. There is also the possibility that participants who benefitted most from an early intervention—in terms of health, cognition, social behavior—were more motivated to seek educational or employment opportunities far from the study’s origins, as seemed to be the case in the Jamaica studies,[18] making follow-up more challenging. Some attrition may also result from refusals by individuals who are not willing to undergo the physical or psychological assessments, perhaps because of the time required to do so or their perceived discomfort in participating in such tests.

Recognizing the heightened potential for attrition for this type of research, greater allowance was made for higher attrition rates in this review of longer-term impacts than in past IEG systematic reviews. Even so, the potential for bias created by differential attrition remains a concern. Therefore, IEG's criteria for inclusion in this systematic review required no study have an attrition rate higher than 40 percent, and if a study had an attrition rate higher than 20 percent, that study needed to demonstrate no differences in baseline characteristics between those who attrite and those who do not and demonstrate that the likelihood of attrition was not related to assignment to the intervention or comparison group. Findings of differential or selected attrition would disqualify a study from inclusion, regardless of the attrition rate.

Even where authors explore for differential attrition, there are some concerns about the methods being used. All of the analyses for attrition bias in studies in this systematic review use Type II statistical tests—that is, the null hypothesis is that there is no difference between those who leave the program and those who do not. The burden is on the data to prove there is no difference. Failure to reject the null of "no difference" is taken to be acceptance of the null, even though failure to reject can be a mere statistical artifact of a lack of power or underlying dispersion in the data. Type I tests would seem more appropriate—a null hypothesis that assumes those who leave and those who do not are different unless the data demonstrate otherwise. If Type II tests are used and the null is not rejected, authors could at least include a calculation of the sample size required to make the observed differences in the test significant at the 95 percent level. Alternately, authors could indicate the size of difference that their analytic sample would detect. If that difference is greater than, say, 20 percent of the baseline value for that characteristic, then authors should be wary that the failure to reject the null is more likely due to having a low-powered test than because there is no differential attrition.

Attrition levels can become extremely high—50 percent or more in some cases[4]. Three programs had impact evaluations with attrition rates between 40 and 45 percent and several more had attrition rates even higher. For transparency, we describe briefly these three studies that were barely excluded for reasons of attrition but otherwise passed the quality review for risk of bias:

- An iron supplementation program for nonstunted infants in Chile resulted in worse scores on every measured outcome at 10 years old (significant for spatial memory and visual-motor integration and suggestive for IQ, arithmetic, visual perception and motor coordination). This perverse result was caused either by differential attrition in the 43 percent dropout rate or by iron poisoning, as those who stayed in the trial were generally healthy and had higher socioeconomic status and developmental scores (Lozoff and others 2012).

- Food supplementation and home visits for cognitive stimulation for Colombian infants produced taller children by age three, but the effect was halved and only marginally significant by age six. The 43 percent attrition rate was found to be "nonrandom with regard to several background factors" in one of three sets of attrition tests (Super and others 1990).
- Vitamin A supplementation was offered to women in Nepal before and during pregnancy, and outcomes were measured when the children were 10–13 years old. At that point there were no differences in cognition or motor development between those children born in the (randomly assigned) treatment villages and placebo villages. Attrition ranged from 40–50 percent (Buckley and others 2014).

In spite of high attrition, some interventions become quite well-known and are studied over several rounds of surveys. The evaluations of the Turkish Early Enrichment Project (TEEP) intervention in Turkey are an example. Because attrition was nearly 50 percent, results from this series of studies was excluded from the main results of this review (see box 9.2).

Box 9.2. Identifying Challenges of Long-Term Evaluations: Attrition in the Turkish Early Enrichment Project

Between 1983 and 1985, the Turkish Early Enrichment Project (TEEP) tracked a group of 255 children between the ages of four and six from low-income areas of Istanbul as they progressed through a variety of preschool environments. The children attended daycare at either an educational or a custodial center or stayed at home, depending on their mothers' occupation. From each type of daycare group, half of the mothers were randomly assigned to training on cognitive stimulation and structured activities to facilitate engagement with their children.

Although long-term follow-up evaluations of TEEP suggest a generally positive trend of cognitive and social growth for participants, the evidence base for these claims is critically weakened by endogenous selection into care facilities and high rates of attrition. These issues challenge the internal validity of the TEEP studies, and their results should be carefully examined before wider lessons can be drawn about the effectiveness of the training curriculum or preschool options.

Nonrandom assignment into treatment groups can confound results. The evaluations of the TEEP program are based on a 3 x 2 matrix comparing outcomes of the children who participated in the three different preschool environments and whose mothers either received training or did not. However, only assignment of the mother training treatment was random. Children were already enrolled in a particular daycare environment based on their mothers' occupation and location. The only valid comparison of outcomes is, therefore, between the children whose mothers received training and those who did not. However, even this comparison may be confounded by the high degree of variability within the groups, since participation in a particular daycare is endogenous.

The loss of participants over time often undermines the ability to confidently attribute effects observed throughout an individual's lifetime to an intervention early in their childhood. One of the most substantial threats to the validity of the TEEP study is the high rate of attrition in the 19-year follow-up. The sample size of 131 reflects a 49 percent attrition rate. Although the authors try to confirm the randomness of attrition, loss of nearly half of the study sample significantly reduces the ability of the evaluation to detect small effect sizes and impairs the external validity of any results. High rates of attrition is one of the most common challenges faced by long-term evaluations, and ECD interventions may be particularly affected if attrition from post-early childhood samples is correlated with the gains achieved throughout childhood and adolescence.

Challenges in Design Allowing for Follow-Up

As noted in the attrition discussion, the evaluations which today constitute the evidence base of long-term effects of ECD interventions were often not designed with that objective. Many evaluation designs were implemented before strong evidence of effects existed across a range of outcomes in the post-early childhood years and were therefore not designed to track participants into adolescence and adulthood. Additionally, universal, standardized measurement across a range of outcomes over the lifespan—cognition and socioemotional development, in particular—are not yet established, making it difficult to know how best to assess these constructs longitudinally. Although in some cases researchers are able to apply econometric methods ex post to tease out lasting effects that can be attributed to the original intervention, the absence of ex ante planning for long-term follow-up at the implementation stage has complicated causal inference from these studies.

For example, the large-scale, locally administered maternal protein-biscuit supplementation program in The Gambia is the only study to investigate the independent effects of maternal supplementation in utero.[5, [1, 20, 21, 22]] Providing the comparison group with access to the intervention undermined the ability to understand any long-term effects of receiving supplementation in utero; subsequent studies would essentially be measuring a dose response comparing the effects of receiving the biscuits for 20 weeks in utero versus 20 weeks during lactation.

However, the authors of follow-up studies have used previous research in The Gambia to convincingly argue that postnatal protein-energy supplementation did not affect the quantity and quality of breast milk (Prentice and others 1983). It is impossible to test this hypothesis for this group of mothers, but if true the integrity of the control group would be preserved for making comparisons of dichotomous treatment. This assumption enabled researchers to measure the longer-term effects of the program when the children were between the ages of 11–17 and 16–22.[1, 20] Although these evaluations were able to provide some insight into possible longer-term effects of maternal supplementation, a gap remains during the preschool and primary school years, which could highlight the lifecycle or fadeout effects of supplementation benefits (table 9.1).

Table 9.1. Outcome Domain by Age at Evaluation in The Gambia Study

Age	Study	Physical Develop-ment	Cognitive Develop-ment	Language Development	Socioemotional Development	Schooling Outcomes	Employment and Labor Market Outcomes
11–17	[20, 22]	2/22	—	—	—	—	—
16–22	[1]	0/1	0/3	0/1	—	0/1	—

Note: Numerator is statistically significant outcome at 10 percent level, and denominator is number of outcomes in the domain.

External Validity

Finally, all evaluations—impact evaluations or otherwise—have challenges of external validity—the ability to apply results found in one study to a different scale, context, or time. Scale elements of external validity are particularly challenging for the selection of impact evaluations reviewed here. Most of the evaluated interventions are somewhat small, and scaling up to a national level may present administrative or other challenges, even if those interventions are not resource intensive. For example, the Jamaica supplementation and psychosocial stimulation project was administered to 90 children. Implementation capacity, monitoring, and quality control are clearly larger challenges if a similar program were to be implemented at the national level.

Impact evaluations, like any other type of evaluative method, can establish credible in-context external validity through appropriate sampling methods. Results found from a random sample of a well-designed sampling frame are externally valid to that frame. Unfortunately, very few of the impact evaluations in this review can claim this type of external validity because the original sample was not a representative, random selection of a larger group.[6] Beyond sample validity, which can take care of some elements of external validity to a specific context (the one from which the sample was drawn), all evaluations have challenges out-of-sample external validity. Evaluative methods of all stripes, impact evaluation and otherwise, are still working out how to apply results from one area to another.

Individual evaluations can offer little on external validity over a prolonged period of time. The interventions included here occurred at a particular point in time—sometimes 30 years ago—and those same contexts are now very different places. There is no guarantee the same results would be observed if the same intervention was implemented in the same context today. Moreover, advances in knowledge, methodology, and measurement in various fields suggest that the same studies would not be implemented in the same way today.

Even so, this systematic review tries to increase understanding of temporal trajectories—how effect sizes may change over time. These streams of impacts can be traced for the six impact evaluations that have longitudinal information. For example, the cognitive advantage of the early stimulation beneficiaries in Jamaica seems to become more apparent over time, as measured by various outcomes (schooling, employment, and earnings). The effects over the range of socioemotional outcomes were not measured consistently (with the same tools),[7] but it is noted that lower scores on the depression scales were sustained from adolescence to early adulthood. Across most measures of socioemotional outcomes, however, advantages were found for the stimulation group at both time points.

While evaluative methods are still working out how to apply results from one scale, context, or time to another, the fact that some of the results described here are consistent across these elements implies a level of robustness in external validity.

---

[1] It should be noted that in addition to the four challenges noted above, many of the evaluations included interventions that were of long duration that required substantial financial investment. Moreover, many of the outcome measures employed in those studies and reviewed in this report—especially those that relied on direct measurements of people—are expensive, time-consuming, and necessitate fairly extensive training to complete. Thus, investigators may have had to compromise on some aspects of study design to carry out and evaluate the intervention. For example, see Martorell, Habicht, and Rivera (1995) for discussion of some of the difficulties relating to design of the supplementary feeding study of INCAP. Subsequent evaluations were often based on the same problematic study designs as the original study without correcting pre-existing differences between the treatment and comparison groups. None of this is unique to follow-up evaluations of early childhood development, but these challenges are worth noting. For sound work to be done, significant time and financial investments are required.

[2] In Martorell, Habicht, and Rivera (1995), the original study designers indicate that the delivery modality of the food supplementation intervention called for twice daily distribution from central feeding stations, precluding the possibility of randomizing at the individual or household levels. Furthermore, budget constraints resulted in the exclusion of two additional control villages and logistical constraints influenced the selection of the village pool.

[3] Pollitt and others (1993) is an example of a study that fell just outside of the inclusion cutoff on methodology grounds. While the authors do control for important differences using covariates, included data are not matched based on those differences, and multivariate regression is not an included impact evaluation method because of the high likelihood of persistent omitted variables bias.

[4] High attrition often seems to afflict evaluations of highly relevant interventions. For example, the only evaluation of a malaria intervention has an attrition rate of 51.6 percent—quite high, especially considering the high-mortality nature of malaria. However, the IE asserts that 17-year old children who had received malaria prophylaxis for two or three seasons when they were under five years old had better cognitive scores than those who received it for just one season or who had received a placebo. While school enrollment was similar between those receiving any prophylaxis and those receiving the placebo, those in the treatment had about a half year more of schooling achievement. (Jukes and others, 2006)

[5] While Guatemala's INCAP supplementary feeding program was available to gravid and lactating mothers, the study design did not allow isolation of the in-utero effects.

[6] Note that impact evaluations using matching methods or data trimming are likely to lose claims of external validity as they drop off-support data.

[7] This may be because the types of socioemotional factors important to measure at various ages change over time.

10. Knowledge Gaps to Fill

The evidence surrounding post-early childhood effects of early childhood interventions is often thin. Between all of the outcome domains across all possible interventions types in all low- and middle-income countries, only 55 studies were of sufficient quality for inclusion. Given the diverse range of early childhood interventions and the impact they can have across physical, cognitive, linguistic, socioemotional, educational, and employment domains throughout an individual's lifespan, there is room for far more evaluation not only on the lasting effectiveness of known programs, but also on the potential synergistic gains from bundling interventions. While the gaps are such that any effort to add high-quality impact evaluation evidence to the existing base would be useful, particular areas have greater need than others. The gaps presented here highlight those areas of concern in an effort to guide future research.

Gaps in the evidence of later-life effects from early childhood intervention exist for various reasons, many of which are factors of the types of interventions being studied or the long time horizon. First, some intervention types are less amenable to impact evaluation methods than others. For example, it can be much more difficult to quantify the physical or cognitive effects of social protection programs (for example, regulatory frameworks, birth registration, and child protection interventions) than to identify the human development effects of a direct micronutrient supplementation program. Secondly, attrition is particularly high in follow-up studies of early childhood interventions as individuals often move, especially as children reach adolescence and early adulthood. Tracking down and re-evaluating participants at multiple stages in their lives can therefore be difficult and costly, if not impossible. A third factor is that long-term funding is often difficult to secure, and funders (and researchers and journals) often implicitly assume that returns are temporally static, such that follow-up work is not needed.

Regardless of the reason, the existence of gaps is a testament to the magnitude of the work still needed to understand the later-life effects of early childhood interventions. It is hoped that identifying these gaps will attract greater interest and resources from researchers, journals, and funders.

Gaps in Outcomes by Intervention Type

Significant research gaps across both the intervention space and outcome domains remain. Even where evidence of an intervention's effect is available, that evidence is typically thin. For instance, as depicted in (figure 10.1), only 1 of the 20 identified intervention types—stimulation—has long-term effects measured in each of the six outcome domains. Furthermore, this report identified only two studies that assessed the impact of sanitation interventions on development outcomes, and not a single evaluation was found of the long-term effects of agricultural or food security programs—both interventions that are likely to affect children's development in lasting ways. The determination of effective interventions thus requires both a more densely populated evidence base and a more expansive evaluation of relevant intervention types and outcome indicators.

As seen in (figure10.1), many intervention types have very little evidence documenting their effects beyond the early childhood period. For example, only a single study measured both the cognitive and language outcomes of a quality early childhood and preprimary program, and just two measured the impact on socioemotional and physical development. Although four studies report schooling outcomes of a quality preprimary program, the lack of evaluation of other outcome domains presents a significant gap in the understanding of why and how these schooling improvements are being achieved.

In the health category, of the six identified types of pregnancy, prenatal, and postnatal interventions, only one—a family planning program[5]—had effects reported across any outcomes. Similarly, only one of the five types of disease treatment programs, deworming, was evaluated for post-early childhood effects in any outcome domain.[36] Other notable gaps include a complete lack of evidence for teacher and caregiver training programs, and educational media.

In certain outcome domains, very little evidence was found for programs in which one would expect a link between the intervention and the indicator. For example, only a single study measured socioemotional outcomes for a breastfeeding promotion intervention, and it did not detect any significant results.[28] The lack of evidence on the relationship between a child's socioemotional development and an intervention that increases the interaction between mother and infant, such as breastfeeding, is troubling and challenges preconceptions of effective ECD strategies. While it could be the case that there is a socioemotional effect from breastfeeding, the single impact evaluation identified by this report, though its sample size is the largest among the 55 impact evaluations, does not support that conclusion. More research is needed before any causal conclusions can be reliably drawn.

The most commonly reported outcome across intervention types was physical: 27 unique studies measured physical effects in the post-early childhood period. The majority of the evidence comes from evaluations of nutritional programs for mothers and children. However, the lack of evidence on the effects of other health programs is surprising. For example, only a single study exists in both the pregnancy, delivery, and postnatal interventions category (family planning)[5] and disease treatment category (deworming).[36]

In addition to needing more evidence in established ECD outcome domains, as evidence of the post-early childhood effects of early childhood interventions continues to emerge, the range of possible outcomes measured must expand to create a comprehensive picture of the impacts at each stage in an individual's life. For instance, despite the clear cognitive and educational gains among children exposed to ECD interventions, which would normally be associated with improved employment possibilities, only one study[19] investigates employment outcomes in adulthood.

While more evidence is needed across the range of possible outcomes, the scope of interventions considered relevant to early childhood development (ECD) must also be expanded. Perhaps the biggest gap is the dearth of evaluations of the post-early childhood effects of programs from outside the traditional ECD sector. For instance, only a single evaluation was found for disease treatment interventions[36] and governance programs,[37] and only two studies[44, 57] assess the development outcomes of sanitation programs. Given the generally positive impacts found by the limited evidence, much could be learned about effective ECD programming if future research included a more expansive range of relevant intervention types.

Given the prevalence of ECD programs that combine multiple interventions, evaluations that can tease out the causal pathways of the combined treatments will provide valuable information about the complex interactions of ECD interventions throughout the post-early childhood timeline. For example, a study of Progresa in Mexico disaggregated effects by both the cash and conditionality components of the program,[32] providing important insight into causal mechanisms. Although bundling of interventions within a program means that evaluations cannot always cleanly attribute impacts to a particular treatment, understanding the beneficial synergies and the most advantageous combinations can constructively influence ECD policy.

Figure 10.1. Intervention and Outcome Matrix

Intervention Type	Intervention Categories		Cognitive Development	Language Development	Socioemotional Development	Physical Development	Schooling	Labor Market	Number of Unique Studies	Number of Unique Projects
			Outcome Domains							
Nutrition	Mother	Micronutrients and iron-folic acid: supplementation and fortification for	2	1		8	2		8	4
Nutrition	Mother	Exclusive breastfeeding promotion	1	1	1	5	1		8	1
Nutrition	Children	Supplemental feeding (preschool, center-base and/or take home rations)	3	3	3	3	4		10	5
Nutrition	Children	Counseling on optimal feeding practices:	2	3	2	2	4		5	2
Nutrition	Children	Micronutrients and fortification for children	1	1		1	1		1	1
Health	Pregnancy, delivery and postnatal	Attended delivery	1				1		1	1
Health	Disease treatment	Deworming	1	1		1			1	1
Health	Access to health care	Planning for family size and spacing	1			1	1		1	1
Health	Access to health care	Well child visits, growth monitoring, screening for developmental delays	4	4	2	3	5		7	4
Health	Access to health care	Immunization	1			1	1		1	1
Water, Sanitation	Water and sanitation	Access to safe water					1		1	1
Water, Sanitation	Water and sanitation	Hygiene or hand-washing			1	1			2	1
Water, Sanitation	Water and sanitation	Adequate sanitation	1	1					1	1
Education	Parent support program		5	4	6	6	3		15	5
Education	Stimulation		7	6	10	6	2	1	18	3
Education	Early childhood and pre-primary programs	Quality teaching, programming or curricula	1	1	2	2	4		6	5
Education	Early childhood and pre-primary programs	Pre-school Infrastructure	1	1	1	1	2		2	2
Social Protection	Childcare / daycare						1		1	1
Social Protection	Social assistance transfer programs	Conditional	3	4	2	3	4		6	3
Social Protection	Social assistance transfer programs	Unconditional/targeted income support (child grants, etc)	1	1		1	1		1	1
Social Protection	Child protection interventions	Orphans and fostering	2	1	6	4			10	1
Gov	Governance reflecting ECD interests	Women's political reservation	1	1					1	1
		Number of Unique Studies	21	19	15	27	19	1	55	
		Number of Unique Projects	16	15	7	14	16	1		25

NOTE: Shown is the frequency of impact evaluations for any given intervention and outcome domain pair. It is based on the intervention types outlined in figure 2 in the Introduction, but only the intervention types that had at least one impact evaluation of sufficient quality are included.

Gaps by Evaluation Components

Impact evaluations on ECD generally, and in particular those cited in this review, offer tremendous value. The systematic review by the Independent Evaluation Group identifies four major gaps in evaluation components of the impact evaluations reviewed here: issues related to time, complementary methods, cost efficiency, and heterogeneous effects.

Three dimensions of time may be critical components to program effectiveness. First is the child's age at exposure. Interventions occurring before or after particular developmental milestones could have very different effects. Because the age of exposure is hypothesized to be an important factor in long-term program effectiveness, it would be useful for studies to explicitly test this hypothesis of critical treatment periods (for example, the "first thousand days"). Second, the length of exposure to the intervention may or may not cause differences in outcomes, as it could be difficult to determine if some null findings could actually be an artifact of too little exposure to an intervention. Finally, increasing and varying the range of ages at the time of evaluation will continue to contribute to the knowledge of post-early childhood effects of interventions during early childhood and can be used to create a trajectory of an individual's development throughout their lifespan.

Very few impact evaluations are able to report on differential effects along these time dimensions—age at exposure, length of exposure, and age at evaluation. The evaluations collected in this report address some of these, but all can be improved in answering these questions. Adding to the body of knowledge of time-varying effects for ECD could greatly benefit the effectiveness of the sector. Each of these dimensions can be tested through multiple treatment arms.

Consistency of analysis across outcome domains of ECD interventions over time is critical to fully understand how effects may change throughout an individual's lifespan. For example, although numerous studies report on the effects of the maternal supplementation program in The Gambia, outcomes were only measured during the first year of life and again when the children were between the ages of 11 and 22.[1, 20, 21, 22] The lack of evidence during younger ages represents a significant gap in the knowledge of how micronutrients received in utero can affect children's development during the primary school years, making it difficult to determine whether the null results at later ages were due to the lack of an effect or to a fadeout of earlier benefits.

Six of the programs identified in this review—the Promotion of Breastfeeding Intervention Trial in Belarus, maternal supplementation in The Gambia, Jamaica's supplementation and stimulation program, the Mauritius Child Health Project, Mexico's Progresa, and the Bucharest foster care program[1]—were evaluated at different points in time to look at a range of outcomes as children aged. However, the majority of the interventions included in this review are only evaluated at a single point in time. Yet the posited value of early childhood development interventions is their ability to influence an individual's development trajectory throughout the post-early childhood timeframe. A comprehensive illustration of the effects of early childhood interventions on the accumulation of human capital throughout a child's life can help researchers and policy makers create and target future programs, and future evaluations should aim to contribute to this knowledge.

Greater use of mixed methods should supplement rigorous econometric modeling of the differential impacts of early childhood interventions over time. Process evaluation, focus groups, and other qualitative methods can help unpack the mechanisms at work in the causal pathways and provide valuable insight into issues surrounding contextual components or implementation. None of the 55 impact evaluations reported both extensive qualitative evaluation and quantitative impact evaluation. Mixed methods can be particularly useful in understanding or eliminating possible reasons for the many null results found in this review and which, if taken at face value, may challenge prior assumptions about the effects of early childhood interventions.

Early childhood interventions may be one of the most cost-effective development strategies available to policy makers (for example, Carneiro and Heckman 2003), yet few studies report cost analysis of any kind. Where they do report these assessments, the quality is generally poor, amounting to back-of-the-envelope calculations, and estimates are rarely comparable due to different methods of calculation. However, research indicates that the period of time before a child enters primary school is the best window of opportunity for interventions to break the intergenerational transmission of poverty (World Bank 2005), and ECD interventions are generally believed to have large equilibrium effects on equity and efficiency. Therefore, it would be beneficial to policy makers if future evaluations provided more comprehensive cost assessments to accurately illustrate the gains of ECD interventions.

Gaps by Region

The 55 unique impact evaluations identified in this report span 5 World Bank Regions and 22 different countries. (Figure 10.2) shows the locations of the ECD programs that had medium- or high-quality impact evaluation (AA or AAA) evidence on post-early childhood outcomes, as determined by this report's rating system; (figure 10.3) depicts the location of every ECD program for which an impact evaluation of any quality measured post-early childhood outcomes.

The most robust evidence base was found in the Europe and Central Asia and the Latin America and Caribbean Regions, with 19 and 18 evaluations, respectively. These Regions enjoy more evidence individually than the remaining four Regions combined: nine in the Sub-Saharan Africa, six in South Asia, and three from East Asia and the Pacific. Not a single robust evaluation of the long-term effects of early childhood interventions was identified from the Middle East and North Africa Region. When the studies that did not pass the quality check are added to the count, the number of represented countries increases by only 7, to 29. The relative densities, however, remain approximately the same across Regions, with high relative representation in Europe and Central Asia and the Latin America and the Caribbean Region, low representation from most of the rest of the Regions, and still no studies from the Middle East and North Africa Region (see figure 10.3).

The distribution of evidence across income level was similarly skewed, with 71 percent (39 of 55) of the medium- and high-quality studies coming from upper-middle-income countries. Of the remaining studies, 10 evaluated programs in low-income countries, and only 6 reported results for the lower-middle-income countries. The dearth of evidence in lower-middle-income countries, which may be poised to make significant investments in ECD, highlights an area where more work is urgently needed.

The inequitable distribution of evidence on the long-term effects of early childhood interventions by Region and income level is surprising. The lack of evidence in the Regions most in need of investment in early childhood development is inconsistent for organizations like the World Bank, with its twin goals of ending poverty and increasing shared prosperity.

Figure 10.2. Locations of ECD Programs with Medium- or High-Quality Impact Evaluations

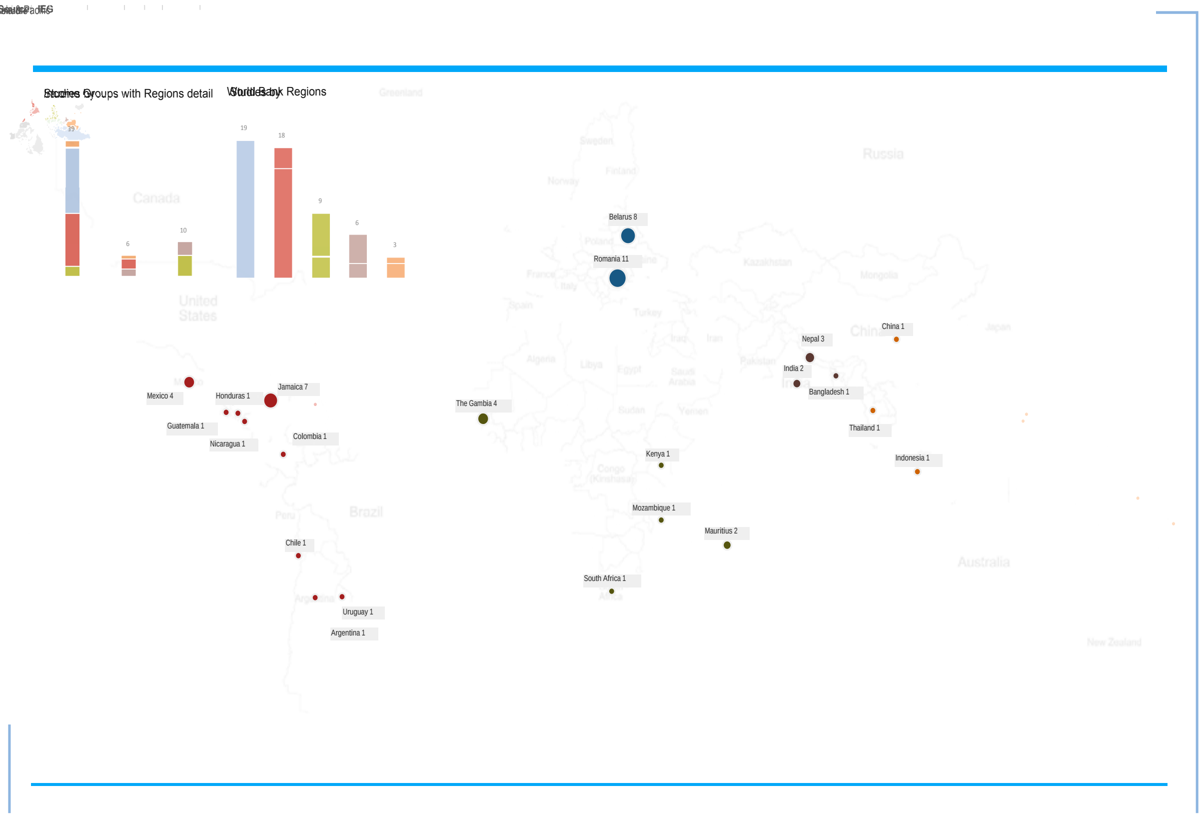
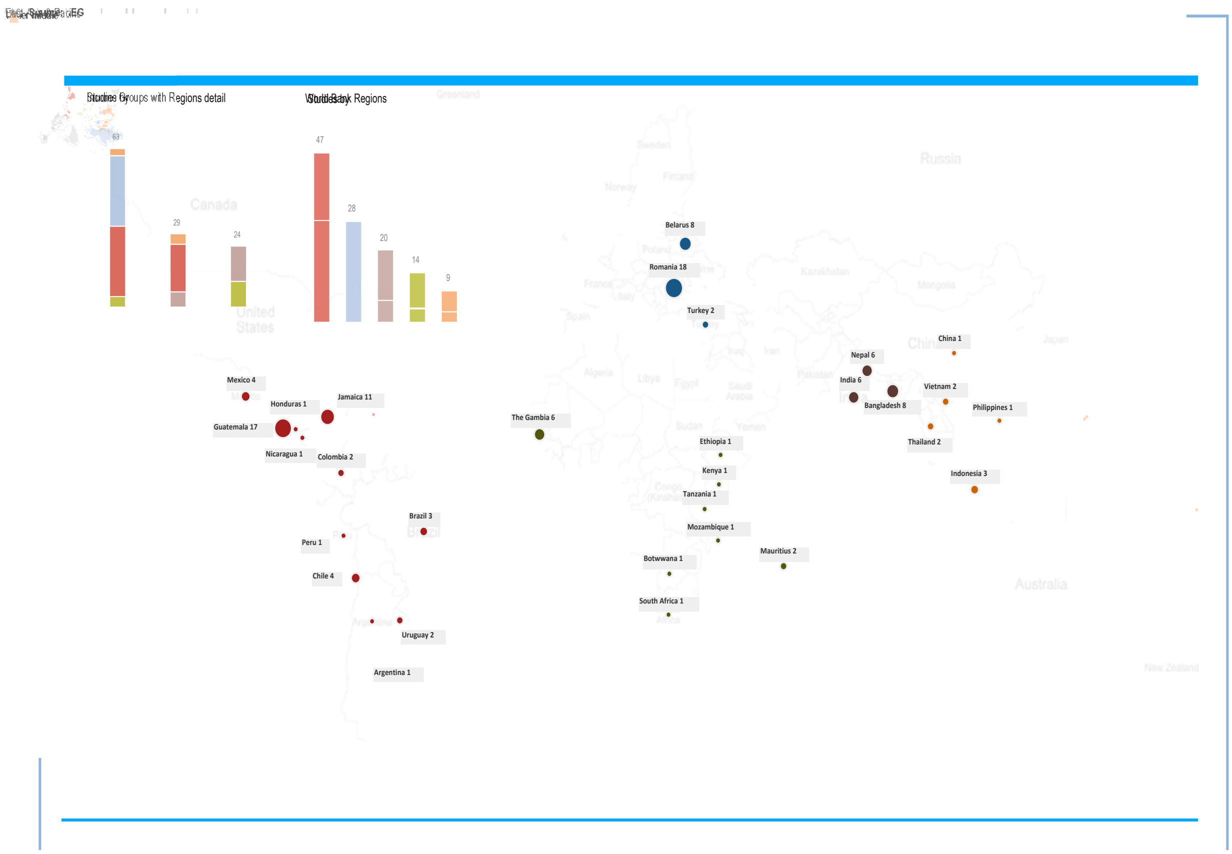


Figure 10.3. Locations of Early Childhood Development Programs with Impact Evaluations of Any Quality



Addressing the Gaps

Recent work has greatly increased the knowledge of the effectiveness of ECD programs across a variety of intervention types and outcome domains (Engle and others 2011; Grantham-McGregor and others 2014). Future research can now build upon this evidence base and increase both the breadth and depth of the knowledge of what works in early childhood development. To the extent that planned impact evaluations touch on the gaps identified above, the holes in knowledge of the post-early childhood effects of early childhood interventions may be ameliorated.

One of the areas in which more research is needed, as discussed previously, is in better understanding the synergistic effects of integrated ECD interventions. While the recent review from the *Annals of the New York Academy of Sciences* (Grantham-McGregor and others 2014) goes a long way toward addressing some of these questions by focusing on integrated nutrition and stimulation interventions, the longer-term effects of such integration are still understudied.

For example, the World Bank's Strategic Impact Evaluation Fund (SIEF) dedicates a significant amount of its focus to evaluations of ECD programs. SIEF's current portfolio includes several studies that specifically address some of the gaps mentioned here. For example, SIEF's upcoming evaluation of the government of Mozambique's expansion of a community-based integrated ECD program could present a valuable opportunity to explore these effects. In Colombia, SIEF will build on the important lessons learned from evaluations of Jamaica's supplementation and stimulation program, and will implement a similar program on a much larger scale. The proposed study design will allow the researchers to evaluate whether interventions during the first two years of life have sustained effects and further investigate the specific causal pathways through which the impacts are achieved. A third study in Indonesia will also assess how access to health services during the early childhood period impacted later schooling and labor market outcomes.

Although the results of studies like these will contribute to the knowledge of long-term effects of early childhood interventions across a range of outcomes, in multiple stages of scale-up and in a variety of contexts, significant gaps remain. As the saying goes, the best time to plant a tree is 10 years ago; the second best time is now. Careful consideration should be given now to design well-identified interventions whose internal validity can persist for decades and which can fill knowledge gaps of geography, intervention and outcome. Concurrent collection of cost data and powering for heterogeneous effects can significantly advance the Results Agenda of the World Bank and data-driven decision-making elsewhere.

---

[1] Two additional project sites were evaluated by multiple impact evaluations—the supplementary feeding program of Guatemala's Instituto de Nutrición de Centroamérica y Panamá and of Matlab, Bangladesh—but in each case only one of those evaluations in the series passed the quality criteria for inclusion in this review (AA or AAA); the others were graded as A quality.

References

The first set of references is a numbered list of the impact evaluation studies that were assigned a quality rating of AA or AAA by the evaluation team. It is followed by relevant literature sources cited in the report. There are 57 individual articles with a quality rating of AA or AAA, but two pairs of related studies are each counted as a single study to avoid the double counting of results. Study [32], evaluating Mexico's Progresa, serves as a robustness check to [16], evaluating the same outcome with the same dataset but using a different method. Study [18], as the working paper to [19], has a wider range of outcomes than [19] and is therefore included along with [19]. Each pair is counted as a single study, bringing the final count of AA or AAA studies to 55.

List of Numbered Studies

1. Alderman, Harold, Sophie Hawkesworth, Mattias Lundberg, Afia Tasneem, Henry Mark, and Sophie E. Moore. 2014. "Supplemental Feeding during Pregnancy Compared with Maternal Supplementation during Lactation Does Not Affect Schooling and Cognitive Development through Late Adolescence." *American Journal of Clinical Nutrition* 99 (1): 122–129.
2. Almas, Alisa N., Kathryn A. Degnan, Anca Radulescu, Charles A. Nelson, Charles H. Zeanah, and Nathan A. Fox. 2012. "Effects of Early Intervention and the Moderating Effects of Brain Activity on Institutionalized Children's Social Skills at Age 8." *Proceedings of the National Academy of Sciences of the United States of America* 109 (Suppl. 2): 17228–17231.
3. Almas, Alisa N., Kathryn A. Degnan, Olga L. Walker, Anca Radulescu, Charles A. Nelson, Charles H. Zeanah, and Nathan A. Fox. 2015. "The Effects of Early Institutionalization and Foster Care Intervention on Children's Social Behaviors at the Age of Eight." *Social Development* 24 (2): 225–239.
4. Attanasio, Orazio, and Marcos Vera-Hernández. 2004. "Medium and Long-Run Effects of Nutrition and Child Care: Evaluation of a Community Nursery Program in Rural Colombia." IFS Working Paper EWP04/06, Institute for Fiscal Studies, London.
5. Barham, Tania. 2012. "Enhancing Cognitive Functioning: Medium-Term Effects of a Health and Family Planning Program in Matlab." *American Economic Journal: Applied Economics* 4 (1): 245–273.
6. Barham, Tania, Karen Macours, John A. Maluccio, Ferdinando Regalia, Veronica Aguilera, and Miriam Enoe Moncada. 2014. *Assessing Long-Term Impacts of Conditional Cash Transfers on Children and Young Adults in Rural Nicaragua.* 3ie Impact Evaluation Report 17. New Delhi: International Initiative for Impact Evaluation.
7. Behrman, Jere R., Susan W. Parker, and Petra E. Todd. 2009. "Schooling Impacts of Conditional Cash Transfers on Young Children: Evidence from Mexico." *Economic Development and Cultural Change* 57 (3): 439–477.

8. Behrman, Jere R., Lia Fernald, Paul Gertler, Lynnette M. Neufeld, and Susan W. Parker. 2008. "Long-Term Effects of Oportunidades on Rural Infant and Toddler Development, Education, and Nutrition after Almost a Decade of Exposure to the Program." In *External Evaluation of Oportunidades 2008, 1997–2007: 10 Years of Intervention in Rural Areas*. Vol. 1, *Impacts of Oportunidades after 10 Years of Operation in Rural Mexico*, 15–57. Mexico City: Secretaría de Desarrollo Social.
9. Berlinski, Samuel, Sebastian Galiani, and Paul Gertler. 2009. "The Effect of Pre-Primary Education on Primary School Performance." *Journal of Public Economics* 93 (1–2): 219–234.
10. Berlinski, Samuel, Sebastian Galiani, and Marco Manacorda. 2008. "Giving Children a Better Start: Preschool Attendance and School-Age Profiles." *Journal of Public Economics* 92(5): 1416–1440.
11. Bos, Karen J., Nathan A. Fox, Charles H. Zeanah, and Charles A. Nelson. 2009. "Effects of Early Psychosocial Deprivation on the Development of Memory and Executive Function." *Frontiers in Behavioral Neuroscience* 3 (16). doi: 10.3389/neuro.08.016.2009.
12. Cas, Ava Gail. 2012. "Early Life Public Health Intervention and Adolescent Cognition: Evidence from the Safe Motherhood Program in Indonesia." Ph.D. dissertation, Sanford School of Public Policy, Duke University.
13. Cortázar Valdés, Alejandra. 2011. "Long-Term Effects of Early Childhood Education on Academic Achievement in Chile." Ph.D. dissertation, Teachers College, Columbia University. ProQuest (AAT 3492908).
14. DSD (Department of Social Development), SASSA (South African Social Security Agency), and UNICEF (United Nations Children's Fund) South Africa. 2012. "The South African Child Support Grant Impact Assessment: Evidence from a Survey of Children, Adolescents and Their Households." Pretoria: UNICEF South Africa.
15. Devakumar, Delan, Shiva Shankar Chaube, Jonathan C.K. Wells, Naomi M. Saville, Jon G. Ayres, Dharma S. Manandhar, Anthony Costello, and David Osrin. 2014. "Effect of Antenatal Multiple Micronutrient Supplementation on Anthropometry and Blood Pressure in Mid-Childhood in Nepal: Follow-Up of a Double-Blind Randomised Controlled Trial." *The Lancet Global Health* 2 (11): e654–e663.
16. Fernald, Lia C.H., Paul J. Gertler, and Lynnette M. Neufeld. 2009. "10-Year Effect of Oportunidades, Mexico's Conditional Cash Transfer Programme, on Child Growth, Cognition, Language, and Behaviour: A Longitudinal Follow-Up Study." *The Lancet* 374 (9706): 1997–2005.
17. Fox, Nathan A., Alisa N. Almas, Kathryn A. Degnan, Charles A. Nelson, and Charles H. Zeanah. 2011. "The Effects of Severe Psychosocial Deprivation and Foster Care Intervention on Cognitive Development at 8 Years of Age: Findings from the Bucharest Early Intervention Project." *Journal of Child Psychology and Psychiatry* 52 (9): 919–928.
18. Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan Chang-Lopez, and Sally Grantham-McGregor. 2013. "Labor Market Returns to Early Childhood Stimulation: A 20-Year Followup to an Experimental Intervention in Jamaica." NBER Working Paper Series No. 19185, National Bureau of Economic Research, Cambridge, Mass.
19. Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M. Chang, and Sally Grantham-McGregor. 2014. "Labor Market

Returns to an Early Childhood Stimulation Intervention in Jamaica." *Science* 344: 998–1001.

20. Hawkesworth, Sophie, Andrew M. Prentice, Anthony J.C. Fulford, and Sophie E. Moore. 2008. "Dietary Supplementation of Rural Gambian Women during Pregnancy Does Not Affect Body Composition in Offspring at 11–17 Years of Age." *Journal of Nutrition* 138 (12): 2468–2473.
21. Hawkesworth, Sophie, Andrew M. Prentice, Anthony J.C. Fulford, and Sophie E. Moore. 2009. "Maternal Protein-Energy Supplementation Does Not Affect Adolescent Blood Pressure in The Gambia." *Journal of Epidemiology* 38 (1): 119–127.
22. Hawkesworth, Sophie, Celia G. Walker, Yankuba Sawo, Anthony J.C. Fulford, Landing M.A. Jarjou, Gail R. Goldberg, Ann Prentice, Andrew M. Prentice, and Sophie E. Moore. 2011. "Nutritional Supplementation during Pregnancy and Offspring Cardiovascular Disease Risk in The Gambia." *American Journal of Clinical Nutrition* 94 (Suppl. 6): 1852S–1860S.
23. Humphreys, Kathryn L., Mary M. Gleason, Stacy S. Drury, Devi M. Murphy, Charles A. Nelson, Nathan A. Fox, and Charles H. Zeanah. Forthcoming. "Effects of Early Deprivation on Psychopathology at Age 12 Years: Follow-Up of a Randomized Controlled Trial." *Lancet Psychiatry.*
24. Kramer, Michael S., Lidia Matush, Irina Vanilovich, Robert W. Platt, Natalia Bogdanovich, Zinaida Sevkovskaya, Irina Dzikovich, Gyorgy Shishko, Jean-Paul Collet, Richard M. Martin, George D. Smith, Matthew W. Gillman, Beverley Chalmers, Ellen Hodnett, and Stanley Shapiro for the Promotion of Breastfeeding Intervention Trial (PROBIT) Study Group. 2007a. "Effects of Prolonged and Exclusive Breastfeeding on Child Height, Weight, Adiposity, and Blood Pressure at Age 6.5 Years: New Evidence from a Large Randomized Trial." *American Journal of Clinical Nutrition* 86 (6): 1717–1721.
25. Kramer, Michael S., Lidia Matush, Irina Vanilovich, Robert W. Platt, Natalia Bogdanovich, Zinaida Sevkovskaya, Irina Dzikovich, Gyorgy Shishko, and Bruce Mazer. 2007b. "Effect of Prolonged and Exclusive Breast Feeding on Risk of Allergy and Asthma: Cluster Randomised Trial." *The BMJ* 335 (7624): 815.
26. Kramer, Michael S., Irina Vanilovich, Lidia Matush, Natalia Bogdanovich, Xun Zhang, Gyorgy Shishko, Michèle Muller-Bolla, and Robert W. Platt. "The Effect of Prolonged and Exclusive Breast-Feeding on Dental Caries in Early School-Age Children. New Evidence from a Large Randomized Trial." *Caries Res* 41 (6): 484–488.
27. Kramer, Michael S., Frances Aboud, Elena Mironova, Irina Vanilovich, Robert W. Platt, Lidia Matush, Sergio Igumnov, Eric Fombonne, Natalia Bogdanovich, Thierry Ducruet, Jean-Paul Collet, Beverley Chalmers, Ellen Hodnett, Sergei Davidovsky, Oleg Skugarevsky, Oleg Trofimovich, Ludmila Kozlova, Stanley Shapiro, and the Promotion of Breastfeeding Intervention Trial Study Group. 2008a. "Breastfeeding and Child Cognitive Development: New Evidence from a Large Randomized Trial." *Archives General Psychiatry* 65 (5): 578–584.
28. Kramer, Michael S., Eric Fombonne, Sergei Igumnov, Irina Vanilovich, Lidia Matush, Elena Mironova, Natalia Bogdanovich, Richard E. Tremblay, Beverley Chalmers, Xun Zhang, and Robert W. Platt for the Promotion of Breastfeeding Intervention Trial (PROBIT) Study Group. 2008b. "Effects of Prolonged and Exclusive Breastfeeding on

Child Behavior and Maternal Adjustment: Evidence from a Large, Randomized Trial." *Pediatrics* 121 (3): e435–e440.
29. Levin, April R., Charles H. Zeanah, Nathan A. Fox, and Charles A. Nelson. 2014. "Motor Outcomes in Children Exposed to Early Psychosocial Deprivation." *Journal of Pediatrics* 164 (1): 123–129.
30. Levin, April R., Nathan A. Fox, Charles H. Zeanah, and Charles A. Nelson. 2015. "Social Communication Difficulties and Autism in Previously Institutionalized Children." *Journal of the American Academy of Child and Adolescent Psychiatry* 54 (2): 108–115.
31. Maluccio, John A., John Hoddinott, Jere R. Behrman, Reynaldo Martorell, Agnes R. Quisumbing, and Aryeh D. Stein. 2009. "The Impact of Nutrition during Early Childhood on Education among Guatemalan Adults." *Economic Journal* 119 (537): 643–975.
32. Manley, James, Lia Fernald, and Paul Gertler. 2012. "Wealthy, Healthy and Wise: Does Money Compensate for Being Born into Difficult Conditions?" Working Paper No. 2012–01, Towson University, Department of Economics, Towson, Md.
33. Martin, Richard M., Rita Patel, Michael S. Kramer, Lauren Guthrie, Konstantin Vilchuck, Natalia Bogdanovich, Natalia Sergeichick, Nina Gusina, Ying Foo, Tom Palmer, Sheryl L. Rifas-Shiman, Matthew W. Gillman, George D. Smith, and Emily Oken. 2013. "Effects of Promoting Longer-Term and Exclusive Breastfeeding on Adiposity and Insulin-Like Growth Factor-I at Age 11.5 Years." *JAMA* 309 (10): 1005–1013.
34. Martin, Richard M., Rita Patel, Michael S. Kramer, Konstantin Vilchuck, Natalia Bogdanovich, Natalia Sergeichick, Nina Gusina, Ying Foo, Tom Palmer, Jennifer Thompson, Matthew W. Gillman, George D. Smith, and Emily Oken. 2014. "Effects of Promoting Longer Term and Exclusive Breastfeeding on Cardiometabolic Risk Factors at Age 11.5 Years: A Cluster-Randomized, Controlled Trial." *Circulation* 129 (3): 321–329.
35. Martínez, Sebastián, Sophie Naudeau, and Vitor Pereira. 2012. *The Promise of Preschool in Africa: A Randomized Impact Evaluation of Early Childhood Development in Rural Mozambique.* Washington, DC: World Bank Group.
36. Ozier, Owen. 2014. "Exploiting Externalities to Estimate the Long-Term Effects of Early Childhood Deworming." Policy Research Working Paper No. WPS 7052, World Bank Group, Washington, DC.
37. Pathak, Yuvraj, and Karen Macours. 2013. "Women's Political Reservation, Early Childhood Development and Learning in India." Young Lives Working Paper 110. http://www.younglives.org.uk/publications/WP/women2019s-political-reservation-early-childhood-development-and-learning-in-india.
38. Pongcharoen, Tippawan. 2010. "Long-Term Effects of Iron and Zinc Supplementation during Infancy on Cognitive Performance and Growth 8 Years Later: A Follow-Up Study." Ph.D. dissertation, Emory University. http://pid.emory.edu/ark:/25593/5czdg.
39. Rackstraw, Emma. 2014. "A Decade Later: An Evaluation of the Longer-Term Impacts of a Honduran Conditional Cash Transfer." Honors thesis collection, Wellesley College. http://repository.wellesley.edu/thesiscollection/215.
40. Raine, Adrian, Peter H. Venables, Cyril Dalais, Kjetil Mellingen, Chandra Reynolds, and Sarnoff A. Mednick. 2001. "Early Educational and Health Enrichment at Age 3–5 Years Is Associated with Increased Autonomic and Central Nervous System Arousal and

Orienting at Age 11 Years: Evidence from the Mauritius Child Health Project." *Psychophysiology* 38 (2): 254–266.

41. Raine, Adrian, Kjetil Mellingen, Jianghong Liu, Peter H. Venables, and Sarnoff A. Mednick. 2003. "Effects of Environmental Enrichment at Ages 3–5 Years on Schizotypal Personality and Antisocial Behavior at Ages 17 and 23 Years." *American Journal of Psychiatry* 160 (9): 1627–1635.
42. Skugarevsky, Oleg, Kaitlin H. Wade, Rebecca C. Richmond, Richard M. Martin, Kate Tilling, Rita Patel, Konstantin Vilchuck, Natalia Bogdanovich, Natalia Sergeichick, George D. Smith, Matthew W. Gillman, Emily Oken, and Michael S. Kramer. 2014. "Effects of Promoting Longer-Term and Exclusive Breastfeeding on Childhood Eating Attitudes: A Cluster-Randomized Trial." *International Journal of Epidemiology* 43 (4): 1263–1271.
43. Smyke, Anna T., Charles H. Zeanah, Mary M. Gleason, Stacy S. Drury, Nathan A. Fox, Charles A. Nelson, and Donald Guthrie. 2012. "A Randomized Controlled Trial Comparing Foster Care and Institutional Care for Children with Signs of Reactive Attachment Disorder." *American Journal of Psychiatry* 169 (5): 508–514.
44. Spears, Dean, and Sneha Lamba. 2013. "Effects of Early-Life Exposure to Sanitation on Childhood Cognitive Skills: Evidence from India's Total Sanitation Campaign." Policy Research Working Paper No. WPS 6659, World Bank Group, Washington, DC.
45. Stewart, Christine P., Parul Christian, Steven C. LeClerq, Keith P. West Jr., and Subarna K. Khatry. 2009a. "Antenatal Supplementation with Folic Acid + Iron + Zinc Improves Linear Growth and Reduces Peripheral Adiposity in School-Age Children in Rural Nepal." *Journal of Nutrition* 90 (1): 132–140.
46. Stewart, Christine P., Parul Christian, Kerry J. Schulze, Steven C. Leclerq, Keith P. West Jr., and Subarna K. Khatry. 2009b. "Antenatal Micronutrient Supplementation Reduces Metabolic Syndrome in 6- to 8-Year-Old Children in Rural Nepal." *Journal of Nutrition* 139 (8): 1575–1581.
47. Todd, J.E., and P. Winters. 2011. "The Effect of Early Interventions in Health and Nutrition on On-Time School Enrollment: Evidence from the Oportunidades Program in Rural Mexico." *Economic Development and Cultural Change* 59 (3): 549–581.
48. Troller-Renfree, Sonya, Jennifer M. McDermott, Charles A. Nelson, Charles H. Zeanah, and Nathan A. Fox. 2014. "The Effects of Early Foster Care Intervention on Attention Biases in Previously Institutionalized Children in Romania." *Developmental Science*: 1–10. doi: 10.1111/desc.
49. Vanderwert, Ross E., Peter J. Marshall, Charles A. Nelson, Charles H. Zeanah, and Nathan A. Fox. 2010. "Timing of Intervention Affects Brain Electrical Activity in Children Exposed to Severe Psychosocial Neglect." *PLoS ONE* 5 (7): e11415.
50. Walker, S.P., S.M. Grantham-McGregor, J.H. Himes, C.A. Powell, and S.M. Chang. 1996. "Early Childhood Supplementation Does Not Benefit the Long-Term Growth of Stunted Children in Jamaica." *Journal of Nutrition* (126): 3017–3024.
51. Walker, S.P., S.M. Grantham-McGregor, C.A. Powell, and S.M. Chang. 2000. "Effects of Growth Restriction in Early Childhood on Growth, IQ, and Cognition at Age 11 to 12 Years and the Benefits of Nutritional Supplementation and Psychosocial Stimulation." *Journal of Pediatrics* 137 (1): 36–41.

52. Walker, S.P., S.M. Chang, C.A. Powell, and S.M. Grantham-McGregor. 2005. "Effects of Early Childhood Psychosocial Stimulation and Nutritional Supplementation on Cognition and Education in Growth-Stunted Jamaican Children: Prospective Cohort Study." *The Lancet* 366 (9499): 19–25.
53. Walker, S.P., S.M. Chang, C.A. Powell, E. Simonoff, and S.M. Grantham-McGregor. 2006. "Effects of Psychosocial Stimulation and Dietary Supplementation in Early Childhood on Psychosocial Functioning in Late Adolescence: Follow-Up of Randomised Controlled Trial." *The BMJ* 333 (7566): 472.
54. Walker, S.P., S.M. Chang, N. Younger, and S.M. Grantham-McGregor. 2010. "The Effect of Psychosocial Stimulation on Cognition and Behaviour at 6 Years in a Cohort of Term, Low-Birthweight Jamaican Children." *Developmental Medicine & Child Neurology* 52 (7): e148–e154.
55. Walker, S.P., S.M. Chang, M. Vera-Hernández, and S.M. Grantham-McGregor. 2011. "Early Childhood Stimulation Benefits Adult Competence and Reduces Violent Behavior." *Pediatrics* 127 (5): 849–857.
56. Windsor, Jennifer, Ana Moraru, Charles A. Nelson, Nathan A. Fox, and Charles H. Zeanah. 2013. "Effect of Foster Care on Language Learning at 8 Years: Findings from the Bucharest Early Intervention Project." *Journal of Child Language* 40 (3): 605–627.
57. Xu, Lixin Colin, and Jing Zhang. 2014. "Water Quality, Brawn, and Education: The Rural Drinking Water Program in China." Policy Research Working Paper No. WPS 7054, World Bank Group, Washington, DC.

Other Citations

Aguero, Jorge, Michael Carter, and Ingrid Woolard. 2007. "The Impact of Unconditional Cash Transfers on Nutrition: The South African Child Support Grant." International Poverty Center Working Paper 39, Brasilia, Brazil.

Alderman, H., J. Konde-Lule, I. Sebuliba, D. Bundy, and A. Hall. 2006: "Effect on Weight Gain of Routinely Giving Albendazole to Preschool Children during Child Health Days in Uganda: Cluster Randomized Controlled Trial." *The BMJ* 333 (7559): 122.

Anderson, Laurie M., Carolynne Shinn, Mindy T. Fullilove, Susan C. Scrimshaw, Jonathan E. Fielding, Jacques Normand, Vilma G. Carande-Kulis, and the Task Force on Community Preventive Services. 2003. "The Effectiveness of Early Childhood Development Programs: A Systematic Review." *American Journal of Preventive Medicine* 24 (3S): 32–46.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–1495.

Baker-Henningham, Helen, and Florencia Lopez Boo. 2010. "Early Childhood Stimulation Interventions in Developing Countries: A Comprehensive Literature Review." IZA Discussion Paper Series No. 5282, Institute for the Study of Labor (IZA), Bonn, Germany.

Barham, Tania. 2011. "A Healthier Start: The Effect of Conditional Cash Transfers on Neonatal and Infant Mortality in Rural Mexico." *Journal of Development Economics* 94 (1): 74–85

Barnett, W. S. 1992. "Benefits of Compensatory Preschool Education." *Journal of Human Resources* 27 (2): 279–312.

Barnett, W. S. 1996. "Lives in the Balance: Benefit-Cost Analysis of the Perry Preschool Program through Age 27." Monographs of the High/Scope Educational Research Foundation. Ypsilanti, MI: High/Scope Press.

Barnett, W. S. 2011. "Effectiveness of Early Educational Intervention." *Science* 333 (6045): 975–978.

Becker, Gary S. 1993. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago, Ill.: University of Chicago Press.

Berrueta-Clement, John R., Lawrence J. Schweinhart, W. Steven Barnett, Ann S. Epstein, and David P. Weikart. 1984. "Changed Lives: The Effects of the Perry Preschool Program on Youths through Age 19." Monographs of the High/Scope Educational Research Foundation. Ypsilanti, MI: High/Scope Press.

Bhutta, Zulfiqar A., Tahmeed Ahmed, Robert E. Black, Simon Cousens, Kathryn Dewey, Elsa Giugliani, Batool A. Haider, Betty Kirkwood, Saul S. Morris, H.P.S. Sachdev, and Meera Shekar for the Maternal and Child Undernutrition Study Group. 2008. "What Works? Interventions for Maternal and Child Undernutrition and Survival." *The Lancet* 371 (9610): 417–440.

Bhutta, Zulfiqar A., Jai K. Das, Arjumand Rizvi, Michelle F. Gaffrey, Neff Walker, Susan Horton, Patrick Webb, Anna Lartey, Robert E. Black, *The Lancet* Nutrition Interventions Review Group, and the Maternal and Child Nutrition Study Group. 2013. "Evidence-Based Interventions for Improvement of Maternal and Child Nutrition: What Can Be Done and at What Cost?" *The Lancet* 382 (9890): 452–477.

Bitler, Marianne P., Hilary W. Hoynes, and Thurston Domina. 2014. "Experimental Evidence on Distributional Effects of Head Start." NBER Working Paper Series No. 20434, National Bureau of Economic Research, Cambridge, Mass.

Black, Maureen M., and Kathryn G. Dewey. 2014. "Promoting Equity through Integrated Early Child Development and Nutrition Interventions." *Annals of the New York Academy of Sciences* 1308: 1–10.

Black, Robert E., Cesar G. Victora, Susan P. Walker, Zulfiqar A. Bhutta, Parul Christian, Mercedes De Onis, Majid Ezzati, Sally Grantham-McGregor, Joanne Katz, Ryenaldo Martorell, and Ricardo Uauy for the Maternal and Child Nutrition Study Group. 2013. "Maternal and Child Undernutrition and Overweight in Low-Income and Middle-Income Countries." *The Lancet* 382 (9890): 427–451.

Black, Robert E., Lindsay H. Allen, Zulfiqar A. Bhutta, Laura E. Caulfield, Mercedes de Onis, Majid Ezzati, Colin Mathers, and Juan Rivera for the Maternal and Child Undernutrition Study Group. 2008. "Maternal and Child Undernutrition: Global and Regional Exposures and Health Consequences." *The Lancet* 371 (9608): 243–260.

Bouguen, Adrien, Deon Filmer, Karen Macours, and Sophie Naudeau. 2013. "Impact Evaluation of Three Types of Early Childhood Development Interventions in Cambodia." Policy Research Working Paper No. WPS 6540, World Bank Group, Washington, DC.

Bradley, Robert H., and Robert F. Corwyn. 2002. "Socioeconomic Status and Child Development." *Annual Review of Psychology* 53: 371–399.

Brooks-Gunn, Jeanne, and Greg J. Duncan. 1997. "The Effects of Poverty on Children." *Children and Poverty* 7 (2): 55.

Camilli, Gregory, Sadako Vargas, Sharon Ryan, and W. Steven Barnett. 2010. "Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development." *Teachers College Record* 112 (3): 579–620.

Campbell, Frances A., Craig T. Ramey, Elizabeth Pungello, Joseph Sparling, and Shari Miller-Johnson. 2002. "Early Childhood Education: Young Adult Outcomes from the Abecedarian Project." *Applied Developmental Science* 6 (1): 42–57.

Campbell, Frances A., Gabriella Conti, James J. Heckman, Seong Hyeok Moon, Rodrigo Pinto, Elizabeth Pungello, and Yi Pan. 2014. "Early Childhood Investments Substantially Boost Adult Health." *Science* 343: 1478–1485.

Cameron, Claire E., Laura L. Brock, William M. Murrah, Lindsay H. Bell, Samantha L. Worzalla, David Grissmer, and Frederick J. Morrison. 2012. "Fine Motor Skills and Executive Function Both Contribute to Kindergarten Achievement." *Child Development* 83 (4): 1229–1244.

Carneiro, Pedro, and James Joseph Heckman. 2003. "Human Capital Policy." IZA Discussion Paper Series No. 821, Institute for the Study of Labor (IZA), Bonn, Germany.

Case, Anne, and Christina Paxson. 2006. "Stature and Status: Height, Ability, and Labor Market Outcomes." NBER Working Paper Series No. 12466, National Bureau of Economic Research, Cambridge, Mass.

Ceesay, Sana M., Andrew M. Prentice, Timothy J. Cole, Frances Foord, Lawrence T. Weaver, Elizabeth M.E. Poskitt, and Roger G. Whitehead. 1997. "Effects on Birth Weight and Perinatal Mortality of Maternal Dietary Supplements in Rural Gambia: 5 Year Randomised Controlled Trial." *The BMJ* 315 (7111): 786–790.

Center on the Developing Child, Harvard University. "Science of Adversity and Resilience." http://developingchild.harvard.edu/activities/sar/.

Couperus, Jane W., and Charles A. Nelson. 2006. "Early Brain Development and Plasticity." In *Blackwell Handbook of Early Childhood Development,* edited by K. McCartney and D. Phillips. Oxford, UK: Blackwell Publishing Ltd.

Croke, Kevin. 2014. "The Long Run Effects of Early Childhood Deworming on Literacy and Numeracy: Evidence from Uganda." http://scholar.harvard.edu/files/kcroke/files/ug_lr_deworming_071714.pdf.

Cunha, Flavio, and James Heckman. 2009. "The Economics and Psychology of Inequality and Human Development." *Journal of the European Economic Association* 7 (2–3): 320–364.

Currie, Janet, and Duncan Thomas. 1998. "School Quality and the Longer-Term Effects of Head Start." NBER Working Paper No. 6362, National Bureau of Economic Research, Cambridge, Mass.

Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." American Economic Journal: Applied Economics 1 (3): 111–134.

Donohue, John, and Peter Siegelman. 1998. "Allocating Resources among Prisons and Social Programs in the Battle against Crime." *Journal of Legal Studies* 27 (1): 1–43.

Denboba, Amina D., Rebecca K. Sayre, Quentin T. Wodon, Leslie K. Elder, Laura B. Rawlings, and Joan Lombardi. 2014. *Stepping up Early Childhood Development: Investing in Young Children for High Returns*. Washington, DC: World Bank Group.

Duncan, Greg J., Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, Leon Feinstein, Mimi Engel, Jeanne Brooks-Gunn, Holly Sexton, Kathryn Duckworth, and Crista Japel. 2007. "School Readiness and Later Achievement." *Developmental Psychology* 43 (6): 1428–1446.

Ebbeling, Cara B., Dorota B. Pawlak, and David S. Ludwig. 2002. "Childhood Obesity: Public-Health Crisis, Common Sense Cure." *The Lancet* 360 (9331): 473–482.

Engle, Patrice L., Maureen M. Black, Jere R. Behrman, Meena Cabral de Mello, Paul J. Gertler, Lydia Kapiriri, Reynaldo Martorell, Mary Eming Young, and the International Child Development Steering Group. 2007. "Strategies to Avoid the Loss of Developmental Potential in More Than 200 Million Children in the Developing World." *The Lancet* 369: 229–242.

Engle, Patrice L., Lia C.H. Fernald, Harold Alderman, Jere Behrman, Chloe O'Gara, Aisha Yousafzai, Meena Cabral de Mello, Melissa Hidrobo, Nurper Ulkuer, Ilgi Ertem, Semil Iltus, and the Global Child Development Steering Group. 2011. "Strategies for Reducing Inequalities and Improving Developmental Outcomes for Young Children in Low-Income and Middle-Income Countries." *The Lancet* 378 (799): 1339–1353.

Evans, G.W., C. Gonnella, L.A. Marcynyszyn, L. Gentile, and N. Salpekar. 2005. "The Role of Chaos in Poverty and Children's Socioemotional Adjustmen." *Psychological Science* 16: 560–565.

Evans, G.W., D. Li, and S.S. Whipple. 2013. "Cumulative Risk and Child Development." *Psychological Bulletin* 139 (6): 1342–1396.

Fernald, Lia C.H., A. Weber, E. Galasso, and L. Ratsifandrihamanana. 2011. "Socioeconomic Gradients and Child Development in a Very Low Income Population: Evidence from Madagascar." *Developmental Science* 14 (4): 832–847.

Fernald, Lia C.H., Paul J. Gertler, and Lynnette M. Neufeld. 2008. "Role of Cash in Conditional Cash Transfer Programmes for Child Health, Growth, and Development: An Analysis of Mexico's Oportunidades." *The Lancet* 378 (799): 1339–1353.

Figueiras, Amira Consuelo, Isabel Cristina Neves de Souza, Viviana Graciela Rios, and Yehuda Benguigui. 2012. *Monitoring Child Development (0–6 Years) in the IMCI Context*. Second edition. Washington, DC: Pan American Health Organization.

Fox, S.E., P. Levitt, and C.A. Nelson. 2010. "How the Timing and Quality of Early Experiences Influence the Development of Brain Architecture." *Child Development* 81: 28–40.

Georgieff, M., and R. Rao. 1999. "The Role of Nutrition in Cognitive Development." In *Handbook of Developmental Cognitive Neuroscience,* edited by C. A. Nelson and M. Luciana, 491–504. Cambridge, Mass.: MIT Press.

Georgieff, Michael K. 2007. "Nutrition and the Developing Brain: Nutrient Priorities and Measurement." *American Journal of Clinical Nutrition* 85 (2): 614S–620S.

Gertler, Paul. 2004. "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment." *American Economic Review Papers and Proceedings* 94 (2): 336–341.

Ghera, Melissa M., Peter J. Marshall, Nathan A. Fox, Charles H. Zeanah, Charles A. Nelson, Anna T. Smyke, and Donald Guthrie. 2009. "The Effects of Foster Care Intervention on Socially Deprived Institutionalized Children's Attention and Positive Affect: Results from the BEIP Study." *Journal of Child Psychology and Psychiatry* 50 (3): 246–253.

Glewwe, Paul, Hanan G. Jacoby, and Elizabeth M. King. 2001. "Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis." *Journal of Public Economics* 81 (3): 345–368.

Gorey, Kevin M. 2001. "Early Childhood Education: A Meta-Analytic Affirmation of the Short- and Long-Term Benefits of Educational Opportunity." *School Psychology Quarterly* 16 (1).

Grantham-McGregor, Sally. 1995. "A Review of Studies of the Effect of Severe Malnutrition on Mental Development." *Journal of Nutrition* 125 (Suppl. 8): 2233–2238.

Grantham-McGregor, S.M., C.A. Powell, S.P. Walker, and J.H. Himes. 1991. "Nutritional Supplementation, Psychosocial Stimulation, and Mental Development of Stunted Children: The Jamaican Study." *The Lancet* 338 (8758): 1–5.

Grantham-McGregor, S., and H. Baker-Henningham. 2005. "Review of the Evidence Linking Protein and Energy to Mental Development." *Public Health and Nutrition* 8 (7A): 1191–1201.

Grantham-McGregor, Sally, Yin Bun Cheung, Santiago Cueto, Paul Glewwe, Linda Richer, Barbara Strupp, and the International Child Development Steering Group. 2007. "Developmental Potential in the First 5 Years for Children in Developing Countries." *The Lancet* 369 (9555): 60–70.

Grantham-McGregor, Sally M., Lia C.H. Fernald, Rose M.C. Kagawa, and Susan Walker. 2014. "Effects of Integrated Child Development and Nutrition Interventions on Child Development and Nutritional Status." *Annals of the New York Academy of Sciences* 1308: 11–32.

Grissmer, David, Kevin J. Grimm, Sophie M. Aiyer, William M. Murrah, and Joel S. Steele. 2010. "Fine Motor Skills and Early Comprehension of the World: Two New School Readiness Indicators." *Developmental Psychology* 46 (5): 1008–1017.

Hackman, Daniel A., and Martha J. Farah. 2009. "Socioeconomic Status and the Developing Brain." *Trends in Cognitive Sciences* 13 (2): 65–73.

Hamill, P.V.V., T.A. Drizd, C.L. Johnson, R.B. Reed, A.F. Roche, and W.M. Moore. 1979. "Physical Growth: National Center for Health Statistics Percentiles." *American Journal of Clinical Nutrition* 32: 607–629.

Hart, Betty, and Todd R. Risley. 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore, Md.: Paul H. Brookes Publishing.

Heckman, James. 2008. "Schools, Skills, and Synapses." *Economic Inquiry* 46 (3): 289.

HHS (U.S. Department of Health and Human Services). 2005. *Head Start Impact Study: First Year Findings.* Washington, DC: HHS.

______. 2010. *Head Start Impact Study: Final Report.* Washington, DC: HHS.

IEG (Independent Evaluation Group). 2013. *Delivering the Millennium Development Goals to Reduce Maternal and Child Mortality: A Systematic Review of Impact Evaluation Evidence.* Washington, DC: World Bank.

______. 2015. *World Bank Support to Early Childhood Development: An Independent Evaluation.* Washington, DC: World Bank.

Irwin, L. G., A. Siddiqi, and C. Hertzman. 2007. *Early Child Development: A Powerful Equalizer: Final Report for the World Health Organization's Commission on Social Determinants of Health*. Vancouver, Canada: Human Early Learning Partnership.

Johnson, Dana E., Donald Guthrie, Anna T. Smyke, Sebastian F. Koga, Nathan A. Fox, Charles H. Zeanah, and Charles A. Nelson. 2010. "Growth and Associations Between Auxology, Caregiving Environment, and Cognition in Socially Deprived Romanian Children Randomized to Foster vs. Ongoing Institutional Care." *Archives of Pediatrics and Adolescent Medicine* 164 (6): 507–516.

Johnson, M. H. 1998. "The Neural Basis of Cognitive Development." In *Handbook of Child Psychology*, edited by D. William, 1–49. Hoboken, N.J.: John Wiley and Sons Inc.

Kagitcibasi, Cigdem, Diane Sunar, Bekman, Nazli Baydar, and Zeynep Cemalcilar. 2009. "Continuing Effects of Early Enrichment in Adult Life: The Turkish Early Enrichment Project 22 Years Later." *Journal of Applied Developmental Psychology* 30 (6): 764–779.

Kalil, A., H. Yoshikawa, and K.M. Ziol-Guest. 2014. "Developmental Psychology and Poverty in Global Contexts: The Role of the Family." In *Barriers to and Opportunities for Poverty Reduction: Prospects for Private Sector-Led Interventions,* 22–37. New York: United Nations Development Programme.

Kessler, R. C., Katie A. McLaughlin, Jennifer Greif Green, Michael J. Gruber, Nancy A. Sampson, Alan M. Zaslavsky, Sergio Aguilar-Gaxiola, Ali Obaid Alhamzawi, Jordi Alonso, Matthias Angermeyer, Corina Benjet, Evelyn Bromet, Somnath Chatterji, Giovanni de Girolamo, Koen Demyttenaere, John Fayyad, Silvia Florescu, Gilad Gal, Oye Gureje, Josep Maria Haro, Chi-yi Hu, Elie G. Karam, Norito Kawakami, Sing Lee, Jean-Pierre Lépine, Johan Ormel, José Posada-Villa, Rajesh Sagar, Adley Tsang, T. Bedirhan Üstün, Svetlozar Vassilev, Maria Carmen Viana, and David R. Williams. 2010. "Childhood Adversities and Adult Psychopathology in the WHO World Mental Health Surveys." *The British Journal of Psychiatry* 197 (5): 378–385.

Kramer, Michael S. 2010. "'Breast Is Best'": The Evidence." *Early Human Development* 86 (11): 729–732.

Kramer, Michael S., Beverley Chalmers, Ellen D. Hodnett, Zinaida Sevkovskaya, Irina Dzikovich, Stanley Shapiro, Jean-Paul Collet, Irina Vanilovich, Irina Mezen, Thierry

Ducruet, George Shishko, Vyacheslav Zubovich, Dimitri Mknuik, Elena Gluchanina, Viktor Dombrovskiy, Anatoly Ustinovitch, Tamara Kot, Natalia Bogdanovich, Lydia Ovchinikova, Elisabet Helsing, and the PROBIT Study Group. 2001. "Promotion of Breastfeeding Intervention Trial (PROBIT): A Randomized Trial in the Republic of Belarus." *Journal of the American Medical Association* 285 (4): 413–420.

Lake, Anthony. 2012. "Investing in Nutrition Security Is Key to Sustainable Development." In *The G8 Camp David Summit 2012: The Road to Recovery,* edited by John Kirton and Madeline Koch, 135–139. London: Newsdesk Media Group.

Leroy, Jef L., Armando García-Guerra, Raquel García, Clara Dominguez, Juan Rivera, and Lynnette M. Neufeld. 2008. "The Oportunidades Program Increases the Linear Growth of Children Enrolled at Young Ages in Urban Mexico." *Journal of Nutrition* 138 (4): 793–798.

Leroy, Jef L., Marie Ruel, Jean-Pierre Habicht, and Edward A. Frongillo. 2014. "Linear Growth Deficit Continues to Accumulate beyond the First 1000 Days in Low- and Middle-Income Countries: Global Evidence from 51 National Surveys." *Journal of Nutrition* 144 (9): 1460–1466.

Maternal and Child Nutrition Study Group. 2013. *Executive Summary of The Lancet Maternal and Child Nutrition Series*. Oxford, UK: Elsevier. http://download.thelancet.com/flatcontentassets/pdfs/nutrition-eng.pdf.

Magnuson, Katherine A., Christopher Ruhm, and Jane Waldfogel. 2007. "Does Prekindergarten Improve School Preparation and Performance?" *Economics of Education Review* 26 (1): 33–51.

Martorell, Reynaldo, Robert E. Klein, and Hernan Delgado. 1980. "Improved Nutrition and Its Effects on Anthropometric Indicators of Nutritional Status." *Nutritional Reports International* 21 (2): 219–230.

Martorell, Reynaldo, Jean-Pierre Habicht, and Juan A. Rivera. 1995. "History and Design of the INCAP Longitudinal Study (1969–77) and Its Follow-Up (1988–89)." *Journal of Nutrition* 125 (Suppl. 5): 1027S–1041S.

Mendez, Michelle A., and Linda S. Adair. 1999. "Severity and Timing of Stunting in the First Two Years of Life Affect Performance on Cognitive Tests in Late Childhood." *Journal of Nutrition* 129 (8): 1555–1562.

Naudeau, S., S. Martínez, P. Premand, and D. Filmer. 2011. "Cognitive Development among Young Children in Low-Income Countries." In *No Small Matter,* edited by H. Alderman, 9–50. Washington, DC: World Bank.

Nelson, Charles A., Charles H. Zeanah, Nathan A. Fox, Peter J. Marshall, Anna T. Smyke, and Donald Guthrie. 2007. "Cognitive Recovery in Socially Deprived Young Children: The Bucharest Early Intervention Project." *Science* 318 (5858): 1937–1940.

Ngure, F. M., B. M. Reid, J. H. Humphrey, M. N. Mbuya, G. Pelto and R. J. Stoltzfus. 2014. "Water, Sanitation, and Hygiene (WASH), Environmental Enteropathy, Nutrition, and Early Child Development: Making the Links." *Annals of the New York Academy of Sciences* 1308: 118–128.

NRC (National Research Council) and IOM (Institute of Medicine). 2000. From *Neurons to Neighborhoods: The Science of Early Childhood Development,* edited by Jack P. Shonkoff and Deborah A. Phillips. Washington, DC: National Academy Press.

Noble, Kimberly G., Suzanne M. Houston, Eric Kan, and Elizabeth R. Sowell. 2012. "Neural Correlates of Socioeconomic Status in the Developing Human Brain." *Developmental Science* 15 (4): 516–527.

Nores, Milagros, and W. Steven Barnett. 2010. "Benefits of Early Childhood Interventions across the World: (Under) Investing in the Very Young." *Economics of Education Review* 29 (2): 271–282.

Olds, D.L., L. Sadler, and H. Kitzman. 2007. "Programs for Parents of Infants and Toddlers: Recent Evidence from Randomized Trials." *Journal of Child Psychology and Psychiatry* 48 (3–4): 355–391.

Pechtel, Pia, and Diego A. Pizzagalli. 2011. "Effects of Early Life Stress on Cognitive and Affective Function: An Integrated Review of Human Literature." *Psychopharmacology* 214 (1): 55–70.

Prendergast, Andrew J., and Jean H. Humphrey. 2014. "The Stunting Syndrome in Developing Countries." *Paediatrics and International Child Health* 34 (4): 250–265.

Prentice, Andrew M. 2006. "The Emerging Epidemic of Obesity in Developing Countries." *International Journal of Epidemiology* 35: 93–99.

Prentice, A.M., S.B. Roberts, A. Prentice, A.A. Paul, M. Watkinson, A.A. Watkinson, and R.G. Whitehead. 1983. "Dietary Supplementation of Lactating Gambian Women. I. Effect on Breast-Milk Volume and Quality." *Hum Nutr Clin Nutr* 37 (1): 53–64.

Raizada, Rajeev D.S., and Mark M. Kishiyama. 2010. "Effects of Socioeconomic Status on Brain Development, and How Cognitive Neuroscience May Contribute to Levelling the Playing Field." *Frontiers in Human Neuroscience* 4:3.

Ramey, Craig T, and Sharon Landesman Ramey. 1998. "Prevention of Intellectual Disabilities: Early Interventions to Improve Cognitive Development." *Preventive Medicine* 27 (2): 224–232.

Reynolds, Arthur J., Judy A. Temple, Dylan L. Robertson, and Emily A. Mann. 2001. "Long-Term Effects of an Early Childhood Intervention on Educational Achievement and Juvenile Arrest: A 15-Year Follow-Up." *Journal of American Medical Association* 285 (18): 2339–2346.

Reynolds, Arthur J., Judy A. Temple, Suh-Ruu Ou, Dylan L. Robertson, Joshua P. Mersky, James W. Topitzes, and Michael D. Niles. 2007. "Effects of a School-Based, Early Childhood Intervention on Adult Health and Well-Being: A 19-Year Follow-Up of Low-Income Families." *Archives of Pediatrics & Adolescent Medicine* 161 (8): 730–739.

Reynolds, Arthur J., and J. A. Temple. 2008. "Cost-Effective Early Childhood Development Programs from Preschool to Third Grade." *Annual Review of Clinical Psychology* 4: 109–139.

Rivera, Juan A., Daniela Sotres-Alvarez, Jean-Pierre Habicht, Teresa Shamah, Salvador Villalpando. 2004. "Impact of the Mexican Program for Education, Health, and Nutrition

(Progresa) on Rates of Growth and Anemia in Infants and Young Children: A Randomized Effectiveness Study." *Journal of the American Medical Association* 291 (21): 2563–2570.

Schady, Norbert, Jere Behrman, Maria Caridad Araujo, Rodrigo Azuero, Raquel Bernal, David Bravo, Florencia Lopez-Boo, Karen Macours, Daniela Marshall, Christina Paxson, and Renos Vakis. 2014. "Wealth Gradients in Early Childhood Cognitive Development in Five Latin American Countries." IDB Working Paper Series No. IDB-WP-482, Inter-American Development Bank, Washington, DC.

Schilling, Elizabeth A., Robert H. Aseltine, and Susan Gore. 2007. "Adverse Childhood Experiences and Mental Health in Young Adults: A Longitudinal Survey." *BMC Public Health* 7 (1): 30.

Schweinhart, L.J. 2007. "Outcomes of the High/Scope Perry Preschool Study and Michigan School Readiness Program." In *Early Childhood Development: From Measurement to Action,* edited by Mary Eming Young with Linda M. Richardson, 87–102. Washington, DC: World Bank.

Schweinhart, L.J., H.V. Barnes, and D.P. Weikart. 1993. *Significant Benefits: The High/Scope Perry Preschool Study through Age 27.* Ypsilanti, Mich.: High/Scope Press.

Schweinhart, L. J., J. Montie, Z. Xiang, W.S. Barnett, C.R. Belfield, and M. Nores. 2005. *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40.* Ypsilanti, Mich.: High/Scope Press.

Shrimpton, Roger, Cesar G. Victora, Mercedes de Onis, Rosângela Costa Lima, Monika Blössner, and Graeme Clugston. 2001. "Worldwide Timing of Growth Faltering: Implications for Nutritional Interventions." *Pediatrics* 107 (5): e75.

Shonkoff, J.P. 2011. "Protecting Brains, Not Simply Stimulating Minds." *Science* 333 (6045): 982–983.

Shonkoff, J.P., W.T. Boyce, and B.S. McEwen. 2009. "Neuroscience, Molecular Biology, and the Childhood Roots of Health Disparities." *Journal of the American Medical Association* 301 (21): 2252–2259.

Shonkoff, J.P., A.S. Garner, B.S. Siegel, M.I. Dobbins, M.F. Earls, L. Mcguinn, J. Pascoe, and D.L. Wood. 2012a. "The Lifelong Effects of Early Childhood Adversity and Toxic Stress." *Pediatrics* 129: e232–e246.

Shonkoff, J. P., L. Richter, J. Van Der Gaag, and Z.A. Bhutta. 2012b. "An Integrated Scientific Framework for Child Survival and Early Childhood Development." *Pediatrics* 129: e460–e472.

Slopen, N., K.C. Koenen, and L.D. Kubzansky. 2014. "Cumulative Adversity in Childhood and Emergent Risk Factors for Long-Term Health." *The Journal of Pediatrics* 164: 631–638.

Smith, James P. 2009. "The Impact of Childhood Health on Adult Labor Market Outcomes." *Review of Economics and Statistics* 91 (3): 478–489.

Sweet, Monica A., and Mark I. Appelbaum. 2004. "Is Home Visiting an Effective Strategy? A Meta-Analytic Review of Home Visiting Programs for Families with Young Children." *Child Development* 75 (5): 1435–1456.

Tanner, Jeffery. 2012. "Smart Choices: Cognitive, Behavioral, and Physical Health Impacts on Children Left Behind by Paternal Migration" in International Labor Flows: Migration Views from the Migrant, the Receiving-Country Economy, and the Sending-Country Family." Ph.D. dissertation, Pardee RAND Graduate School, Santa Monica, Calif. http://www.rand.org/content/dam/rand/pubs/rgs_dissertations/2012/RAND_RGSD303.pdf.

UNICEF (United Nations Children's Fund). 2007. Progress for Children: A World Fit for Children Statistical Review. No. 6. New York: UNICEF. http://www.unicef.org/progressforchildren/2007n6/files/Progress_for_Children_-_No._6.pdf.

______. 2013. "Improving Child Nutrition—the Achievable Imperative for Global Progress." New York: UNICEF.

Victora, Cesar G., Linda Adair, Caroline Fall, Pedro C. Hallal, Reynaldo Martorell, Linda Richter, and Harshpal Singh Sachdev. 2008. "Maternal and Child Undernutrition: Consequences for Adult Health and Human Capital." *The Lancet* 371 (9609): 340–357.

Wachs, Theodore D., Michael Georgieff, Sarah Cusick, and Bruce S. McEwen. 2014. "Issues in the Timing of Integrated Early Interventions: Contributions from Nutrition, Neuroscience, and Psychological Research." *Annals of the New York Academy of Sciences* 1308: 89–106.

Waddington, Hugh, Howard White, Birte Snilstveit, Jorge Garcia Hombrados, Martina Vojtkova, Philip Davies, Ami Bhavsar, John Eyers, Tracey Perez Koehlmoos, Mark Petticrew, Jeffrey C. Valentine, and Peter Tugwell. 2012. "How to Do a Good Systematic Review of Effects in International Development: A Tool Kit." *Journal of Development Effectiveness* 4 (3): 359–387.

Wagstaff, Adam. 2011. "Fungibility and the Impact of Development Assistance: Evidence from Vietnam's Health Sector." *Journal of Development Economics* 94 (1): 62–73.

Waldfogel, Jane, and Elizabeth Washbrook. 2011. "Early Years Policy." *Child Development Research,* vol. 2011, Article ID 343016, 12 pages, 2011. doi:10.1155/2011/343016.

Walker, Dale, Charles Greenwood, Betty Hart, and Judith Carta. 1994. "Prediction of School Outcomes Based on Early Language Production and Socioeconomic Factors." *Child Development* 65 (2): 606–621.

Walker, Susan P., Susan M. Chang, Christine A. Powell, and Sally M. Grantham-McGregor. 2004. "Psychosocial Intervention Improves the Development of Term and Low-Birth-Weight Infants." *The Journal of Nutrition* 134: 1417-1423.

Walker, Susan.P., Christine A. Powell, Sally M. Grantham-McGregor, John H. Himes, and Susan M. Chang. 1991. "Nutritional Supplementation, Psychosocial Stimulation, and Growth of Stunted Children: The Jamaican Study." *American Journal of Clinical Nutrition* 54 (4): 642–648.

Walker, Susan P., Theodore D. Wachs, Julie Meeks Gardner, Betsy Lozoff, Gail A. Wasserman, Ernesto Pollitt, Julie A. Carter, and the International Child Development Steering Group. 2007. "Child Development: Risk Factors for Adverse Outcomes in Developing Countries." *The Lancet* 369 (9556): 145–157.

Walker, Susan P., Theodore D. Wachs, Sally Grantham-McGregor, Maureen M. Black, Charles A. Nelson, Sandra L. Huffman, Helen Baker-Henningham, Susan M. Chang, Jena D. Hamadani, Betsy Lozoff, Julie M. Meeks Gardner, Christine A. Powell, Atif Rahman, and Linda Richter. 2011. "Inequality in Early Childhood: Risk and Protective Factors for Early Childhood Development." *The Lancet* 378 (9799): 1325–1338.

Wechsler, D. 1991. Wechsler Intelligence Scale for Children—Third Edition. San Antonio, TX: The Psychological Corporation.

WHO (World Health Organization). 2006. "Who Child Growth Standards Based on Length/Height, Weight and Age." *Acta Paediatrica Supplement* 450: 76–85.

_______. 2014. Global Database on Child Growth and Malnutrition: Child Growth Indicators and Their Interpretation. Available at http://www.who.int/nutgrowthdb/about/introduction/en/index2.html.

Windsor, Jennifer, Joann P. Benigno, Christine A. Wing, Patrick J. Carroll, Sebastian F. Koga, Charles A. Nelson, Nathan A. Fox, and Charles H. Zeanah. 2011. "Effect of Foster Care on Young Children's Language Learning." *Child Development* 82 (4): 1040–1046.

Woolcock, Michael. 2013. "Using Case Studies to Explore the External Validity of 'Complex' Development Interventions." *Evaluation* 19 (3): 229–248.

World Bank. 2005. *World Development Report 2006: Equity and Development*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/5988 License: CC BY 3.0 IGO.

———. 2010. *Stepping up Skills for More Jobs and Higher Productivity*. Washington, DC: World Bank.

———. 2011. *Impact Evaluation in Practice*. Washington, DC: World Bank.

———. 2013. *World Bank Goals: End Extreme Poverty and Promote Shared Prosperity*. Washington, DC: World Bank.

———. 2015. *World Development Report 2015: Mind, Society, and Behavior.* Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/20597 License: CC BY 3.0 IGO.

Yoshikawa, Hirokazu. 1994. "Prevention as Cumulative Protection: Effects of Early Family Support and Education on Chronic Delinquency and Its Risks." *Psychological Bulletin* 115 (1): 28.

Yousafzai, Aisha K., Muneera A. Rasheed, and Zulfiqar A. Bhutta. 2013. "Annual Research Review: Improved Nutrition—a Pathway to Resilience." *Journal of Child Psychology and Psychiatry* 54 (4): 367–377.

Zigler, Edward, and Sally J. Styfco. 1994. "Is the Perry Preschool Better than Head Start? Yes and No." *Early Childhood Research Quarterly* 9 (3–4): 269–287.

Zigler, Edward, Cara Taussig, and Kathryn Black. 1992. "Early Childhood Intervention: A Promising Preventative for Juvenile Delinquency." *American Psychologist* 47 (8): 997.

Appendix A. Descriptions of Interventions from Included Studies

Country, *Name of Intervention*	Study	Description	Counterfactual	Age at Intervention	Age at Evaluation	Outcomes
Argentina *Preschool*	Berlinski, Galiani, and Gertler 2009 [9]	In 1993, the government of Argentina made preprimary school compulsory and began a massive public school construction program. Given the need to build so many new schools, the preprimary educational requirement was phased in over six years. Children are eligible for up to 3 years of public preprimary education. Length of exposure: 1–3 years Scale: National	Children in areas where schools had not yet been built	3–5 years old	8–9 years old	Students paying attention in class**; Students putting a large amount of effort into understanding explanations**; Students regularly participating in class**; Students well-disciplined in classroom; Mathematics test**; Language test**
Bangladesh *Matlab Family Planning Program*	Barham 2012 [5]	Started in 1977. Local female health workers made monthly home visits to provide free contraception and advice on contraceptives, nutrition, hygiene and breastfeeding. Women were also eligible for a tetanus toxoid vaccine, folic acid and iron during the last trimester of pregnancy. Starting in 1982, half of the treatment communities received child health programs as well, in particular measles vaccines. The other half of the treatment areas began receiving it in 1985. Starting in 1986, additional child health interventions, including immunizations, vitamin A supplementation for children under five, and nutritional rehabilitation for those who were nutritionally at risk were added. All interventions were administered in the home of the recipient. Control areas continued to receive government healthcare. In this study, children 8–11 years old would have benefited from all of the child health programs as well as the effects of the family	Children in control areas	8–19 year olds: Birth 20–24 year olds: 20 years old	8–24 years old	Attention-concentration (8–14 years old **, 15–19 years old, 20–24 years old); Highest grade completed (8–14 years old***, 15–19 years old, 20–24 years old); Language test (8–14 years old, 15–19 years old; 20–24 years old); Orientation (8–14 years old**, 15–19 years old, 20–24 years old); Recall (8–14 years old, 15–19 years old, 20–24 years old); Registration (8–14 years old ***, 15–19 years old, 20–24 years old)

		planning. Children 12–14 would have been eligible for vaccinations and the benefits of family planning, those 15–19 would have benefits from the family planning program, and those 20–24 were born after the program began but could still benefit from smaller family sizes through the family planning program. Length of exposure: Varies Scale: Approximately 100,000 people				
Belarus *Promotion of Breastfeeding Intervention Trial (PROBIT)*	Kramer and others 2007a [24] Kramer and others 2007b [25] Kramer and others 2007c [26] Kramer and others 2008a [27] Kramer and others 2008b [28]	1996–1997. Thirty-two pairs of maternity hospitals and associated polyclinics from across Belarus were randomly assigned to participate, and women who expressed, upon admission to the postpartum ward, an intentions to breastfeed and had given birth to a full-term, normal birth weight infant were eligible to participate. Midwives, nurses and doctors were trained in lactation and instructed women in the treatment group on the benefits of exclusive breastfeeding and proper duration. Length of exposure: 1 year Scale: 8547 mother-infant pairs	Children whose mothers attended the maternity hospitals or polyclinics that were randomized to not receive the program	<1 year old	6.5 years old	Allergies (overall, Alternaria, birch pollen, cat, house dust mite, northern grasses); Block designs; BMI; Conduct problems; Diastolic blood pressure; Asthma; Eczema; Emotional symptoms; Hay fever symptoms; Wheezing; Externalizing behavior; Full-scale IQ*; Hay fever; Head circumference*; Height; Hip circumference; Hyperactivity; Incisor DMFT; Internalizing behavior; Leg length; Mathematics test; Mid-upper-arm circumference; Midthigh circumference**; Peer problems; Performance IQ; Incisor DMFT ≥ 1; DMFT ≥ 1; Incisor DMFT ≥ 2; DMFT ≥ 2; Prosocial behavior; Raven's matrices; Reading*; Recurrent itchy rash; Similarities**; Subscapular skinfold thickness; Triceps skinfold thickness; Strength and difficulties questionnaire; Systolic blood pressure; Total DMFT; Number of teeth; Teeth with cavities; Teeth with fillings; Verbal IQ**; Vocabulary**; Waist circumference; Waist:hip ratio; Writing
	Martin and others 2013 [33] Martin and				11.5 years old	Adiponectin; Apolipoprotein A1; BMI; BMI (≥ 85th percentile**, ≥ 95th percentile*); Body fat; Diastolic blood pressure; Dieting**; Eating attitudes**; Fasting insulin; FFMI; FMI;

	others 2014 [34] Skugarevsky and others 2014 [42]					Food preoccupation; Glucose; Head circumference; Height; Hip circumference; IGF-I; Leg length; Mid-upper-arm circumference; Metabolic syndrome; Pressure from others**; Purging and vomiting*; Restriction and control; Subscapular skinfold thickness; Triceps skinfold thickness; Waist circumference; Waist:hip ratio; Weight preoccupation**
Chile *Early Childhood Care and Education (ECCE)*	Cortázar Valdés 2011 [13]	Ongoing program. Center-based pre-kindergarten programs for 2–4 year olds, targeted to those from the two lowest income quintiles. The program included educational activities involving the child, educator and family, as well as a feeding program and health component. The family involvement included promoting home-based activities, inviting parents to participate in classrooms and helping to organize extracurricular activities. Length of exposure: 1–3 years Scale: National	Children who did not participate in any pre-kindergarten education program	2–4 years old	Fourth graders (about 10 years old)	Mathematics test**; Reading***; Social science***
China *Rural Drink Water Program*	Xu and Zhang 2014 [57]	1980s–present. Built water treatment plants and pipelines to provide clean drinking water to rural residents. The central government set general standards, and local governments were in charge of program implementation. Length of exposure: Varies Scale: National	Children who lived in municipalities that built treatment plants after the children turned 6 years old	0–2 years old or 3–5 years old	18–25 years old	Years of schooling (0–2 years old***, 3–5 years old)
Colombia *Hogares Comunitarios*	Attanasio and Vera-Hernández 2004 [4]	Initiated from 1984–1986. Parents of children 0–6 years old from poor households in poor neighborhoods and localities were encouraged to form ‘parents associations,’ which elected a madre comunitaria (community mother). This	Distance to nearest hogar comunitario and average distance for the town were used as	0–6 years old	8–17 years old	Probability of attending primary school; Probability of attending secondary school**; Probability of passing a grade (8–12 years old, 13–17 years old*)

		mother had to have at least a basic education and a large enough house to host up to 15 children. The parents would then pay a small monthly salary to the madre, and the government would provide funds for lunch, two snacks and a nutrition beverage per day for the children (the menu was established by a nutritionist and provided 50–70 % of the advisable daily amount of calories). The program was designed to promote children's physical, social and cognitive development, and madres attended a 40-hour training on child development. Length of exposure: Varies Scale: National, although it is unclear how many localities qualified	instruments for number of months attended and exposure (number of months attended/age in months)			
Gambia, The *Maternal supplementation*	Hawkesworth and others 2008 [20] Hawkesworth and others 2009 [21] Hawkesworth and others 2011 [22] Alderman and others 2014 [1]	1989–1994. Women of childbearing age (15–45y) from 28 villages in the West Kiang region were randomized into intervention (biscuits provided from around 20 weeks gestation to term) and control (biscuits provided for 20 weeks after delivery) groups. Women were given 2 biscuits/day that provided extra energy, protein, fat, calcium and iron. Biscuits were prepared by village women and distributed by two birth attendants in each village who intensively encouraged consumption and in whose presence the biscuits had to be consume. Attendants recorded compliance. Women in both groups received routine ANC. The subjects of the impact evaluations were the children of the women who participated. Length of exposure: 20 weeks Scale: 1460 women in both treatment and control	Dose response between receiving biscuits during pregnancy and receiving them for 20 weeks after delivery	At birth	11–17 years old 16–22 years old	BMI; Body fat; Diastolic blood pressure; FFMI; FMI; Glucose**; HDL cholesterol; LDL cholesterol; LMI; Log insulin; Mean arterial pressure; Pulse pressure*; Total cholesterol; Systolic blood pressure; Triglycerides; Trunk fat; Weight Digit span backward; Digit span forward; Highest grade completed; Raven's matrices; Vocabulary

Guatemala *Guatemala's Instituto de Nutrición de Centroamérica y Panamá (INCAP)*	Maluccio and others 2009 [31]	1969–1977. Two sets of similar village pairs were selected in rural eastern Guatemala, and one village from each pair was randomly assigned to receive as a dietary supplement a high protein-energy drink atole (53%, 91 kcal and 6.4 g protein/100 mL). The other two villages received a low-energy drink called fresco (47%, 33 kcal/100 mL, no protein). The drinks were distributed at centrally located feeding centers twice a day to any member of the village. All villages' residents were offered free medical care throughout the intervention. All villages simultaneously received preventative services including immunization and antiparasite efforts. Length of exposure: 3–7 years Scale: One large (~900 residents) and one small (~500 resident) village	Children who lived in villages that were assigned to receive fresco	<36 months	25–42 years old	Highest grade completed (men, women**); Raven's matrices**; Reading**
Honduras *Programa de Asignación Familiar-II (PRAF-II)*	Rackstraw 2014 [39]	2000–2002. Randomized CCT implemented by the Honduran government targeted munici-palities with the lowest mean height-for-age scores. Pregnant women and families with children aged 0–3 received up to two health transfers of 644 Lempiras (US$40) each per year, conditional on regular health center visits. Children aged 6–12 who had not yet completed 4th grade were eligible for up to three education transfers of 828 Lempiras (US$58) each per year. The average transfer amount was 5 percent of median per capital expenditure. Length of exposure: 2 years	Children who were born in municipalities that were randomized not to receive treatment.	0–3 years old	13–15 years old	Literacy; School attendance***; Hours worked; Worked 1 hour or more; Years of schooling***

		Scale: 40 municipalities				
India *Women's Political Participation*	Pathak and Macours 2013 [37]	1992–present. The Indian constitution was amended to require that one-third of the seats in local councils, as well as one-third of Pradhan (leadership) positions be reserved for women. This was done by randomly reserving seats in local elections, such that in each election, certain localities would have reserved seats and others would not. Length of exposure: Varies Scale: National	Three groups were compared to each other: 1. Children born in 2001–2002 in municipalities that had seats reserved from 1995–2001 2. Children born in 2001–2002 in municipalities that had seats reserved from 2001–2006 3. Children born in 2001–2002 in municipalities that had seats reserved from 2006–2011	Varies	8 years old	Reading (in utero)*; Reading (0–5 years old); Numeracy; Vocabulary
India *Total Sanitation Campaign (TSC)*	Spears and Lamba 2013 [44]	2001–present. The Indian government promoted the construction and use of low-cost pit latrines in rural areas. This was done through various strategies, including motivating high-caste villagers to build latrines for their village and providing monetary incentives for villages that could verify they were open-defecation free. Length of exposure: 6 years Scale: Rural communities throughout India	Children who lived in villages that did not have a latrine during their first year of life	<1 year old	6 years	Mathematics (recognized numbers)*; Reading (recognized letters)*

Indonesia	Cas 2012 [12]	1990–1996. More than 54,000 nursing school graduates were gradually deployed in most of Indonesia's nonmetropolitan villages as midwives with the objective of substantially increasing women's access to health care and safe delivery services. The village midwives provide antenatal, postnatal and general health care, working with traditional birth attendants and referring complicated obstetric cases to health centers and hospitals. Length of exposure: Varies (up to 3 years) Scale: 4575	Children born in 1986 to 1989 who did not receive the intervention prior to age 3. (Treatment group are those born in 1993 to 1996 who received the treatment from 0 to age 3)	At birth	11–14 years old	Highest grade completed**; Mathematics test***; Probability of enrolling primary school at age 6; Raven's matrices***; Schooling (in school)
Jamaica *Psychosocial stimulation in LBW-T children*	Walker and others 2010 [54]	1999–2001. Low birth weight infants from lower socioeconomic households (based on mother's education) were recruited from the main public maternity hospital in Kingston and randomly assigned to treatment or control. Treatment was divided into two phases, with community health workers visiting the mothers and children weekly. In the first phase (the first 8 weeks of life), the focus was on improving mothers' responsiveness to their infants. In the second phase (from 7–24 months), the health worker demonstrated play techniques, taught proper discipline habits, and provided toys. Length of exposure: 2 years Scale: 140 infants (70 T and 70 C)	Low birth weight infants who were randomly assigned to the control group	At birth	6 years old	Attention (map search); Attention (opposite-same switching); Corsi blocks***; Digit span; Early reading; PPVT; Reading; Strength and difficulties questionnaire**; WPPSI Full-Scale; WPPSI Performance**; WPPSI Verbal
Jamaica *Psychosocial Stimulation and Nutritional*	Walker and others 1996 [50]	1986–1987. 129 stunted children were randomly assigned to one of three treatment groups or to a control group. The three treatment groups were: 1. Supplementation with 1 kg milk based formula each week, 2. Weekly play sessions with mother	Supplementation: groups 1 and 3 v. group 2 and control Stimulation: groups 2 and 3 v. group 1	9–24 months	7–8 years old	Height-for-age z-score; Weight-for-age z-score

Supplementation	Walker and others 2000 [51]	and child, including weekly home visits to help the mothers improve their verbal interactions with the child, and 3. Both supplementation and weekly play sessions. Length of exposure: 2 years Scale: 96	and control Physical outcomes: Each group was compared to the others Cognitive outcomes: Stimulation: groups 2 and 3 v. group 1 and control	11–12 years old	BMI; Corsi blocks; Digit span backward; Digit span forward; Height-for-age z-score; PPVT; Raven's matrices**; Search; Stroop; Verbal analogies; Vocabulary**; WISC-R Full-scale**; WISC-R Performance*; WISC-R Verbal**
	Walker and others 2005 [52]		Stimulation: groups 2 and 3 v. group 1 and control	17–18 years old	Digit span backward^; Digit span forward^; Full-scale IQ^; Mathematics^; PPVT^; Performance IQ^; Raven's matrices^; Reading^; Verbal analogies^; Verbal IQ^; Visual spatial working memory^
	Walker and others 2006 [53]		Supplementation: groups 1 and 3 v. group 2 and control Stimulation: groups 2 and 3 v. group 1 and control		Anti-social behavior; Anxiety***; Attention deficit**; Cognitive problems or lack of attention; Depression**; Hyperactivity; Oppositional behavior*; Self-esteem**
	Gertler and others 2013 [18] Gertler and others 2014 [19]		Stimulation: groups 2 and 3 v. group 1 and control	22 years old	Passed ≥ 1 CAP exam**; Passed ≥ 1 CXC exam*; Passed ≥ 4 CXC exams; HOME scores***; Looking for work; Migration (baseline*, follow-up**); Any college*; Any vocational schooling; Schooling (in school)***; In school (full time)***; Cognitive skills***; Skills (ever expelled)**; Skills (externalizing behavior); Skills (internalizing behavior)**; Years of schooling*; Employed*; Employed full-time; Employed in non-temporary job; Log earnings

	Walker and others 2011 [55]		Supplementation: groups 1 and 3 v. group 2 and control Stimulation: groups 2 and 3 v. group 1 and control			Anxiety^; Depression^; Full-scale IQ^; General knowledge^; Highest grade completed^; Involved in fight^; Involved in violent crime^; Mathematics^; Pass secondary level exams^; Performance IQ^; Reading^; Social inhibition^; Verbal IQ^; Weapon use^
Kenya *Primary School Deworming Project*	Ozier 2014 [36]	1998–2001. Schoolchildren in southern Busia, a poor and densely populated farming region in western Kenya, in randomly assigned to receive free deworming treatment. One-third of the schools were assigned to receive it in both 1998 and 1999, one-third in just 1999, and one-third in 2001. Length of exposure: 1–2 years Scale: 73 primary schools	Younger siblings of children who were dewormed are compared to younger siblings of those who were not	<1 year old	8–15 years old	Digit span backward; Digit span forward; General intelligence**; Height; Height-for-age z-score; PPVT*; Raven's matrices***; Stunting; Verbal fluency (animals**, foods*)
Mauritius *Experimental nursery schooling*	Raine and others 2001 [40]	1973–1976. The enrichment program took place in two specially constructed nursery schools. In addition to regular educational activities, the children were instructed in hygiene and nutrition and received regular medical inspections from doctors who visited the nurseries. Teachers were specially trained in basic kindergarten education as well as physical health, art, drama, and music. Parents were encouraged to participate in parent-teacher associations. Length of Exposure: 2 years Scale: 100 participants	Children who attended the traditional Mauritian petites écoles	3 years old	11 years old	EEG during continuous performance task; Resting EEG; Skin conductance arousal***; Skin conductance orienting (amplitude***, latency, recovery times***, rise time***)
	Raine and others 2003 [41]				17 years old 23 years old	Anxiety; Attention deficit; Cognitive disorganization**; Cognitive perceptual; Conduct disorder***; Court-reported criminal offenders*; Disorganized; Interpersonal deficits; Motor excess**; Positive Schizotypal personality**; Psychotic behavior**; Schizotypal personality; Self-reported criminal offenders**; Socialized aggression

Mexico *Progresa* Note: The program was renamed to Oportunidades in 2002. This report refers to it as Progresa since the IEs included are evaluating the effect of the program during the Progresa time period.	Todd and Winters 2011 [47]	Early treatment began in April, 1998; late treatment began in November, 1999. Within communities assigned to receive treatment, eligible household (those that qualified as poor based on a marginality index) receive bi-monthly cash transfer, conditional on family members completing regular health checkups and children attending school. Children become eligible for the educational scholarships when they enter third grade. The transfer is, on average, equivalent to an average of 20% of household consumption. Recipients—usually mothers—are required to attend regular meetings (pláticas) in which health and nutrition practices are discussed. In addition to the cash transfer, there is a nutritional component, which includes the provision of nutritional supplements to pregnant and lactating women and to children 4–24 months old or up to 59 months old if signs of malnutrition are detected by the clinic personnel. A prerequisite for receiving nutritional complements is ongoing growth monitoring of preschool children. Length of exposure: Varies. (NB: Children are eligible for some portion of Progresa through the end of high school, but for the purposes of this review, the estimates are limited to those who enrolled during the early childhood period) Scale: 320 villages in the early treatment, 186 villages in the late treatment	Cohorts of children who turned 3 years old around the time the late treatment started (isolates the impact of receiving an additional 18 months exposure before 3 years old)	0–6 years old	6–9 years old	Number of days of school missed in a usual month***; Probability of enrolling primary school at age 6*
	Manley, Fernald, and Gertler 2012 [32] Fernald, Gertler, and Neufeld 2009 [16]				8–10 years old	BMI z-score; Cognitive assessment score*; Height-for-age z-score; Strength and difficulties questionnaire (cash, conditionalities**); Verbal assessment score***
	Behrman and others 2008 [8]				7–11 years old	BMI z-score; Cognitive development; Highest grade completed; Language development**; Overweight; Progressing through school on time; Reading; Strength and difficulties questionnaire**; Stunting
	Behrman, Parker, and Todd 2009 [7]		Experimental sample: the effect of an additional 18 months of program participation Matched sample: children who were in the early treatment group were matched with children who had never enrolled in		7–11 years old†	Age at first grade entry^; Highest grade completed^; Processing through school on time^

			Progresa			
Mozambique *Early Childhood Development Program*	Martínez, Naudeu, and Pereira 2012 [35]	2008–2010. After a successfully pilot program that started in 2005, Save the Children scaled up a community-based preschool model in Gaza province wherein randomly selected communities committed space for classroom construction, as well as locally available construction material and 100 percent of the labor, in exchange for technical assistance and materials for up to three classrooms, playgrounds, child-sized latrines and washing stations. Community members were also to form a committee to encourage parent participation and enrollment and to select two volunteer teachers for each class. Save the Children conducted training for the teachers. Preschool typically lasted 3 hours and 15 minutes with activities designed to stimulate child development. Length of exposure: 1–3 years Scale: 30 communities	Children in participated in the preschool program were compared to those in communities that were not assigned to receive one	3–5 years old	5–9 years old	Appropriate grade for age**; Able to count to 20*; Able to recognize geometric shapes**; Able to remember things easily**; Able to say which number is bigger between two***; Able to sort and classify objects by common characteristics**; Able to use one-to-one correspondence***; Experimenting with writing tools**; Interested in games involving numbers*; Interested in mathematics**; Cognitive development and language**; Communication*; Communication and general knowledge; Cough; Emotional maturity; Ever gone to school**; Sickness**; Gross motor coordination; Had diarrhea; Had skin problems; Height-for-age z-score; Physical health and well-being; Precise motor coordination*; Primary school drop-out; Primary school enrollment***; Problem solving**; Changana (raw, standardized); Portuguese (raw, standardized); Social competence; Stunting; Time use^; ASQ score**; Wasting; Weight-for-age z-score
Nepal *Maternal Micronutrient Supplementation*	Devakumar and others 2014 [15]	2002–2004. Women attending Janakpur Zonal Hospital for antenatal care were randomly allocated to receive either a multivitamin supplement (containing vitamins A, B1, B2, B6, B12, C, D, and E and niacin, folic acid, iron, zinc, copper, selenium, and iodine) or a control supplement of iron and folic acid. Supplements	Children of women who were randomized to receive the control supplement of iron and folic acid	In utero	8 years old	Biceps thickness; BMI z-score; Chest circumference; Diastolic blood pressure; Fat mass; Head circumference; Height; Height-for-age z-score; Hip circumference; Lean mass; Mid-upper-arm circumference; Renal dimension; Subscapular thickness; Suprailiac thickness; Systolic blood pressure; Triceps

		were taken every day from 12 to 20 weeks' gestation (average 15.9 weeks) until delivery, and women were assessed every two weeks. Length of exposure: 20–28 weeks Scale: 1200 women				thickness; Trunk length; Upper leg circumference; Waist circumference; Weight; Weight-for-age z-score
Nepal *Nepal Nutrition Intervention Project–Sarlahi*	Stewart and others 2009a [45] Stewart and others 2009b [46]	1999–2001. Pregnant women in the rural southeastern Sarlahi district of Nepal were randomly assigned to a daily supplementation or control group from early pregnancy to 3 months postpartum. There were four supplementation groups: 1. Folic acid, 2. Folic acid and iron, 3. Folic acid, iron and zinc, and 4. Folic acid, iron, zinc and 11 additional vitamins and minerals. All women—both in the treatment and control groups—were given vitamin A. Length of intervention: Up to 1 year Scale: 4047 women‡	Children of women who were in groups 1–4 were compared to children of women who just received vitamin A	In utero	6–8 years old	Blood pressure (SBP or DBP > 90th percentile)^; BMI^; Diastolic blood pressure^; Fasting insulin^; Height^; Glucose^; Glucose ≥ 85th percentile^; Glycated hemoglobin^; HDL cholesterol^; HDL cholesterol < 0.9 mmol/L^; Homeostasis model assessment^; Risk for metabolic syndrome^; Risk of microalbuminuria^; Systolic blood pressure^; Total cholesterol^; Triglycerides^; Triglycerides ≥ 1.7 mmol/L^; Urinary microalbumin: creatinine^; Waist circumference^; Waist circumference ≥ 85th percentile^
Nicaragua *Red de Protección Social (RPS)*	Barham and others 2014 [6]	Early treatment group began in 2000; late treatment group began in 2003. Forty-two poor localities were randomized into equal treatment and control (late treatment) groups wherein poor households received bimonthly transfers equal to approximately 18 percent of households' pre-program expenditures. The transfers were conditional on meeting health and educational (starting at 7 years old) requirements, including regular preventative healthcare visits for children under five. The designated caregiver in each house was also required to attend bimonthly meetings on nutrition and health. Health services were free and delivered by private health providers contracted by RPS. While households in the early treatment group were no longer eligible for the cash transfers after 2003, they continued	Dose response of receiving RPS during the first 1,000 days versus receiving it from 2–5 years old	At birth	10 years old	Anthropometric index; Cognition index (boys***, girls); Digit span backward; Digit span forward; Height-for-age z-score; Long-term memory; Process speed**; Raven's matrices; Short-term memory*; Vocabulary**; Weight-for-age z-score

		to receive free private health services through 2005. Length of exposure: 3 years Scale: Rural communities in six municipalities in central and northern Nicaragua				
Romania *Bucharest Early Intervention Project (BEIP)*	Windsor and others 2013 [56] Fox and others 2011 [17] Bos and others 2009 [11] Levin and others 2014 [29] Almas and others 2012 [2] Vanderwert and others 2010 [49] Smyke and others 2012 [43] Troller-Renfree and others 2014 [48]	2001–2005. Abandoned or orphaned children were placed in the institutions at an average age of 3 months (although that does not take into account the time the children may have spent in maternity hospitals after being abandoned). To be eligible for the study, children had to pass a pediatric and neurological assessment and were excluded if they had known genetic syndromes or signs of fetal alcohol syndrome. Eligible children were then randomized to either care as usual (i.e., staying in the institution) or else foster care. Children were randomized with their siblings. The average age for entering foster care was 23.6 months. Foster parents were recruited and trained in collaboration with US practitioners, and social workers were available to help foster parents manage problems and create good relationships with the children. Children stayed in their assigned places (institutions or foster care) until at least 54 months old, at which point children in institutions could enter foster care. Children in foster care were guaranteed not to be returned to institutions. Children in either group could be adopted or return to their biological parents at any point during the study, Length of intervention: Varies Scale: 136 children	Children randomized to remain in institutional care	6–31 months old	8 years old	Balance; Bilateral coordination; Executive function; Fine motor integration; Fine motor precision; Manual dexterity; Mean length of utterance**; Motor skills; Nonword repetition; Nonword repetition (FG placed by 25 months)**; Running speed and agility; Sentence repetition**; Sentence repetition (FG placed by 25 months); Word identification**; Word identification (FG placed by 25 months)**; Strength; Upper-limb coordination; Visual memory; WISC-IV index (full-scale*, perceptual reasoning, processing speed, verbal comprehension**, working memory); Social skills; EEG power (alpha band, beta band, theta band); Disinhibited type of reactive attachment disorder (disturbances of attachment interview***, foster care placement before/after 24 months); Inhibited type of reactive attachment disorder (disturbances of attachment interview***, foster care placement before/after 24 months); Positive bias**; Threat bias; JP (competitive) task orientation; JP (cooperative) social engagement; LP (cooperative) social engagement; LP (cooperative) task orientation; Speech reticence**

	Almas and others 2015 [3]					
	Levin and others 2015 [30]				10 years old	Communication; Reciprocal social interaction*; Restricted, repetitive, and stereotyped patterns of behavior; SCQ score***
	Humphreys and others (forthcoming) [23]				12 years old	ADHD; Externalizing disorder**; Internalizing disorder
South Africa *Child Support Grant (CSG)*	DSD, SASSA, and UNICEF South Africa 2012 [14]	1998–present. Caregivers of vulnerable children received a monthly unconditional cash transfer, which can start soon after birth and last until the child is 18 years old (originally, children were only eligible until they were seven years old, but the threshold has increased over the years). Eligibility is based on a means test. Length of exposure: Varies Scale: National	Physical: Children who started CSG before they were 2 years old compared to children who started after they were two Other outcomes: Children who started CSG at birth compared to those who started at six years old	<2 years old	10 years old	Ability to read a story; DPT (number of immunizations, proportion of children receiving all immunizations); Height-for-age z-score; Hepatitis (number of immunizations, proportion of children receiving all immunizations); Highest grade completed; Hours spent doing housework; Hours spent studying; Mathematics (arithmetic*, overall, shape recognition); Number of days ill in last 15 days; Polio (number of immunizations, proportion of children receiving all immunizations); Probability of delayed enrollment; Probability of grade repetition Probability of illness in last 15 days; Reading; Stunting; Whether the child does housework; Whether the child studied in the last week

Thailand *Iron and Zinc Supplementation*	Pongcharoen 2010 [38]	1998. Infants in Khon Kaen province (northeast) were randomly divided into one of four groups for a double blind, placebo-controlled trial: 1. 10 mg of iron, 2. 10 mg of zinc, 3. Both iron and zinc, 4. placebo. Infants received supplements daily for 6 months. All children received one dose of 1,500 µg retinol equivalent (RE) vitamin A at the beginning of the study. Length of exposure: 6 months Scale: 609 infants, including the placebo group	Each group was compared independently to each of the other groups	4–6 months old	9 years old	Age at first grade entry^; BMI^; BMI z-score^; English^; Full-scale IQ^; Height^; Mathematics test^; Percentage in current grade 2, 3, 4^; Performance IQ^; Raven's matrices^; Science^; Thai language^; Verbal IQ^; Weight^; WISC-III index^ (freedom from distractibility, perceptual organization, processing speed, verbal comprehension)
Uruguay *Preschool*	Berlinski, Galiani and Manacorda 2008 [10]	Started in 1995. The evaluated intervention is whether a child attended either a public or private preschool. The increase in preschool attendance was triggered by a government program to build new classrooms and hire more preschool teachers. Length of exposure: 1–3 years Scale: National	Children who did not attend either public or private preschool	3–5 years old	7–15 years old	Primary school drop-out^; Probability of completing primary school^; School attendance^; Years of schooling^

Scale: indicates the size of the treatment group.
* *p*<0.1; ** *p*<0.05; *** *p*<0.01.
^ These studies report several results by treatment arm or subgroup for this outcome. See appendix B for detailed results.
† Some of the matched sample outcomes are estimated for children up to 14 years old. However, since this review focuses on the effect of receiving Progresa during the early childhood period, only estimates for children up to 11 years old (5 years old at baseline) are used.
‡ When the control group, which received vitamin A supplements, is included, there were 4926 participants.

Appendix B. Complete List of Outcomes Appearing in Included Studies

This appendix lists all outcomes estimated in the 55 studies that passed the inclusion criteria and quality check in the systematic review. Outcomes are listed in alphabetical order as the authors described them. Outcomes in bold were analyzed in the body of the systematic review. Study identification numbers correspond to the numbered studies in References. In the main body of the report, they are denoted in [brackets] but are not bracketed here for better readability. Information in parentheses indicates the population or comparators for whom a particular outcome is estimated. A study without a parenthetical population note implies that its outcome was measured for the entire study population.

The letter "v" indicates versus and gives the explicit counterfactual where multiple tests are reported; T= treatment group; C = control group. Statistical significance is noted for each study in which an outcome appears by the standard convention:

* $p<0.1$; ** $p<0.05$; *** $p<0.01$.

List of Unique Outcomes	Source (Study ID)
Ability to read a story	14
Adiponectin	34
Age at first grade entry	7 (7–8 boys, T1998 v C2003); 7 (7–8 boys, T1998 v T1999); 7 (7–8 girls, T1998 v C2003)*; 7 (7–8 girls, T1998 v T1999)*; 7 (7–11 boys, T1998 v C2003); 7 (7–11 boys, T1998 v T1999); 7 (7–11 girls, T1998 v C2003); 7 (7–11 girls, T1998 v T1999); 7 (9–11 boys, T1998 v C2003); 7 (9–11 boys, T1998 v T1999); 7 (9–11 girls, T1998 v C2003); 7 (9–11 girls, T1998 v T1999); 38 (iron); 38 (iron and zinc); 38 (zinc)
Allergies (at least one antigen)	25
Allergies to Alternaria	25
Allergies to birch pollen	25
Allergies to cat	25
Allergies to house dust mite	25

List of Unique Outcomes	Source (Study ID)
Allergies to northern grasses	25
Anthropometric index	6 (boys); 6 (girls)
Anti-social behavior	53
Anxiety	41; 53***; 55 (stimulation); 55 (supplementation)*
Apolipoprotein A1	34
Appropriate grade for age	35**
Attention-concentration	5 (8–14 yr)**; 5 (15–19 yr); 5 (20–24 yr)
Attention (map search)	54 (map search); 54 (opposite-same switching)
Attention deficit	41; 53**
Attention deficit / hyperactivity disorder (ADHD)	23
Balance	29 (standing on one leg on a balance beam, eyes open); 29 (walking forward on a line)
Biceps thickness	15
Bilateral coordination	29 (jumping in place); 29 (tapping feet and fingers)
Block designs	27
Blood pressure (SBP or DBP > 90th percentile)	46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
BMI	20 (female); 20 (male); 22; 24; 33; 38 (iron); 38 (iron and zinc); 38 (zinc); 46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient); 51 (supplementation)
BMI ≥ 85th percentile	33**
BMI ≥ 95th percentile	33*
BMI z-score	8; 15; 32; 38 (iron); 38 (iron and zinc); 38 (zinc)
Body fat	20 (female); 20 (male); 33
Chest circumference	15

List of Unique Outcomes	Source (Study ID)
Child is able to count to 20	35*
Child is able to recognize geometric shapes	35**
Child is able to remember things easily	35**
Child is able to say which number is bigger between two	35***
Child is able to sort and classify objects by a common characteristics (e.g., shape, color, size)	35**
Child is able to use one-to-one correspondence	35***
Child is experimenting with writing tools	35**
Child is interested in games involving numbers	35*
Child is interested in mathematics	35**
Cognition index	6 (boys)***; 6 (girls)
Cognitive assessment score	32*
Cognitive development	8
Cognitive development and language	35**
Cognitive disorganization	41**
Cognitive perceptual	41
Cognitive problems or lack of attention	53
Communication	30; 35*
Communication and general knowledge	35
Conduct disorder	41***
Conduct problems	28 (assessed by parent); 28 (assessed by teacher)
Corsi blocks	51 (stimulation + both interventions); 54***

List of Unique Outcomes	Source (Study ID)
Cough in the last 4 weeks	35
Court-reported criminal offenders	41*
Depression	53**; 55 (stimulation)**; 55 (supplementation)
Diastolic blood pressure	15; 21; 22; 24; 34; 46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Dieting	42**
Digit span	54
Digit span backward	1; 6; 36; 51 (stimulation + both interventions); 52 (stimulation); 52 (supplementation)
Digit span forward	1; 6; 36; 51 (stimulation + both interventions); 52 (stimulation); 52 (supplementation)
Disinhibited type of reactive attachment disorder	43 (disturbances of attachment interview)***; 43 (foster care placement before/after 24 months)
Disorganized	41
DPT: number of immunizations	14
DPT: proportion of children receiving all immunizations	14
Early reading	54
Eating attitudes	42**
EEG (electroencephalogram) during continuous performance task	40
EEG power in the alpha band	49
EEG power in the beta band	49
EEG power in the theta band	49
Emotional maturity	35
Emotional symptoms	28 (assessed by parent); 28 (assessed by teacher)
Employed	19*

List of Unique Outcomes	Source (Study ID)
Employed full-time	19
Employed in nontemporary job	19
English	38 (iron); 38 (iron and zinc); 38 (zinc)
Ever gone to school	35**
Ever had asthma	25
Ever had eczema	25
Ever had hay fever symptoms	25
Ever had wheezing	25
Ever sick in the last 4 weeks	35**
Exams	18 (passed at least one Caribbean Advanced Proficiency Examination)**; 18 (passed at least one CXC exam)*; 18 (passed 4 or more CXC exams)
Executive function	11 (spatial working memory); 11 (stockings of Cambridge)
Externalizing behavior scores	28 (assessed by parent); 28 (assessed by teacher)
Externalizing disorder	23**
Fasting insulin	34; 46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Fat mass	15
FFMI	20 (female); 20 (male); 33
Fine motor integration	29 (copying a square); 29 (copying a start)
Fine motor precision	29 (drawing lines through crooked paths); 29 (folding paper)
FMI	20 (female); 20 (male); 22; 33
Food preoccupation	42
Full-scale IQ	27*; 38 (iron); 38 (iron and zinc); 38 (zinc); 52 (stimulation)**; 52 (supplementation); 55 (stimulation)***; 55 (supplementation)

List of Unique Outcomes	Source (Study ID)
General intelligence - first component	36**
General knowledge	55 (stimulation)***; 55 (supplementation)
Glucose	22**; 34; 46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Glucose ≥ 85th percentile	46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Glycated hemoglobin	46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Gross motor coordination	35
Had diarrhea in the last 4 weeks	35
Had skin problems in the last 4 weeks	35
Hay fever symptoms in past 12 months	25
HDL cholesterol	22; 46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
HDL cholesterol < 0.9 mmol/L	46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Head circumference	15; 24*; 33
Height	15; 24; 33; 36; 38 (iron); 38 (iron and zinc); 38 (zinc); 45 (folic acid+iron+zinc)*; 45 (folic acid, folic acid+iron, multiple micronutrient)
Height-for-age z-score	6; 14; 15; 32; 35; 36; 38 (iron); 38 (iron and zinc)*; 38 (zinc); 50 (stimulation); 50 (stimulation + supplementation); 50 (supplementation); 51 (stimulation); 51 (stimulation + supplementation); 51 (supplementation)
Hepatitis: number of immunizations	14
Hepatitis: proportion of children receiving all immunizations	14
Highest grade completed	1; 5 (8–14 yr)***; 5 (15–19 yr); 5 (20–24 yr); 7 (6–8 boys, T1998 v C2003); 7 (6–8 boys, T1998 v T1999); 7 (6–8 girls, T1998 v T1999); 7 (9–11 boys, T1998 v C2003)***; 7 (9–11 boys, T1998 v T1999); 7 (9–11 girls, T1998 v C2003)***; 7 (9–11 girls, T1998 v T1999); 8; 12**; 14; 31 (men); 31 (women)**; 55 (stimulation)**; 55 (supplementation)

List of Unique Outcomes	Source (Study ID)
Hip circumference	15; 24; 33
HOME scores at end of trial	18***
Homeostasis model assessment	46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Hours spent doing housework per week	14
Hours spent studying per day	14
How many of your students pay attention in class?	9**
How many of your students put a large amount of effort into understanding explanations?	9**
How many of your students regularly participate in your class?	9**
How many students are well disciplined in the classroom?	9
Hyperactivity	53
Hyperactivity (SDQ)	28 (assessed by parent); 28 (assessed by teacher)
IGF-I	33
Incisor DMFT	26
Inhibited type of reactive attachment disorder	43 (disturbances of attachment interview)***; 43 (foster care placement before/after 24 months)
Internalizing behavior scores	28 (assessed by parent); 28 (assessed by teacher)
Internalizing disorder	23
Interpersonal deficits	41
Involved in a physical fight	55 (stimulation)**; 55 (supplementation)
Involved in violent crime	55 (stimulation)**; 55 (supplementation)
JP (competitive) task orientation	3
JP (cooperative) social engagement	3

List of Unique Outcomes	Source (Study ID)
Language development	8**
Language test score	5 (8–14 yr); 5 (15–19 yr); 5 (20–24 yr); 9**
LDL cholesterol	22
Lean mass	15
Leg length	24; 33
Literacy	39
LMI	22
Log earnings (various versions)	19
Log insulin	22
Long-term memory	6
Looking for work	18
LP (cooperative) social engagement	3
LP (cooperative) task orientation	3
Manual dexterity	29
Mathematics (numeracy)	37
Mathematics (Annual Status of Education Report)	44 (recognizes numbers 1 to 9 at 6 years old)*; 44 (recognizes numbers 1 to 9 at 7 years old)*; 44 (recognizes numbers 1 to 9 at 8 years old)*; 44 (recognizes numbers 10 to 99 at 6 years old)*; 44 (recognizes numbers 10 to 99 at 7 years old)*; 44 (recognizes numbers 10 to 99 at 8 years old)*
Mathematics (Early Grade Mathematics Assessment)	14 (arithmetic)*; 14 (overall); 14 (shape recognition)
Mathematics (from Indonesian Family Life Survey 3 & 4)	12***
Mathematics (school-specific test)	9**; 13***; 27; 38 (iron); 38 (iron and zinc); 38 (zinc)
Mathematics (Wide Range Achievement Test)	52 (stimulation); 52 (supplementation); 55 (stimulation)**; 55 (supplementation)

List of Unique Outcomes	Source (Study ID)
Mean arterial pressure	21
Mean length of utterance	56**; 56 (FG placed by 25 months)**
Mid-upper-arm circumference	15; 24; 33; 38 (iron); 38 (iron and zinc); 38 (zinc)
Midthigh circumference	24**
Migration	18 (full baseline sample)*; 18 (sample found at follow-up)**
Motor excess	41**
Motor skills	11; 29
No. secondary level examination passes	55 (stimulation)*; 55 (supplementation)
Nonword repetition	56; 56 (FG placed by 25 months)**
Number of days ill in the last 15 days	14
Number of days of school missed in a usual month	47***
Oppositional behavior	53*
Orientation	5 (8–14 yr)**; 5 (15–19 yr); 5 (20–24 yr)
Other subjects	27
Overweight	8
Peabody Picture Vocabulary Test (receptive vocabulary)	36*; 51 (stimulation + both interventions); 52 (stimulation)**; 52 (supplementation); 54
Peer problems	28 (assessed by parent); 28 (assessed by teacher)
Percentage in each current grade- 2,3,4	38 (iron); 38 (iron and zinc); 38 (zinc)
Performance IQ	27; 38 (iron); 38 (iron and zinc); 38 (zinc); 52 (stimulation)*; 52 (supplementation); 55 (stimulation)***; 55 (supplementation)
Physical health and well-being	35
Polio: number of immunizations	14

List of Unique Outcomes	Source (Study ID)
Polio: proportion of children receiving all immunizations	14
Positive Schizotypal personality	41**
Positive bias	48**
Precise motor coordination	35*
Presence of metabolic syndrome	34
Pressure from others	42**
Primary school drop-out	10 (7-yr)***; 10 (8-yr)**; 10 (9-yr); 10 (10-yr); 10 (11-yr); 10 (12-yr); 10 (13-yr)***; 10 (14-yr)**; 10 (15-yr)***; 35
Primary school enrollment	35***
Probability of attending primary school	4
Probability of attending secondary school	4**
Probability of completing primary school	10 (7-yr); 10 (8-yr); 10 (9-yr); 10 (10-yr); 10 (11-yr)***; 10 (12-yr); 10 (13-yr)***; 10 (14-yr)***; 10 (15-yr)***
Probability of delayed enrollment	14
Probability of enrolling primary school at age 6	12; 47*
Probability of grade repetition	14
Probability of illness in the last 15 days	14
Probability of passing a grade between 2002 and 2003	4 (8–12 yrs); 4 (13–17 yrs)*
Problem solving	35**
Process speed	6**
Progressing through school on time	7 (9–11 boys, T1998 v C2003)**; 7 (9–11 boys, T1998 v T1999); 7 (9–11 girls, T1998 v C2003)***; 7 (9–11 girls, T1998 v T1999); 8
Proportion of children with incisor DMFT ≥ 1	26
Proportion of children with total DMFT ≥ 1	26

List of Unique Outcomes	Source (Study ID)
Proportion with incisor DMFT ≥ 2	26
Proportion with total DMFT ≥ 2	26
Prosocial behavior	28 (assessed by parent); 28 (assessed by teacher)
Psychotic behavior	41**
Pulse pressure	21*
Purging and vomiting	42*
Raven's matrices (nonverbal cognition)	1; 6**; 12***; 27; 31**; 36***; 38 (iron); 38 (iron and zinc); 38 (zinc); 51 (stimulation + both interventions)**; 52 (stimulation)*; 52 (supplementation)
Raw Changana	35
Raw Portuguese	35
Reading	8; 13***; 14; 27*; 31**; 37 (in utero)*; 37 (0–5 years); 44 (recognized letters at 6 years old)*; 44 (recognized letters at 7 years old)*; 44 (recognized letters at 8 years old)*; 52 (context, stimulation)***; 52 (context, supplementation); 52 (sentence, stimulation)***; 52 (sentence, supplementation); 54; 55 (stimulation)***; 55 (supplementation)
Recall	5 (8–14 years old); 5 (15–19 years old); 5 (20–24 years old)
Reciprocal social interaction	30*
Recurrent itchy rash	25
Registration	5 (8–14 years old)***; 5 (15–19 years old); 5 (20–24 years old)
Renal dimension	15 (left anteroposterior distance); 15 (left length); 15 (right anteroposterior distance); 15 (right length)
Resting EEG	40
Restricted, repetitive, and stereotyped patterns of behavior	30
Restriction and control	42
Risk for metabolic syndrome	46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)

List of Unique Outcomes	Source (Study ID)
Risk of microalbuminuria	46 (folic acid)*; 46 (folic acid + iron); 46 (folic acid + iron + zinc)*; 46 (multiple micronutrient)
Running speed and agility	29
Schizotypal personality total score	41
School attendance	10 (7-yr)**; 10 (8-yr)**; 10 (9-yr)*; 10 (10-yr); 10 (11-yr); 10 (12-yr)**; 10 (13-yr)***; 10 (14-yr)***; 10 (15-yr)***; 39***
Schooling (any college)	18*
Schooling (any vocational training)	18
Schooling (in school full time)	18***
Schooling (in school)	12; 18***
Science	38 (iron); 38 (iron and zinc); 38 (zinc)
Search	51 (stimulation + both interventions)
Sentence repetition	56**; 56 (FG placed by 25 months)
Self-esteem	53**
Self-reported criminal offenders	41**
Short-term memory	6*
Similarities	27**
Skills (cognitive factor)	18***
Skills (ever expelled from school)	18**
Skills (externalizing behavior factor)	18
Skills (internalizing behavior factor)	18**
Skin conductance arousal - level	40***
Skin conductance orienting - amplitude	40***

List of Unique Outcomes	Source (Study ID)
Skin conductance orienting - latency	40
Skin conductance orienting - recovery times	40***
Skin conductance orienting - rise time	40***
Skinfold thickness	24 (subscapular); 24 (triceps); 33 (subscapular); 33 (triceps)
Social competence	35
Social inhibition	55 (stimulation)**; 55 (supplementation)
Social science	13***
Social skills	2
Socialized aggression	41
Speech reticence	3**
Standardized Changana	35
Standardized Portuguese	35
Strength	29 (push-ups); 29 (sit-ups)
Strength and Difficulties Questionnaire (total difficulties)	8**; 28 (assessed by parent); 28 (assessed by teacher); 32 (cash); 32 (conditionalities)**; 54**
Stroop (Executive Function)	51 (stimulation + both interventions)
Stunting	8; 14; 35; 36
Subscapular thickness	15
Suprailiac thickness	15
Systolic blood pressure	15; 21; 22; 24; 34; 46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Thai language	38 (iron); 38 (iron and zinc); 38 (zinc)
Threat bias	48

List of Unique Outcomes	Source (Study ID)
Time use (caring for children, elders and sick)	35
Time use (community meetings)	35***
Time use (household chores)	35
Time use (play)	35
Time use (school and homework)	35***
Time use (sleep)	35
Time use (work at family's plot)	35**
Total ASQ score	35**
Total cholesterol	22; 46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Total DMFT	26
Total number of teeth	26
Total number of teeth with cavities	26
Total number of teeth with fillings	26
Total SCQ score	30***
Triceps thickness	15
Triglycerides	22; 46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Triglycerides ≥ 1.7 mmol/L	46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Trunk fat	20 (female); 20 (male)
Trunk length	15
Upper leg circumference	15
Upper-limb coordination	29 (dribbling a ball, alternating hands); 29 (dropping and catching a ball, both hands)

List of Unique Outcomes	Source (Study ID)
Urinary microalbumin: creatinine	46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Verbal analogies	51 (stimulation + both interventions); 52 (stimulation)**; 52 (supplementation)
Verbal assessment score	32***
Verbal fluency: animals	36**
Verbal fluency: foods	36*
Verbal IQ	27**; 38 (iron); 38 (iron and zinc); 38 (zinc); 52 (stimulation)**; 52 (supplementation); 55 (stimulation)***; 55 (supplementation)
Visual memory	11 (delayed matching sample); 11 (paired associated learning)
Visual spatial working memory	52 (stimulation); 52 (supplementation)
Vocabulary	1; 6**; 27**; 37; 51 (stimulation + both interventions)**
Waist circumference	15; 24; 33; 46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Waist circumference ≥ 85th percentile	46 (folic acid); 46 (folic acid + iron); 46 (folic acid + iron + zinc); 46 (multiple micronutrient)
Waist:hip ratio	24; 33
Wasting	35
Weapon use	55 (stimulation); 55 (supplementation)
Weight	15; 20 (female); 20 (male); 38 (iron); 38 (iron and zinc); 38 (zinc)
Weight-for-age z-score	6; 15; 35; 38 (iron); 38 (iron and zinc); 38 (zinc); 50 (stimulation); 50 (supplementation + stimulation); 50 (supplementation)
Weight preoccupation	42**
Wheezing in past 12 months	25
Whether the child does housework	14
Whether the child studied in the last week	14
WISC-III index scores freedom from distractibility	38 (iron); 38 (iron and zinc); 38 (zinc)

List of Unique Outcomes	Source (Study ID)
WISC-III index scores perceptual organization	38 (iron); 38 (iron and zinc); 38 (zinc)
WISC-III index scores processing speed	38 (iron); 38 (iron and zinc); 38 (zinc)
WISC-III index scores verbal comprehension	38 (iron); 38 (iron and zinc); 38 (zinc)
WISC-IV index scores full-scale IQ	17*
WISC-IV index scores perceptual reasoning	17
WISC-IV index scores processing speed	17
WISC-IV index scores verbal comprehension	17**
WISC-IV index scores working memory	17
WISC-R Full-scale	51 (stimulation + both interventions)**
WISC-R Performance	51 (stimulation + both interventions)*
WISC-R Verbal	51 (stimulation + both interventions)**
Word identification	56**; 56 (FG placed by 25 months)**
Work	39 (hours worked); 39 (one hour or more)
WPPSI Full-Scale	54
WPPSI Performance	54**
WPPSI Verbal	54
Writing	27
Years of schooling	10 (7-yr)**; 10 (8-yr); 10 (9-yr); 10 (10-yr); 10 (11-yr)**; 10 (12-yr)***; 10 (13-yr)***; 10 (14-yr)***; 10 (15-yr)***; 18*; 39***; 57 (0–2 years old)***; 57 (3–5 years old)

*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.10$.

Appendix C. Search Strategy Details—Identifying Relevant Studies

The data for the systematic review originate from completed (or nearly completed) impact evaluations (IEs) of early childhood development (ECD) interventions that provided estimates on post-ECD outcomes. Potential IEs were identified through a detailed search strategy developed for *Delivering the Millennium Development Goals to Reduce Maternal and Child Mortality: A Systematic Review of Impact Evaluation Evidence* (IEG 2013) and built on existing systematic review frameworks (Card and others 2010; Drabo, Perez-Arce, and Yoong 2011; IEG 2010, 2011). The search strategy included three rounds of data collection. Potential IEs were then submitted to a full text review and, if they qualified as impact evaluations of ECD interventions with post-ECD outcomes, the coding process (see coding strategy for additional information).

Three categories of search terms were identified to capture studies of interest, using a pilot test to modify the terms as necessary. These terms provided the basis for the bibliographic database search that took place during the first round and identified the majority of potential IEs. Round A also included searches of other research, IE, and donor organizations, World Bank databases, top economic journals, and the curriculum vitae of top health economists. Where possible, the title, subject, and abstract for each result were examined. From Round 1, after eliminating duplicates, 1,937 potential IEs were gathered.

Rounds B and C served as comprehensiveness checks. During Round B, the reference lists were reviewed of systematic reviews that focused on early childhood development. Looking first at title, then subject and abstract (if available), 12 potential IEs were identified. Round C involved reviewing the curriculum vitae of the most prolific authors for missing studies and conducting a snowballing activity in which the reference lists for all of the studies selected for coding were reviewed. As a result, 193 potential IEs were found.

Search Terms

Three categories of search terms were used based on the authors' knowledge of ECD literature to allow capture of studies that used appropriate experimental or quasi-experimental methods to estimate the impact on ECD outcomes in low- or middle-income countries.[1,2] The three search term categories are outcomes (A), methods (B), and low- or middle-income country (C). Each search term category has a universe of related search terms likely to be found in the title, subject, or abstract of relevant studies. Note that country names in this third category are search terms only and are not official country names. The search term categories should be joined by AND, while the list of search terms should be joined by OR:

A. Outcomes

(early childhood development) OR (early childhood education) OR (early childhood care) OR (infant development) OR (child cognitive development) OR (child linguistic development) OR (child socioemotional development) OR (child physical development) OR (child growth)

AND

B. Methods

(impact) OR (effectiveness) OR (randomized control trial) OR (randomized trial) OR (control trial) OR (RCT) OR (counterfactual) OR (natural experiment) OR (experimental) OR (quasi experimental) OR (difference in difference) OR (double difference) OR (regression discontinuity) OR (matching) OR (instrumental variable) OR (fixed effects) OR (control area) OR (treatment area) OR (control group) OR (treatment group) OR (panel data)

AND

C. Low- or Middle-Income Country

(Afghanistan) OR (Albania) OR (Algeria) OR (American Samoa) OR (Angola) OR (Antigua and Barbuda) OR (Argentina) OR (Armenia) OR (Azerbaijan) OR (Bangladesh) OR (Belarus) OR (Belize) OR (Benin) OR (Bhutan) OR (Bolivia) OR (Bosnia and Herzegovina) OR (Botswana) OR (Brazil) OR (Bulgaria) OR (Burkina Faso) OR (Burundi) OR (Cambodia) OR (Cameroon) OR (Cape Verde) OR (Central African Republic) OR (Chad) OR (Chile) OR (China) OR (Colombia) OR (Comoros) OR (Congo) OR (Republic of Congo) OR (Costa Rica) OR (Côte d'Ivoire) OR (Ivory Coast) OR (Cuba) OR (Djibouti) OR (Dominica) OR (Dominican Republic) OR (Ecuador) OR (Egypt) OR (El Salvador) OR (Eritrea) OR (Ethiopia) OR (Fiji) OR (Gabon) OR (Gambia) OR (Georgia) OR (Ghana) OR (Grenada) OR (Guatemala) OR (Guinea) OR (Guinea-Bissau) OR (Guyana) OR (Haiti) OR (Honduras) OR (India) OR (Indonesia) OR (Iran) OR (Iraq) OR (Jamaica) OR (Jordan) OR (Kazakhstan) OR (Kenya) OR (Kiribati) OR (Korea) OR (Kosovo) OR (Kyrgyz Republic) OR (Kyrgyzstan) OR (Laos) OR (Latvia) OR (Lebanon) OR (Lesotho) OR (Liberia) OR (Libya) OR (Lithuania) OR (Macedonia) OR (Madagascar) OR (Malawi) OR (Malaysia) OR (Maldives) OR (Mali) OR (Marshall Islands) OR (Mauritania) OR (Mauritius) OR (Mayotte) OR (Mexico) OR (Micronesia) OR (Moldova) OR (Mongolia) OR (Montenegro) OR (Morocco) OR (Mozambique) OR (Myanmar) OR (Namibia) OR (Nepal) OR (Nicaragua) OR (Niger) OR (Nigeria) OR (Pakistan) OR (Palau) OR (Panama) OR (Papua New Guinea) OR (Paraguay) OR (Peru) OR (Philippines) OR (Romania) OR (Russia) OR (Rwanda) OR (Samoa) OR (São Tomé and Principe) OR (Senegal) OR (Serbia) OR (Seychelles) OR (Sierra Leone) OR (Solomon Islands) OR (Somalia) OR (South Africa) OR (South Sudan) OR (Sri Lanka) OR (St. Kitts and Nevis) OR (St. Lucia) OR (St. Vincent and the Grenadines) OR (Sudan) OR (Suriname) OR (Swaziland) OR (Syrian Arab Republic) OR (Tajikistan) OR (Tanzania) OR (Thailand) OR (Timor-Leste) OR (East Timor) OR (Togo) OR (Tonga) OR (Tunisia) OR (Turkey) OR (Turkmenistan) OR (Tuvalu) OR (Uganda) OR (Ukraine) OR (Uruguay) OR (Uzbekistan) OR

(Vanuatu) OR (Venezuela) OR (Vietnam) OR (West Bank and Gaza) OR (Yemen) OR (Zambia) OR (Zimbabwe) OR (Latin America) OR (Central America) OR (Caribbean) OR (Eastern Europe) OR (South Asia) OR (Southeast Asia) OR (Africa)

Round A: Bibliographic Databases

The search term categories guided the search of relevant bibliographic databases. The ideal search strategy was the combination of A, B, and C, although the exact search terms were adapted to each database to exclude high numbers of irrelevant results. General statistics on the number of results yielded from each database are provided in table C.1.

Table C.1. Statistics on the Number of Results Yielded from Each Database

Database	Total Results	Potential Impact Evaluations
EconLit	39	14
Science Direct*	7,785 (3,350) 5,580 (4,976)	916 205
PopLine	1,591 (442)	103
Dialog	55	31
PubMed/MedLine*	1435 2200	121 105
ERIC	105	14
ArticleFirst	4	2
WorldCat	52	10

Note: Two separate searches were performed in Science Direct and PubMed because, when reviewing references from systematic reviews, the original search strategy had missed many relevant IEs that were in these two databases. Consequently a second, more exhaustive search was undertaken. Each row of numbers reflects a different search, and the total results are undoubtedly inflated due to duplicates between the two different searches.

From the total results for each database, titles, subjects, and abstracts were reviewed to identify the relevant IEs. As Science Direct and PopLine returned such a high number of results, an automatic export feature was used to export all of the results into an Excel file where irrelevant studies were quickly eliminated.[3] The number given in parenthesis is the number of studies for which the title and abstract were reviewed.

IE-Focused Organizations

Three organizations were identified that offered online databases of impact evaluations. Broad search terms were used to find publications on the following two websites and 13 relevant IEs were identified:

- Abdul Latif Jameel Poverty Action Lab (J-PAL)
- Innovations for Poverty Action (IPA)

Relevant Research Organizations

Institutions that are involved in relevant ECD research were identified, although they are not focused solely on impact evaluation. Their websites were searched for publications, and the results are presented in table C.2.

Table C.2. Statistics on the Number of Results Yielded from Each Database

Organization Name	Potential IEs
Population Council	6
IFPRI	13
RAND	10
IZA	5
GDN	4
BREAD	4

Supplemental Sources

While the majority of results were expected to be found through the bibliographic databases noted above, a number of searches in various supplemental sources were conducted. Additional potential IEs were obtained, and this method also served as an additional check on the comprehensiveness of the initial bibliographic search strategy.

The Lancet

ECD experts identified two relevant series from *The Lancet* and reviewed the references to identify potential impact evaluations. Given that the determination was made strictly on title, more emphasis was placed on inclusion and about 300 were identified. The series were Child Development in Developing Countries (2007) and Child Development in Developing Countries 2 (2011).

World Bank Databases

The bibliographic database searches yielded some impact evaluations of World Bank projects or implemented by World Bank staff. Nonetheless, World Bank databases were checked to identify additional IEs that included outcomes of interest. Reliance was placed primarily on an IE database compiled by IEG for the report entitled *The World Bank Group Impact Evaluations: Relevance and Effectiveness,* which includes IEs from the Development Impact Evaluation database, IEs identified by literature reviews for previous IEG reports, or other IEs provided to the IEG team by World Bank staff (IEG 2012). Also searched was the World Bank Safety Nets Publications Database. Together, these databases yielded 61 potential IEs.

Other

The remaining potential impact evaluations were found through seminars attended, personal knowledge of the team, or word of mouth.

Round B: Systematic Reviews

Examined were reference lists of 39 systematic reviews that focused on ECD. Each title was examined, followed by an assessment of the abstracts of those that seemed relevant. Those that passed this scrutiny were marked for a full-text review.

Round C: Most Prolific Authors and Snowballing

Subsequent to undertaking the extensive search in Rounds A and B, the team refined the scope of the report to focus on post-ECD effects of ECD interventions. Consequently, Round C was particularly important in the search process as the initial database search terms were concentrated on ECD outcomes.

Most Prolific Authors

The impact evaluations of ECD interventions that estimate effects on non-ECD outcomes were used to create a list of the most prolific authors who focused on post-ECD outcomes (table C.3). This was done by tabulating the number of times an author's name appeared in the byline of these studies. Their curriculum vitae were browsed to find additional relevant publications. Although those of Susan P. Walker, Susan M. Chang, and Christine A. Powell were not available online, most of their previous work had been done in conjunction with Sally Grantham-McGregor. In sum, 39 new potential IEs were identified, four of which were passed on for full coding.

Table C.3. Most Prolific Authors Who Focus on Post-ECD Outcomes

Names of Most Prolific Authors	
Sally Grantham-McGregor	Huiman Xie Barnhart
Susan P. Walker	Paul Gertler
Susan M. Chang	Michael S. Kramer
Christine A. Powell	John Hoddinott
Jere Behrman	John Maluccio

Snowballing

As part of the coding process, the reference list of each relevant impact evaluation was examined. Similar to the approach for the reference lists of the systematic reviews, relevant titles were identified, subjects and abstracts checked, and duplicates eliminated. As a result, 154 potential IEs were gathered, nine of which were selected for full coding.

2015 Update

The original search ended in fall, 2013, and was updated in February–March 2015. All of the bibliographic databases listed in table C.1, as well as the Development Impact Evaluation database, new systematic reviews, and the CVs of the most prolific authors were re-searched for

IEs published from 2013–2015. This produced more than 3000 results, 184 of which were exported for additional screening after a review of the title and abstract.

Screening Studies

Inclusion and exclusion criteria were applied to the titles, subjects, and abstracts of the studies to generate a list of potential IEs:

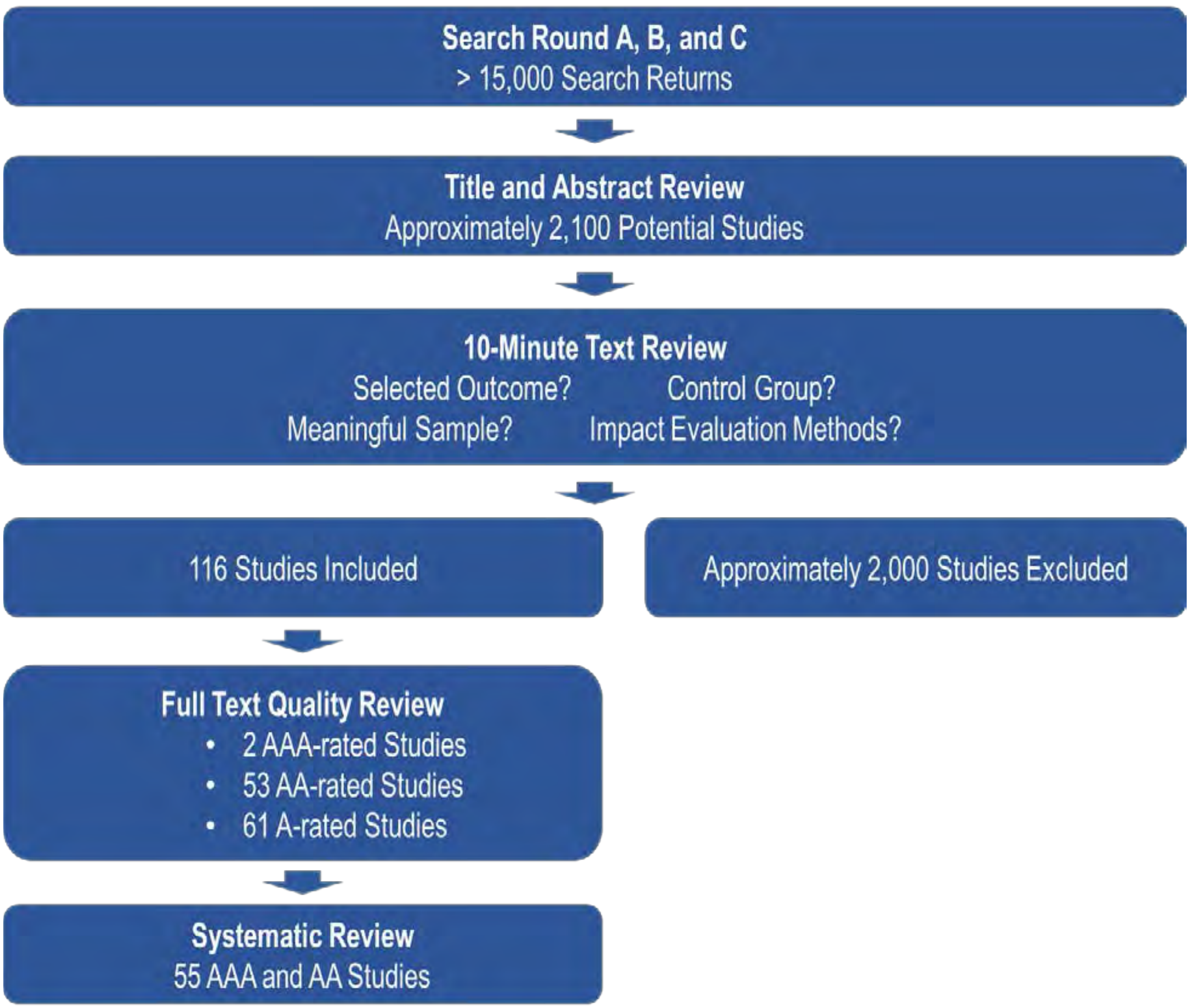
- Outcomes: Studies that evaluate ECD outcomes were included.
- Study design: Studies that evaluate interventions based on quantitative experimental or quasi-experimental IE design with a well-defined counterfactual were included.
- Location: Studies were selected that occurred in a country meeting the low- and middle-income specifications for the International Bank for Reconstruction and Development and the International Development Association.
- Language: The search focused on studies in English, though studies in Spanish, French, and Portuguese were found and included.
- Publication date: Studies published since January 1, 1990, and after were included.
- Unit of analysis: Studies that use regional or national time series data were excluded.
- Peer Review: IEs that have been subjected to peer review (for example, published in a quality journal or a book) or are in the process of eliciting feedback from the research community (such as working papers or papers presented in conferences) were be included.
- Nonclinical interventions: Following the advice in the World Bank's handbook *Impact Evaluation in Practice* (Gertler and others 2011) that results are most useful for government and development workers when they are the result of interventions that take place "under normal circumstances, using regular implementation channels," interventions were only included if they were implemented using local capacity.

The titles, subjects, and abstracts did not always provide enough information to determine if the study met the selection criteria, particularly regarding outcomes and study design. When unclear, the study was included as a potential IE for more consideration. A full text review of all potential IEs were performed, and those that qualified as IEs of the selected outcomes were coded (see coding strategy).

Search Results

IEG reviewed more than 15,000 search results across Rounds A, B, and C (figure C.1). From a title and abstract review of these, approximately 2,100 potential studies were identified. After a further 10-minute text review of these studies, more than 500 employed an impact evaluation method on a population from a middle- or low-income country and were related to ECD. Of these, 116 were included as IEs of the post-early childhood effect of early childhood interventions. Following a full-text review for quality and risk of bias, this report classified 116 studies into 2 AAA, 53 AA, and 61 A rated studies. The systematic review uses only the 55 AA and AAA studies for evidence.

Figure C.1. Search Results



References

Card, David, Jochen Kluve, and Andrew Weber. 2010. "Active Labor Market Policy Evaluations: A Meta-Analysis." *Economic Journal* 120: F452–F457.

Drabo, E.F., F. Perez-Arce, and J. Yoong. 2011. *Proven and Promising Interventions to Improve Adolescent Reproductive Health in Developing Countries: A Systematic Review*. Arlington, Va.: Rand Corporation.

Gertler, Paul J., Sebastián Martínez, Patrick Premand, Laura B. Rawlings, and Christel M.J. Vermeersch. 2011. *Impact Evaluation in Practice.* Washington, DC: World Bank.

IEG (Independent Evaluation Group). 2010. *What Can We Learn from Nutrition Impact Evaluations?* Washington, DC: World Bank.

______. 2011. *Evidence and Lessons Learned from Impact Evaluations on Social Safety Nets.* Washington, DC: World Bank.

______. 2012. *World Bank Group Impact Evaluations: Relevance and Effectiveness.* Washington, DC: World Bank.

______. 2013. *Delivering the Millennium Development Goals to Reduce Maternal and Child Mortality: A Systematic Review of Impact Evaluation Evidence.* Washington, DC: World Bank.

---

[1] The team first searched for impact evaluations (IEs) that provided estimates for any early childhood development (ECD) outcome. The full-text review was then used identify IEs that evaluated interventions that took place from birth to five years old and reported estimates for post-ECD outcomes.

[2] Initially, the team used a broad set of outcome terms, which resulted in the identification of studies measuring non-ECD outcomes. However, further research revealed that the outcome domains adequately identified studies reflecting a range of interventions. The following is a broader set of key terms used for the pilot study: (pre-primary education) OR (early stimulation) OR (nutrition) OR (pre-natal care) OR (parent education) OR (early childhood development) OR (early childhood education) OR (early childhood care) OR (immunization) OR (child health) OR (conditional cash transfer) OR (reduced drop-out) OR (cognitive development) OR (age appropriate school entry) OR (earnings) OR (reduced criminal behavior) OR (birth weight) OR (anthropometric indicators) OR (child mortality) OR (stunting) OR (breast feeding) OR (growth monitoring) OR (maternal mortality).

[3] To search studies in Excel, the title and abstract were exported and a search run that would identify if certain terms appeared in either. If the study contained any of these terms in the title or abstract, it was passed on for a manual title and abstract review. The terms included were: "early childhood development" "early childhood education" "early childhood care" "early child care" "infant development" "child cognitive development" "child linguistic development" "child socioemotional development" "child socio emotional development" "child physical development" "child growth" "child health" "child nutrition" "preschool" "pre school" "pre-school" "neonatal" "child development" "child" "infant" "under 6" "under 5" "under five" "under six."

Appendix D. Coding Protocol and Instruments

After identifying studies based on a title and abstract review as outlined in the search strategy, the potential impact evaluations were retrieved and read in full, and the following approach was used to code the documents. This coding strategy was based on the one used in "Delivering the Millennium Development Goals to Reduce Maternal and Child Mortality: A Systematic Review of Impact Evaluation Evidence."

Step 1: 10-Minute Review

Each study received a 10-minute review, or a brief full-text examination, to find the information described below. If the answer to any of the questions below was no, the study was not included for coding.

- Intervention of interest: Does the study evaluate an intervention that targeted either pregnant women or children under six years old.
- Counterfactual: Does the study use a counterfactual, that is, information on others who do not receive the treatment to proxy for what would have happened in the absence of the treatment? Studies that use time-series or before-after observations on the same treatment group but do not have a comparison group are not to be included. Studies without a control group but that provide convincing exogenous variation in the treatment were placed in a separate folder in EndNote for potential use.
- Impact evaluation method: Does the study utilize an appropriate impact evaluation method: randomized experiment, double or triple difference, matching, instrumental variable, regression discontinuity, or other IE method?
- Representative: Does the study take place in a country that meets the low- and middle-income specifications of the International Bank for Reconstruction and Development and the International Development Association, and outside of a controlled environment and in a real-world context (for example, not in a lab)?
- Effectiveness: Was the interventions implemented using local capacity so that it can be replicated by local implementers?
- Post-ECD outcomes: Does the study estimate the effect of the intervention on the recipients for outcomes that are measured after the beginning of primary school (approximately six years old).

Step 2: Quality Rating

For each study that passed the 10-minute review, the full text was read and a quality rating determined. Internal validity was the primary consideration, but also factored in were any other major concerns with the study (for example, data collection methods, sample size and representativeness, power, policy replicability).

Each study was double coded (two junior coders filled out the Quality and Evaluation Design section and provided a quality rating), and any disputes were settled after a third reading and rating by a senior coder.

The attached coding instrument provides each variable included in the Quality and Evaluation Design section, while key questions requiring additional guidance are outlined below.

- Assessment of internal validity/quality of evaluation design is done vis-à-vis the estimation strategy used to evaluate impacts (see box D.1). The starting point is then to first identify the evaluation design (randomized or quasi-experimental) and IE methods (difference-in-differences, matching, instrumental variables, or regression discontinuity) used to identify impacts by the study. Subsequently, reviewers assessed whether the relevant identification assumptions have been satisfied or adequately discussed, and coded this information as all, some, or none (that is, all, some, or none of the identification assumptions have been satisfied) for each of the methods used in study. Note that in adjacent columns where reviewers coded the extent to which assumptions have been satisfied, they were expected to document the reasons why they coded it as all, some, or none.
- Assessment of the strength and stability of findings, usually achieved through various types of robustness checks. After determining the main impact evaluation method, robustness checks were coded based on the following questions:
    - Did the IE use multiple estimation methods?
    - Did the IE use multiple specifications?
    - Did the IE perform other robustness analyses (for instance, falsification tests, alternative ways to measure the treatment, multiple control groups, sensitivity analysis and bounds)?
    - Was there a problem with missing data (for example, attrition above 40 percent, refusal to participate?)
    - Was the study likely to suffer from John Henry, Hawthorne or Pygmalion effects?
    - Did the IE protect against spillover, was it free of selective analysis or outcome reporting, and were the standard errors appropriate?
    - Were there any concerns about construct validity?
- Rate overall quality of IEs based on the assessment of internal validity, robustness of findings, and any other major concerns. Reviewers rated the

impact evaluations as being low (not meeting most of the criteria), medium (meeting some, but not all, of the criteria), or high (meeting most of the criteria).

Box D.1. Quality of Evaluation Design and Internal Validity

The quality of evaluation design is the most critical determinant of impact evaluation quality; it is the linchpin for estimating the share of the change in outcomes attributable to the intervention. The starting point for this exercise is identifying the evaluation design (experimental or quasi-experimental) and method (difference-in-difference, matching, instrumental variables, or regression discontinuity) used to identify program impacts. The next step is assessing whether the report provides a discussion of the assumptions or conditions under which the estimation method is valid:

Assumptions under randomized experiment: (i) balanced treatment and control groups (the two groups having no statistically significant difference in main baseline or time-invariant characteristics); and (ii) noncompliance or attrition (minimal incidence of beneficiaries not receiving treatment or leaving the program, and vice versa).

Assumptions under double difference: (i) parallel trending (the treatment and control groups progress similarly in terms of the outcomes of interests); and (ii) time-varying confounders (no time-variant variables that may affect the progress of the outcomes other than the intervention).

Assumptions under matching: (i) common support (the overlap in terms of propensity scores or matching variables between the treatment and control group); (ii) balancing checks (the treatment and control groups having no statistically significant difference in main observable characteristics); (iii) matching on outcomes or covariates (the variables used to match are not affected by the intervention); and (iv) selection on unobservables (there should be a discussion of potential selection bias due to unobservable differences between the treatment and control).

Assumptions under instrumental variables: (i) first stage tested (the relationship between the intervention and the instrument is statistically significant; F-test or Wald test); and (ii) exclusion restriction (the instrument affects the outcome only via the intervention).

Assumptions under regression discontinuity: (i) sorting around the assignment rule (beneficiaries tricking the rule to be eligible for the treatment); and (ii) balanced covariates at discontinuity (the two subgroups above and below the eligibility cutoff have statistically similar characteristics).

Step 3: Code Relevant Information

After determining a rating for each study, additional information from medium and high studies was coded. Low studies were not coded further. Since many of the programs were evaluated multiple times, each program was assigned a number, and the program information and questions on external validity were coded once for each program in a separate document. The program number was then used in the full coding document.

- Study information:
    - Country, author, year and type of publication

  - Type of World Bank involvement

- Program or intervention information:
  - Program or intervention name, description, and start and end dates
  - Intervention classification in up to three categories
  - Intervention duration, length of exposure, delivery modality, delivery location, implementer, and level of operation
  - Program targeting (by age and gender)
  - Adherence or take-up rate
  - Whether the intervention occurred in an low-income or fragile or conflict country
- External validity:
  - Whether the program was a pilot program
  - Barriers and enablers to intervention implementation, scaling up, and sustainability
- Model:
  - Did the authors discuss their logic model? If so, did it explicitly incorporate age?
- Data:
  - Sample size, data source, years of data collection, and length of evaluated exposure and of delays to implementation
  - Unit of analysis, use of retrospective data and baseline data
  - Sampling strategy and representativeness of the sample
- Cost analysis:
  - Presence of cost analysis
  - Cost analysis methodology and estimates, given by outcome
- Findings:
  - Specific outcome, the outcome domain, and the age group for which it was estimated
  - Baseline value and estimate type
  - Estimate, significance level, and interpretation
- Heterogeneous Effects:
  - Estimates of any heterogeneous effects and for what subgroups, given by outcome

The full dataset will be published separately and will include the full coding instrument.

Appendix E. Anthology of Identified ECD Systematic Reviews

Author	Title	Year	Source
Aiello and others	Effects of Hand Hygiene on Infectious Disease Risk in the Community Setting: A Meta-Analysis	2008	3ie
Ainsworth	What Can We Learn from Nutrition Impact Evaluations? Lessons from a Review of Interventions to Reduce Child Malnutrition in Developing Countries	2010	World Bank
Arnold and Colford	Treating Water with Chlorine at Point-Of-Use to Improve Water Quality and Reduce Child Diarrhea in Developing Countries: A Systematic Review and Meta-Analysis	2007	PubMed: *American Journal of Tropical Medicine and Hygiene*
Berti and others	A Review of the Effectiveness of Agricultural Interventions in Improving Nutrition Outcomes	2004	3ie (Public Health Nutrition)
Bhutta and others	Prevention of Diarrhea and Pneumonia by Zinc Supplementation in Children in Developing Countries: Pooled Analysis of Randomized Controlled Trials	1999	*Journal of Pediatrics*
Bhutta and others	Community-Based Interventions for Improving Perinatal and Neonatal Health Outcomes in Developing Countries: A Review of the Evidence	2005	*Journal of Pediatrics*
Bhutta and others	What Works? Interventions for Maternal and Child Undernutrition and Survival	2008	*Lancet*
Bhutta and others	Evidence-Based Interventions for Improvement of Maternal and Child Nutrition: What Can Be Done and at What Cost?	2013	*Lancet*
Lengeler	Insecticides-Treated Bed Nets and Curtains for Preventing Malaria	2004	3ie/Cochrane Collaboration
Dewey and Adu-Ararwuah	Efficacy and Effectiveness of Complementary Feeding Interventions in Developing Countries	2008	PubMed: *Maternal and Child Nutrition*
Engle and others	Strategies to Avoid the Loss of Developmental Potential in More than 200 Million Children in the Developing World	2007	*Lancet*
Engle and others	Strategies for Reducing Inequalities and Improving Developmental Outcomes for Young Children in Low and Middle Income Countries	2011	*Lancet*
Fewtrell and others	Water, Sanitation, and Hygiene Interventions to Reduce Diarrhoea in Less Developed Countries: A Systematic Review and Meta-Analysis	2005	*Lancet*
Gaarder, Glassman, and Todd	Conditional Cash Transfer: Unpacking the Causal Chain	2010	3ie

Gamble and others	Insecticide-Treated Nets for the Prevention of Malaria and Pregnancy: A Systematic Review of Randomised Controlled Trials	2007	*PLoS Medicine*
Grantham-McGregor and others	Effects of Integrated Child Development and Nutrition Interventions on Child Development and Nutritional Status	2014	*Annals of the New York Academy of Sciences*
Gunaratna and others	A Meta-Analysis of Community Based Studies on Quality Protein Maize	2010	3ie
Hundley and others	Are Birth Kits a Good Idea? A Systematic Review of the Evidence	2011	3ie
Iannotti and others	Iron Supplementation in Early Childhood: Heath Benefits and Risks	2006	PubMed: *American Journal of Clinical Nutrition*
Indad, Yakoob, and Bhutta	Impact of Maternal Education about Complimentary Feeding and Provision of Complimentary Foods on Child Growth in Developing Countries	2011	PubMed: *BMC Public Health*
Kramer and Kakuma	Energy and Protein Intake during Pregnancy	2010	3ie/Cochrane Collaboration
Lassi, Heider, and Bhutta	Community-Based Intervention Packages for Reducing Maternal and Neonatal Morbidity and Mortality and Improving Neonatal Outcomes	2010	Cochrane Library
Lassi and others	Systematic Review of Complementary Feeding Strategies amongst Children Less than Two Years of Age	2013	DFID
Leroy and others	The Impact of Daycare Programs on Child Health, Nutrition and Development in Developing Countries: A Systematic Review	2011	3ie
Leroy, Ruel, and Verhofstad	The Impact of Conditional Cash Transfer Programs on Child Nutrition: A Review of Evidence Using Program Theory Framework	2009	3ie
Manley, Gitter, Slavchevska	How Effective Are Cash Transfers at Improving Nutritional Status?	2013	*World Development*
Masset and others	A Systematic Review of Agricultural Interventions That Aim to Improve Nutritional Status of Children	2011	3ie
Maulik and Darmstadt	Community-Based Interventions to Optimize Early Childhood Development in Low Resource Settings	2009	PubMed: *Journal of Perinatology*
Nores and Barnett	Benefits of Early Childhood Interventions across the World: (Under) Investing in the Very Young	2010	ScienceDirect: *Economics of Education Review*
Piroska	The Positive Deviance/Hearth Approach to Reducing Child Malnutrition: Systematic Review	2011	PubMed: *Tropical Medicine and International Health*
Sachdev, Gera, and Nestel	Effect of Iron Supplementation on Physical Growth in Children: Systematic Review of Randomized Controlled Trials	2005	WHO: (Public Heath Nutrition)
Sguassero and others	Community-Based Supplementary Feeding for Promoting the Growth of Children under Five Years of Age in Low and Middle Income Countries (*Updated SR - older version published in 2005*)	2012	Cochrane Library

Tanner and others	Delivering the MDGs on Maternal and Child Mortality	2013	IEG
Waddington and others	Effectiveness and Sustainability of Water, Sanitation and Hygiene in Combating Diarrhoea	2009	3ie
Walker and others	Child Development: Risk Factors for Adverse Outcomes in Developing Countries	2007	*Lancet*
Walker and others	Inequality in Early Childhood: Risk and Protective Factors for Early Childhood Development	2011	*Lancet*
Walker	Promoting Equity through ECD Interventions for Children from Birth through Three Years of Age	2011	World Bank

Appendix F. List of Studies Given an A Rating

Aboud, Frances E., and Kamal Hossain. 2011. "The Impact of Preprimary School on Primary School Achievement in Bangladesh." *Early Childhood Research Quarterly* 26 (2): 237–246.

Aboud, Frances E., Kamal Hossain, and Chloe O'Gara. 2008. "The Succeed Project: Challenging Early School Failure in Bangladesh." *Research in Comparative and International Education* 3 (3).

Aguilar, R., and R. Tansini. 2010. "Preschool Education and School Performance the Case of Public Schools in Montevideo." Working Papers in Economics No. 434, University of Gothenburg, Gothenburg, Sweden.

Behrman, Jere R., Maria C. Calderon, Samuel H. Preston, John F. Hoddinott, Reynaldo Martorell, and Aryeh D. Stein. 2009. "Nutritional Supplementation in Girls Influences the Growth of Their Children." *American Journal of Clinical Nutrition* 90 (5): 1372–1379.

Bloom, D.E., D. Canning, and E.S. Shenoy. 2012. "The Effect of Vaccination on Children's Physical and Cognitive Development in the Philippines." *Applied Economics* 44 (21): 2777–2783.

Buckley, Gillian J., Laura E. Murray-Kolb, Subama K. Khatry, Steven C. Leclerq, Lee Wu, Keith P. West Jr., and Parul Christian. 2013. "Cognitive and Motor Skills in School-Aged Children Following Maternal Vitamin a Supplementation During Pregnancy in Rural Nepal: A Follow-Up of a Placebo-Controlled, Randomised Cohort." *The BMJ Open* 3 (5). pii: e002000. doi: 10.1136/bmjopen-2012-002000.

de Brauw, A., D. Gilligan, J. Hoddinott, V. Moreira, and S. Roy. 2012. The Impact of Bolsa Família on Child, Maternal, and Household Welfare. Washington, DC: International Food Policy Research Institute.

Chang, S.M., S.P. Walker, S. Grantham-McGregor, and C.A. Powell. 2002. "Early Childhood Stunting and Later Behaviour and School Achievement." *Journal of Child Psychology and Psychiatry* 43 (6): 775–783.

Chaudhuri, Anoshua. 2008. "Revisiting the Impact of a Reproductive Health Intervention on Children's Height-for-Age with Evidence from Rural Bangladesh." *Economic Development and Cultural Change* 56 (3): 619–656.

Christian, P., L.E. Murray-Kolb, S.K. Khatry, J. Katz, B.A. Schaefer, P.M. Cole, S.C. Leclerq, and J.M. Tielsch. 2010. "Prenatal Micronutrient Supplementation and Intellectual and Motor Function in Early School-Aged Children in Nepal." *JAMA* 304 (24): 2716–2723.

Christian P., M.E. Morgan, L. Murray-Kolb, S.C. LeClerq, S.K. Khatry, B. Schaefer, P.M. Cole, J. Katz, and J.M. Tielsch. 2011. "Preschool Iron-Folic Acid and Zinc Supplementation in Children Exposed to Iron-Folic Acid in Utero Confers No Added Cognitive Benefit in Early School-Age." *Journal of Nutrition* 141 (11): 2042–2048.

Congdon, Eliza L., Alissa Westerlund, Cecilia R. Algarin, Patricio D. Peirano, Matthew Gregas, Betsy Lozoff, and Charles A. Nelson. 2012. "Iron Deficiency in Infancy Is Associated with Altered Neural Correlates of Recognition Memory at 10 Years." *Journal of Pediatrics* 160 (6): 1027–1033.

Conlisk, A.J., H.X. Barnhart, R. Martorell, R. Grajeda, and A.D. Stein. 2004. "Maternal and Child Nutritional Supplementation Are Inversely Associated with Fasting Plasma Glucose Concentration in Young Guatemalan Adults." *Journal of Nutrition* 134 (4): 890–897.

Cueto, S., and J. Diaz. 1999. "Impacto de la Educación Inicial en el Rendimiento en Primer Grado de Primaria en Escuelas Públicas Urbanas de Lima." *Revista de Psicología* 17: 74–91.

de Felicio, Fabiana, Rafael Terra, and Ana Carolina Zoghbi. 2012. "The Effects of Early Childhood Education on Literacy Scores Using Data from a New Brazilian Assessment Tool." *Estudos Economicos* 42 (1): 97–128.

Field, Erica, Omar Robles, and Maximo Torero. 2009. "Iodine Deficiency and Schooling Attainment in Tanzania." *American Economic Journal: Applied Economics* 1 (4): 140–169.

Foster, Andrew D., and Nikhil Roy. 1997. "The Dynamics of Education and Fertility: Evidence from a Family Planning Experiment." Draft. http://adfdell.pstc.brown.edu/papers/educ.pdf.

Gautam, Hazarika, and Vejoya Viren. 2013. "The Effect of Early Childhood Developmental Program Attendance on Future School Enrollment and Grade Progression in Rural North India." *Economics of Education Review* 34: 146–161.

Grantham-McGregor, S.M., C. Powell, S.P. Walker, S.M. Chang, and P. Fletcher. 1994. "The Long-Term Follow-Up of Severely Malnourished Children Who Participated in an Intervention Program." *Child Dev* 65 (2 Spec. No.): 428–439.

Grantham-McGregor, S.M., S.P. Walker, S.M. Chang, and C.A. Powell. 1997. "Effects of Early Childhood Supplementation with and without Stimulation on Later Development in Stunted Jamaican Children." *American Journal of Clinical Nutrition* 66 (2): 247–253.

Griffen, Andrew S. 2013. "Height and Calories in Early Childhood." GCC Working Paper Series, GCC 13-09, University of Pennsylvania, Philadelphia.

Haas, J.D., E.J. Martínez, S. Murdoch, E. Conlisk, J.A. Rivera, and R. Martorell. 1995. "Nutritional Supplementation during the Preschool Years and Physical Work Capacity in Adolescent and Young Adult Guatemalans." *Journal of Nutrition* 125 (Suppl. 4): 1078S–1089S.

Hoddinott, John, John A. Maluccio, Jere R. Behrman, Rafael Flores, and Reynaldo Martorell. 2008. "Effect of a Nutrition Intervention During Early Childhood on Economic Productivity in Guatemalan Adults." *The Lancet* 371 (9610): 411–416.

Johnes, G. 2008. "Early Year Education and Subsequent Schooling in Rural India: An Economic Analysis." Working Paper No. 018, Department of Economics, Lancaster University, Lancaster, U.K.

Joshi, Shareen, and T. Paul Schultz. 2007. "Family Planning as an Investment in Development: Evaluation of a Program's Consequences in Matlab, Bangladesh." Center Discussion Paper No. 951, Economic Growth Center, Yale University, New Haven, Conn.

Jukes, M.C., M. Pinder, E.L. Grigorenko, H.B. Smith, G. Walraven, E.M. Bariau, R.J. Sternberg, L.J. Drake, P. Milligan, Y.B. Cheung, B.M. Greenwood, and D.A. Bundy. 2006. "Long-Term Impact of Malaria Chemoprophylaxis on Cognitive Abilities and Educational Attainment: Follow-Up of a Controlled Trial." *PLos Clin Trials* 1 (4): e19.

Kagitcibasi, Cigdem, Diane Sunar, and Sevda Bekman. 2001. "Long-Term Effects of Early Intervention: Turkish Low-Income Mothers and Children." *Journal of Applied Developmental Psychology* 22 (4): 333–361.

Kagitcibasi, Cigdem, Diane Sunar, Sevda Bekman, Nazli Baydar, and Zeynep Cemalcilar. 2009. "Continuing Effects of Early Enrichment in Adult Life: The Turkish Early Enrichment Project 22 Years Later." *Journal of Applied Developmental Psychology* 30 (6): 764–779.

Kinra, S., K.V. Rameshwar Sarma, Ghafoorunissa, V.V. Rao Mendu, R. Ravikumar, V. Mohan, I.B. Wilkinson, J.R. Cockcroft, G. Davey Smith, and Y. Ben-Shlomo. 2008. "Effect of Integration of Supplemental Nutrition with Public Health Programmes in Pregnancy and Early Childhood on Cardiovascular Risk in Rural Indian Adolescents: Long Term Follow-Up of Hyderabad Nutrition Trial." *The BMJ* 337: a605.

Klein, P.S., and H. Rye. 2004. "Interaction-Oriented Early Intervention in Ethiopia: The Misc Approach." *Infants and Young Children* 17: 340–354.

Kulkarni, Bharati, Hannah Kuper, K.V. Radhakrishna, Andrew P. Hills, Nuala M. Byrne, Amy Taylor, Ruth Sullivan, Liza Bowen, Jonathan C. Wells, Yoav Ben-Shlomo, George D. Smith, Shah Ebrahim, and Sanjay Kinra. 2014. "The Association of Early Life Supplemental Nutrition with Lean Body Mass and Grip Strength in Adulthood: Evidence From APCAPS." *American Journal of Epidemiology* 179 (6): 700–709.

Li, H., H.X. Barnhart, A.D. Stein, and R. Martorell. 2003. "Effects of Early Childhood Supplementation on the Educational Achievement of Women." *Pediatrics* 112 (5): 1156–1162.

Lozoff, Betsy, Marcela Castillo, Katy M. Clark, and Julia B. Smith. 2012. "Follow-Up of a Randomized Controlled Trial of Iron-Fortified (12.7 mg/L) vs. Low-Iron (2.3 mg/L) Infant Formula: Developmental Outcome at 10 Years." *Archives of Pediatrics and Adolescent Medicine* 166 (3): 208–215.

Lozoff, Betsy, Marcela Castillo, Katy M. Clark, Julia B. Smith, and Julie Sturza. 2014. "Iron Supplementation in Infancy Contributes to More Adaptive Behavior at 10 Years of Age." *Journal of nutrition* 144 (6): 838–845.

Mackintosh, U., D. Marsh, and D. Schroeder. 2002. "Sustained Positive Deviant Child Care Practices and Their Effects on Child Growth in Viet Nam." *Food and Nutrition Bulletin* 23 (Suppl. 4): 18–27.

Maluccio, J., J. Hoddinott, J. Behrman, R. Martorell, A. Stein, and A. Quisumbing. 2006. "The Impact of an Experimental Nutritional Intervention in Childhood on Education among Guatemalan Adults." FCND Discussion Paper 207, International Food Policy Research Institute, Washington, DC.

McDermott, Jennifer M., Alissa Westerlund, Charles H. Zeanah, Charles A. Nelson, and Nathan A. Fox. 2012. "Early Adversity and Neural Correlates of Executive Function:

Implications for Academic Adjustment." *Developmental Cognitive Neuroscience* 2 (1): S59–S66.

McDermott, Jennifer M., Sonya Troller-Renfree, Ross Vanderwert, Charles A. Nelson, Charles H. Zeanah, and Nathan A. Fox. 2013. "Psychosocial Deprivation, Executive Functions, and the Emergence of Socio-Emotional Behavior Problems." *Frontiers in Human Neuroscience* 7:167.

McLaughlin, Katie A., Margaret A. Sheridan, Warren Winter, Nathan A. Fox, Charles H. Zeanah, and Charles A. Nelson. 2014. "Widespread Reductions in Cortical Thickness Following Severe Early-Life Deprivation: A Neurodevelopmental Pathway to Attention-Deficit/Hyperactivity Disorder." *Biological Psychiatry* 76 (8):629–638.

McLaughlin, Katie A., Nathan A. Fox, Charles H. Zeanah, and Charles A. Nelson. 2011. "Adverse Rearing Environments and Neural Development in Children: The Development of Frontal Electroencephalogram Asymmetry." *Biological Psychiatry* 70 (11): 1008–1015.

Nelson, Charles A., Alissa Westerlund, Jennifer Martin McDermott, Charles H. Zeanah, and Nathan A. Fox. 2013. "Emotion Recognition Following Early Psychosocial Deprivation." *Development and Psychopathology* 25 (2): 517–525.

Nobles, J., and E. Frankenberg. 2009. "Mothers' Community Participation and Child Health." *Journal of Health and Social Behavior* 50 (1): 16–30.

Pickett, Kate E., Jere D. Haas, Scott Murdoch, Juan A. Rivera, and Reynaldo Martorell. 1995. "Early Nutrition Supplementation and Skeletal Maturation in Guatemalan Adults." Journal of Nutrition 125 (4): 1097S–1103S.

Pollitt, E., A. Jahari, and H. Walka. 2000. "A Developmental View of the Effects of an Energy and Micronutrient Supplement in Undernourished Children in Indonesia." *Eur J Clin Nutr* 54 (Suppl 2.): S107–S113.

Pollitt, E., K.S. Gorman, P.L. Engle, R. Martorell, and J. Rivera. 1993. "Early Supplementary Feeding and Cognition: Effects over Two Decades." Monogr Soc Res Child Dev 58 (7): 1–99; 111–118.

Pollitt, E., K.S. Gorman, P.L. Engle, J.A. Rivera, and R. Martorell. 1995. "Nutrition in Early Life and the Fulfillment of Intellectual Potential." *Journal of Nutrition* 125 (Suppl. 4): 1111S–1118S.

Pollitt, E., W.E. Watkins, and M.A. Husaini. 1997. "Three-Month Nutritional Supplementation in Indonesian Infants and Toddlers Benefits Memory Function 8 Y Later." American Journal of Clinical Nutrition 66 (6): 1357–1363.

Ramakrishnan, U., H. Barnhart, D.G. Schroeder, A.D. Stein, and R. Martorell. 1999. "Early Childhood Nutrition, Education, and Fertility Milestones in Guatemala." *Journal of Nutrition* 129: 2196–2202.

Raudenbush, S.W., S. Kidchanapanish, and S.J. Kang. 1991. "The Effects of Preprimary Access and Quality on Educational-Achievement in Thailand." *Comparative Education Review* 35 (2): 255–273.

Rivera, Juan A., Reynaldo Martorell, Marie T. Ruel, Jean-Pierre Habicht, and Jere D. Haas. 1995. "Nutritional Supplementation during Preschool Years Influences Body Size and Composition of Guatemalan Adolescents." *Journal of Nutrition* 125 (Suppl. 4): 1068S–1077S.

Rodrigues, C.G., C.X. C. Pinto, and D.D. Santos. 2010. "The Impact of Daycare Attendance on Math Test Scores for a Cohort of 4th Graders in Brazil." No. 290, Textos para discussão, Escola de Economia de São Paulo, Getulio Vargas Foundation, Brazil.

Schultz, Paul T. 2009. "How Does Family Planning Promote Development? Evidence from a Social Experiment in Matlab, Bangladesh, 1977–1996." Paper presented at Population Association of America meetings, Detroit, Mich., April 30, 2009.

Sheridan, Margaret A., Nathan A. Fox, Charles H. Zeanah, Katie A. McLaughlin, and Charles A. Nelson. 2012. "Variation in Neural Development as a Result of Exposure to Institutionalization Early in Childhood." *Proceedings of the National Academy of Sciences of the United States of America* 109 (32): 12927–12932.

Sinha, Nistha. 2005. "Fertility, Child Work, and Schooling Consequences of Family Planning Programs: Evidence from an Experiment in Rural Bangladesh." *Economic Development and Cultural Change* 54 (1): 97–128.

Super, C.M., M.G. Herrera, and J.O. Mora. 1990. "Long-Term Effects of Food Supplementation and Psychosocial Intervention on the Physical Growth of Colombian Infants at Risk of Malnutrition." *Child Development* 61 (1): 29–49.

Stein, A.D., M. Wang, A. DiGirolamo, R. Grajeda, U. Ramakrishnan, M. Ramírez-Zea, K.M. Yount, and R. Martorell, R. 2008. "Nutritional Supplementation in Early Childhood, Schooling and Intellectual Functioning in Adulthood: A Prospective Study in Guatemala." *Arch Pediatr Adolesc Med* 162 (7): 612–618.

Stein, A. D., M. Wang, M. Ramirez-Zea, R. Flores, and R. Grajeda. 2006."Exposure to a Nutrition Supplementation Intervention in Early Childhood and Risk Factors for Cardiovascular Disease in Adulthood: Evidence from Guatemala." *American Journal of Epidemiology* 164, 12: 1160–1170.

Taiwo, A.A., and J.B. Tyolo. 2002. "The Effect of Pre-School Education on Academic Performance in Primary School: A Case Study of Grade One Pupils in Botswana." *International Journal of Educational Development* 22 (2): 169–180.

Watanabe, K., R. Flores, J. Fujiwara, and L.T. Huong Tran. 2005. "Early Childhood Development Interventions and Cognitive Development of Young Children in Rural Vietnam." *Journal of Nutrition* 135 (8): 1918–1925.

Webb, A.L., A.J. Conlisk, H.X. Barnhart, R. Martorell, R. Grajeda, and A.D. Stein. 2005. "Maternal and Childhood Nutrition and Later Blood Pressure Levels in Young Guatemalan Adults." *Int J Epidemiol* 34 (4): 898–904.

Zuilkowski, Stephanie S., and Matthew C. H. Jukes. 2014. "Early Childhood Malaria Prevention and Children's Patterns of School Leaving in The Gambia." *British Journal of Educational Psychology* 84 (3): 483–501.

Appendix G. Intervention-Outcome Gap Map

Similar to Figure 11.1, this gap map shows understudied interventions and outcomes.

Intervention Type	Intervention Categories		Cognitive Development	Language Development	Socioemotional Development	Physical Development	Schooling	Labor Market	Number of Unique Studies	Number of Unique Projects
Nutrition	Mother	Counseling on adequate diet during pregnancy								
		Micronutrients and iron-folic acid: supplementation and fortification for	2	1		8	2		8	4
		Exclusive breastfeeding promotion	1	1	1	5	1		8	1
	Children	Complementary feeding								
		Supplemental feeding (preschool, center-base and/or take home rations)	3	3	3	3	4		10	5
		Counseling on optimal feeding practices:	2	3	2	2	4		5	2
		Therapeutic zinc supplementation for diarrhea								
		Growth monitoring and promotion - prevention and treatment for acute malnutrition								
		Micronutrients and fortification for children	1	1		1	1		1	1
Health	Pregnancy, delivery and postnatal	Antenatal visits								
		Attended delivery	1				1		1	1
	Disease treatment	Disease prevention								
		Deworming	1	1		1			1	1
	Access to health care	Planning for family size and spacing	1			1	1		1	1
		Well child visits, growth monitoring, screening for developmental delays	4	4	2	3	5		7	4
		Immunization	1			1	1		1	1
	Prevention and maternal depression									
Water, Sanitation	Water and sanitation	Access to safe water					1		1	1
		Hygiene or hand-washing			1	1			2	1
		Adequate sanitation	1	1					1	1
Education	Parent support program		5	4	6	6	3		15	5
	Stimulation		7	6	10	6	2	1	18	3
	Early childhood and pre-primary programs	Quality teaching, programming or curricula	1	1	2	2	4		6	5
		Pre-school Infrastructure	1	1	1	1	2		2	2
	Transition to quality primary schools									
Social Protection	Birth registration									
	Parental leave and adequate childcare or daycare						1		1	1
	Social assistance transfer programs	Conditional	3	4	2	3	4		6	3
		Unconditional/targeted income support (child grants, etc)	1	1		1	1		1	1
	Child protection interventions	Orphans and fostering	2	1	6	4			10	1
Governance	Governance reflecting ECD interests	Women's political reservation	1	1					1	1
	Policy or regulation in nutrition, health, education and social protection									
		Number of Unique Studies	21	19	15	27	19	1	55	
		Number of Unique Projects	16	15	7	14	16	1		25

Intervention Type / Outcome Domains