**WORKING PAPER** · NO. 2024-119

# Perry Preschool at 50: What Lessons Should Be Drawn and Which Criticisms Ignored?

*Alison W. Baulos, Jorge Luis García, and James J. Heckman*

SEPTEMBER 2024

THE UNIVERSITY OF CHICAGO
BECKER FRIEDMAN INSTITUTE FOR ECONOMICS

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

Perry Preschool at 50: What Lessons Should Be Drawn and Which Criticisms Ignored?
Alison W. Baulos, Jorge Luis García, and James J. Heckman
September 2024
JEL No. C53, I24, I32, J15


## ABSTRACT

The Perry Preschool Project, the longest-running experimental study of an early childhood education program, demonstrates how such interventions can yield long-term personal, societal, and intergenerational benefits for disadvantaged populations. The evidence is clear: investments in high-quality early childhood education and parental engagement can deliver returns even 50 years later. The program's findings remain scientifically robust, particularly when analyzed through rigorous small-sample inference methods. The program's findings also contradict common criticisms of preschool, as, when measured correctly, treatment effects on IQ do not fadeout. This paper draws insights from both the original founders and recent empirical studies, emphasizing the critical role of parental involvement in early education. The authors advocate for a scientific agenda focused on understanding the mechanisms behind treatment effects, rather than replicating specific programs. The analysis also underscores the broader implications of early childhood interventions for social mobility and human capital formation. Analysts of early childhood education should recognize that although credentials and formal curricula contribute to successful programs, the true measure of quality lies in adult-child interactions, which play an essential role.



Alison W. Baulos
University of Chicago
abaulos@uchicago.edu

Jorge Luis García
Department of Economics
Texas A&M University
2935 Research Parkway, Office 275
College Station, TX 77843
and NBER
jlgarcia@tamu.edu

James J. Heckman
Center for the Economics of
Human Development
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and IZA
and also NBER
jjh@uchicago.edu

The Perry Preschool study is a deep intellectual well that keeps giving. It is the longest-running experimental study of an early childhood intervention. The next wave—and we earnestly hope there will be one—will take participants into their golden years. The Perry study informs how early childhood interventions produce long-run intergenerational growth. No other study has its range or provides a guide to the life-cycle impacts of early childhood programs. Its detailed and insightful data collection instructs scholars and practitioners about the dynamics of early childhood skill formation. The study highlights the mechanisms that promote successful lives for disadvantaged children.

Despite its many merits, some scholars and policymakers, and a recent paper in Science magazine, dismiss the Perry findings and argue that the study is too small in sample size and not relevant to early childhood education today. The mentality that dismisses Perry replaces serious science with a statistical pseudoscience that seeks the "best" program among a batch of candidate programs evaluated by random assignment. This line of argument ignores the large body of research that applies rigorous, exact small sample inference that confirms the conclusions of a more conventional large sample statistical methodology. Even more importantly, this judgment ignores the primary goal of any scientific study: to understand the mechanisms that produce treatment effects.

This essay makes three general points. (1) We first review the perceptive commentary on the program by the founders of Perry. They qualitatively enunciate the ingredients of the program, which later quantitative work confirms. Early on, the Perry pioneers realized that engaging the parent actively in the learning life of the child is an essential ingredient in successful early childhood programs. (2) We next illustrate the severe limitations of an approach to child development that seeks to identify and replicate successful programs previously implemented. We propose an alternative approach that builds new programs based on effective *mechanisms* adapted to specific contexts. (3) We then summarize the lessons learned from the full body of Perry data through age 50.

## 1. *The Principles Underlying the Perry Preschool Program*

The Perry Preschool Study was founded on the belief that parent-child and teacher-child interactions are critical to the intellectual development of a child. The program actively engaged

parents in the learning life of their child. It did not follow a rigid formula in constructing its curriculum. The program enriched the environments of young children and fostered learning and active parenting in various ways. The program's teachers promoted scaffolding of children in the classroom and supplemented it with weekly afternoon visits with each child.

These visits took place both in and out of the home and gave individualized attention to the child. They broadened the child's background knowledge and encouraged applying newly mastered skills (e.g., trips to a library, museum or zoo). For many children in the program, home environments were substantially supplemented. The program worked directly with the parents and took their feedback on the design of the program. Weikart wrote:

> "*It became evident that when cooperation and participation of the parent were secured, the child's motivation to learn was considerably augmented… The mother's participation proved to be helpful both in stimulating the child and in enhancing her own understanding of the teaching process."*
>
> —Weikart (1966)

Afternoon visits with the child enhanced the benefits of preschool classroom activities, facilitated assessing a child's skill development and fostered relationships with families. (Wakabayashi et al., 2024). In the at-home visits, teachers engaged with the mother, siblings, or other caregivers in the household when available to stimulate and play with the child using activities from the preschool program and to suggest and model ways to incorporate science and math into everyday routines, such as food preparation or plant care. Through these visits, teachers learned about the children's home environments and aligned teaching strategies to accommodate them. Warm, frequent, language-rich, and attenuated interactions between teachers and children, as well as mutual engagement between the teachers and parents and between the parents and children, were critical to the success of the entire enterprise.

Parental engagement in the child's learning was also promoted through monthly group meetings. These informal and relaxed gatherings, which did not include classroom teachers or administrators, let parents exchange views about parenting and their children's education. Mothers visited museums the children had visited so they could directly share in their children's learning experiences. Fathers requested their semi-monthly gathering, and they, too, took an active role in their child's educational development. For example, a craft project for fathers gave them an opportunity to make educational toys for the classroom. In Waves 0, 1 and 2, 77% of

mothers and 80% of fathers in the experimental group attended at least one group meeting, and 55% of mothers and 50% of fathers attended over half of the meetings (Weikhart et al., 1964).[1]

Weikart often spoke about educators' obsession with IQ scores and program evaluation based on standardized tests in the field of education. The program was launched in the heyday of cognitive psychology. Teachers and administrators often face the same environment today. This focus on cognition crowded out the essential features of the program. Weikart described the Perry approach in simple terms:

> *"The [Perry] program is a permissive but teacher-structured, intended to guide the children toward increased cognitive development." (Weikart, 1966).*
>
> *"Emphasis was placed on the teacher's flexibility in gearing classroom activities to individual children's level of development. Heavier emphasis was placed on verbal stimulation and interaction, socio-dramatic play, and field trips than on social behavior and other traditional concerns of nursery schools." (Weikart et al, 1971).*

As a leader in early childhood education, the HighScope group continued its research beyond the Perry Preschool Study. It implemented a project to compare curriculum approaches. The Curriculum Demonstration Study (1967–1970) compared (a) DISTAR (Direct Instructional System for Teaching and Remediation), (b) Perry (the Perry Preschool program had evolved into the HighScope Cognitively Oriented Curriculum), and (c) a traditional nursery school model that prioritized play and social interactions. Each of the three treatment groups also received bi-weekly 90-minute home visits from one of the children's teachers. Teachers and administrators at HighScope felt that an educationally supportive home environment was essential and would foster growth in child skills beyond preschool.

The comparison study found that all three curriculum types boosted the IQs of disadvantaged children. However, when the children were followed to age 15, the Direct Instruction group had significantly more delinquency, including five times as many acts of property violence and lower self-worth, with 49% were diagnosed by their school as needing special services for emotional problems (Schweinhart, 1986). Similar negative effects were observed when participants were 23 (Schweinhart, 1997).

[1] Data for later waves are not available.

An article in the New York Times discussed the comparison study and denounced "early education pressure cookers." The author concluded,

> *"The new study goes a long way to show that the question is no longer whether early childhood education is desirable, but rather how to eliminate undesirable side effects (Hechinger, 1986)."*

Today's debates about preschool classrooms echo these discussions: what curricula should be used, what criteria make something "evidence-based" and therefore eligible for public funding, what teacher training and qualifications are required, and how program "effectiveness" should be determined. The research demonstrating that Perry promoted a variety of skills for disadvantaged children provides useful information to these discussions.

Scholars and policymakers should also look to the Perry Preschool given negative experiences for some children and pressures on teachers in today's preschool classrooms. For example, a high number of Black boys are expelled or suspended from preschool and early childhood programs. In 2020-2021, Black boys accounted for 9% of preschool enrollment but represented 23% of preschool children who received one or more out-of-school suspensions and 20% of preschool children who were expelled (U.S. Department of Education, *2020-21 Civil Rights Data*).[2] It is notable that in Perry, no children were suspended or expelled. This was intentional. It is a feature of many HighScope classrooms today. In his memoir, Weikart described that within the first six weeks of Perry, a "big-for-his-age" four-year-old boy threw a chair across the room, which startled everyone.

> *"In a very real sense, young [he] was the instigator of the HighScope Curriculum, for he experienced what we initially had to offer and found it lacking. Our choice was either to blame him or to reform our practice. Our decision was to accept that program opportunities and management, not the child, were the problem" (Weikart, 2004).*

This flexibility in the classroom and a whole child approach was a key aspect of Perry.

> *"With a renewed sense of purpose, the teachers and other special services staff went to work on developing a unified curriculum. It was clear that what we wanted didn't fit in a continuum of rigid lessons or, at the other end, of wholly free play. Instead, I wanted teachers to have ideas that enabled children to play as they wished and allowed the teachers to talk with children to develop their language skills" (Weikart, 2004).*

There was always tension between a focus on structured activities designed to promote cognitive growth and a focus on play activities to promote socio-emotional growth. To have a chance at

[2] These data do not include privately supplied programs.

scaling up, the program needed to demonstrate impacts on test scores. In all activities, Perry prioritized relationships and connections. Constance Kamii, a researcher on the project, brought back specific activities from Piaget's laboratory, but Weikart discontinued their use because teachers found them too rigid (Beatty, 2013). He coined the phrase "verbal bombardment" as a guide to the teachers on how to interact with children, but de-intensified the focus on language after the children "complained that the teachers talked too much" (Weikart, 2004). In this way the program was guided by the latest science on learning, but also on the feedback of the teachers and children.

In the Perry Preschool, social and emotional adjustment to the preschool classroom and the adults in it was recognized as a necessary foundation for the later intellectual growth of the child. The real-world consequences of boosting these socioemotional skills are chronicled in Section 3.

## 2. *How Not to Use the Perry Preschool Studies*

Studies of the Perry Preschool offer general lessons, rather a lesson about the effectiveness of a specific program. Evaluations of programs often suggest that the specifics of that program should be scaled up. This narrow view takes an "up or down" approach, where curriculum programs are either accepted or rejected based on their performance in treatment-control comparisons. Analyses conducted in this way are statistical exercises, not science. An "up or down" approach recommends which programs to scale up when a study shows positive results without considering the broader context to which the program might be applied or alternative approaches. In contrast, our approach seeks to place the findings from Perry in a broader context. We aim to understand how the findings from Perry can guide early childhood and preschool policy today, not how to replicate Perry as *the policy.*

Comparing one program to another is not sufficient. To effectively guide public policy, serious scholars should examine mechanisms in successful omnibus programs and isolate their commonalities. Science aims to understand the mechanisms producing observed outcomes. Perry enables analysts to discover the mechanisms that produce positive personality and social development. Perry is a guide to the ingredients of an effective approach.

### *2.1. Is Perry Irrelevant Today?*

García and Heckman (2023) examine the representativeness of participants in both the Perry study and a similar program, the Carolina Abecedarian Project (Ramey, 1984). They note that the samples from Perry and Abecedarian represent disadvantaged, low-income populations that would benefit from early childhood programs. Specifically, they indicate that 10% of children today are eligible for Perry based on the most recent US census data, which reflects the percentage of children born in households meeting the program's eligibility criteria. García and Heckman (2023) argue that the similarity in eligibility criteria across these programs suggests that a similar percentage of children in the US would be eligible for Abecedarian. Further, many Head Start and early childhood education (including preschool) classrooms today use HighScope or Learning Games, developed by the founders of the Perry and Abecedarian studies (Sparling, Lewis, & Ramey, 2001). Therefore, the findings from Perry and Abecedarian have immediate policy relevance and applicability.

### *2.2. Confronting the Critics with Hard Facts*

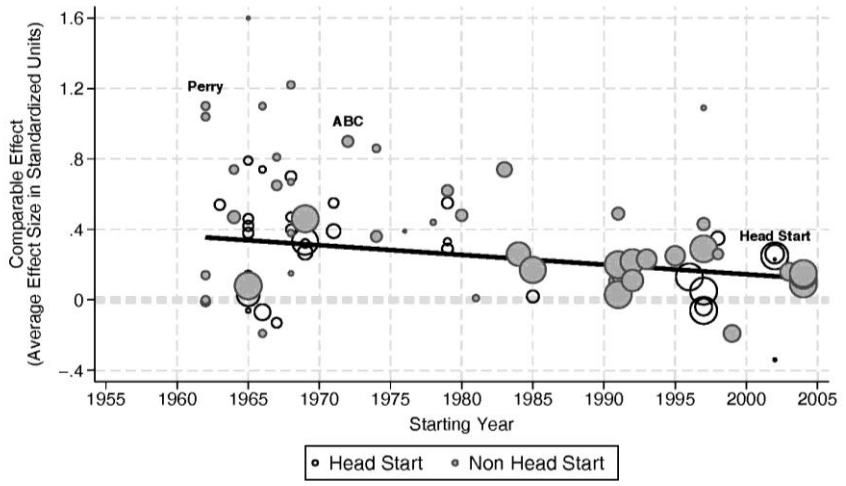
Persistent and influential critics of the research on Perry claim, among other things, that the program is too small, too old, and that times have changed (eg, Burchinal et al., 2024). Indeed, the world is different today than when Perry was implemented. Family sizes are smaller, there is a larger proportion of single parents, and childcare is more available. Parents and caregivers are generally better educated and resourced. This is why our studies aim to understand mechanisms, not specificities. The objection regarding the study's age overlooks the universal factors that promote human development across different times and places. While practices and values may vary, the fundamental principles of fostering positive child development remain constant. This has been understood by countless philosophers, theologians, and individuals with common sense throughout history.

Human development is a life-cycle process. An approach with the greatest prospects for success searches for common *mechanisms* across successful programs. Such mechanisms are transportable and can guide policy everywhere. Child development has a common dynamic across eras, cultures, and ethnic groups (Ertem et al., 2018; Fernald et al., 2017). One should

ask how to bolster these mechanisms—not to recommend a specific policy off the shelf, but to have a template for assessing and developing successful policies appropriate for targeted populations. It is unlikely that any successful program in one context can be transported without modification to another context. Many early childhood investments and policies today are guided by an "up or down" philosophy. The fixation on treatment effects and decision theory to select among existing programs implements this "up or down" approach. A better approach extracts general lessons and builds tools that model the impacts of context and allow analysts to account for it. Long-run studies are central to this approach, as are recently developed methods that can reliably forecast long-run future skills and life outcomes for newly collected samples of program participants (e.g., García et al., 2020).

The analysis of Duncan and Magnuson (2013) is cited frequently, and it is typical of the "up or down" mindset. They compare programs that differ greatly regarding target populations, interventions administered, and measures used to gauge success. They compare the incomparable without rigorous methodological standardizations or any understanding of the mechanisms producing successful child development.

**Figure 1.** Average Impact of Early Childcare Programs at End of Treatment




**Source:** Duncan and Magnuson (2013). **Note (directly from source):** This figure shows the distribution of 84 program-average treatment effect sizes for cognitive and achievement outcomes, measured at the end of each program's treatment period by the calendar year the program began. Reflecting their approximate contribution to weighted results, "bubble" sizes are proportional to the inverse of the squared standard error of the estimated program impact. There is a weighted regression line of the effect size by calendar year.

Figure 1, taken from their work, illustrates their approach. It compares outcomes at the immediate end of *each intervention* for an assortment of programs of unspecified duration and lifetime impact by the date the program originated without attempting to standardize the populations studied. They do not consider lifetime consequences as studies of Perry do. They do not assess the quality of the programs studied, the quality of the reported estimates in terms of methodology, replicability, comparability in measures used, the quality of the investigators reporting results, the quality of the programs themselves, or the autonomy of the evaluators from the originators of the programs. It is unwise for policymakers to conclude from this figure and others like it that early childhood programs today are less effective than they were or that the Perry studies are irrelevant to today's context. A narrow "up or down" approach to program evaluation and scale-up will not lead to an effective early childhood education policy.

The "best practice" approach does not search for common developmental mechanisms and considers programs as stand-alone. In this view, policy evaluation searches for the "best" program to implement. The *What Works and What Does Not?* archive[3] and What Works Clearinghouse[4] are founded on this principle. "Meta-analysis" is built on this approach. Treatment effects from diverse programs, assessed using diverse measures on diverse populations, are "synthesized," forcing comparisons of incomparables. In this approach, statistics replaces science.

Research on Perry provides a basis for creating general principles underlying child development. The process of developing a deeper understanding of these principles continues today. A recent example is Dehaene (2020). Understanding these mechanisms should be front and center in the science of child development and the design of effective social policy, not the quest for the best policy among those already tried.

Many policymakers have cited the results from Perry Preschool to advocate for universal preschool programs (e.g., Nir, 2017). All our work has demonstrated the greatest returns to these investments is for disadvantaged[5] populations. By disregarding the characteristics of these programs, policy will fall short. Publicly funded early childhood and preschool programs provide

[3] Washington State Institute for Public Policy, 2015
[4] https://ies.ed.gov/ncee/wwc/
[5] Eligibility criteria in the Perry study was that children had a low socioeconomic status (based on an index of parental education and employment, and rooms per household member), a Stanford-Binet IQ score one standard deviation below the population average, and live in the Perry school district. Eligibility criteria in the Abecedarian study rated children on a High Risk Index determined by maternal and paternal education levels, family income, father's presence, and other indictors of family status and functioning.

different classroom experiences for children and teachers given high economic and racial segregation. Public mandates shape differences in how preschoolers spend their time and the kind of instruction they receive (Stockstill, 2023). Therefore, effective policy to promote mobility for under-resourced households should look to research that isolates the common mechanisms across successful programs and maintain the key components, for example direct engagement and empowerment of parents/caregivers (Derman-Sparks et, al., 2016a), (Derman-Sparks et al., 2016b).

## 3. Lessons Learned

When making investments in early childhood programs, policymakers look to "evidence-based curriculum" and "validated assessments" to ensure programs meet child development outcomes and reduce inequality in education. Unfortunately, these efforts can be misguided due to a blind reliance on achievement test scores and rigid, burdensome fidelity requirements. Therefore, creating scientifically grounded syntheses and exploring the differences and similarities among the various studies in a systematic way is critical for selecting investments in social programs. A research agenda at the University of Chicago's Center for the Economics of Human Development (CEHD) proposes models of skill development that synthesize data from multiple programs, and empirically grounded analytical models that offer the most robust approach for policy analysis.

By integrating and unifying the human capital and early life development literatures and developing dynamic models of the acquisition of skills (including health), scholars can make empirically grounded policy recommendations for programs that efficiently shape skills that promote flourishing lives in society. It is essential that further research investigates how to measure skills and how families and environments moderate or mediate the effects of interventions. These have important practical consequences for a successful early childhood education policy at scale.

### 3.1. The Economics of Early Childhood Development

Cunha and Heckman (2007, 2008, 2010) introduced the "technology of skill formation," which describes the process through which skills and abilities are developed over an individual's life course. Dynamic complementarity explains why skills acquired at one stage of development increase the productivity of investments at subsequent stages, creating a compounding effect on

human capital accumulation. Since early investments in human capital can enhance the effectiveness of later investments, they argued that interventions in the early childhood period can have high rates of return and be important mechanisms for reducing inequality and developing human potential. Analysis of Perry and later work at CEHD has demonstrated the various positive outcomes from effective high-quality early childhood programs, including improved educational attainment, higher earnings, reduced crime rates, and better health. These benefits promote social mobility, reduce social costs, and are more effective and less costly than later life remediation. Benefits extend beyond the children enrolled in the program, improving outcomes of younger siblings, parents, and offspring and generating returns to society.

Current research at CEHD examines dynamic analytical frameworks (common across studies) to analyze longitudinal data on the growth of comparable measures of skills and outcomes using data from major early childhood interventions targeted to disadvantaged children, controlling for family and environmental conditions. Mediation analyses of a number of studies investigate the channels of influence that produce outcomes. The longitudinal data from Perry and similar studies allow scholars to estimate the causal effects of education, parenting practices, and family investments and environments on behaviors of persons over their life cycles and across generations.

One instance of this approach is Garcia and Heckman (2024), which compares the Perry and the Carolina Abecedarian studies to home-visiting programs. Home-visiting programs scaffold the parent or caregiver to engage positively and use household materials to offer cognitive stimulation to the child, similar to that in Perry, as described in Derman-Sparks (2016a,b). When valid comparisons of child and adult outcomes are made at comparable ages, scholars can examine omnibus and focused home visiting programs to isolate and understand the importance of the home-visiting component in Perry. This crucial component of Perry—home-visiting—is effective in many different contexts worldwide (see Figure 3).

### *3.2. A Review of the Main Findings of Perry*

We present an overview of the key findings and a summary of the life cycles of Perry children and how they compare with findings in the wider literature. Table 1 presents a profile of lifecycle outcomes for Perry and the Carolina Abecedarian Project. It displays the average in the control group, the average treatment-control difference, as well as inference for this difference.

Gains are substantial across domains over the life cycle.

**Table 1.** Baseline Characteristics, Outcomes, and Fertility: Perry and Abecedarian

| | Perry | | | Abecedarian | | |
|---|---|---|---|---|---|---|
| | Control Mean | Mean Difference (MD) | MD p-value | Control Mean | Mean Difference (MD) | MD p-value |
| ***Panel a.*** **Baseline** | | | | | | |
| IQ (Perry) or Mother's IQ (Abecedarian) | 78.54 | 1.03 | 0.387 | 83.49 | 1.83 | 0.399 |
| Socioeconomic Index | 8.62 | 0.17 | 0.530 | 21.82 | -1.93 | **0.089** |
| Mother Does not Work¶ | 0.69 | 0.22 | **0.002** | 0.39 | -0.22 | **0.010** |
| Mother's Year of Birth | 1959.97 | 0.03 | 0.950 | 1974.35 | -0.15 | 0.674 |
| ***Panel b.*** **Midlife Skills**† | | | | | | |
| Cognitive | 0.00 | 0.48 | **0.005** | 0.00 | 0.34 | **0.031** |
| Non-Cognitive | 0.00 | 0.50 | **0.011** | 0.00 | 0.47 | **0.031** |
| ***Panel c.*** **Midlife Education**‡ | | | | | | |
| High-School Graduate | 0.52 | 0.20 | **0.021** | 0.53 | 0.20 | **0.025** |
| College Graduate | 0.05 | 0.02 | 0.453 | 0.09 | 0.21 | **0.007** |
| ***Panel d.*** **Midlife Outcomes*** | | | | | | |
| Married | 0.25 | 0.09 | **0.082** | 0.42 | 0.01 | 0.486 |
| Labor Income (2021 USD) | 16,298.91 | 7,826.94 | **0.018** | 37,527.95 | 13,044.70 | **0.098** |
| Household Labor Income (2021 USD) | 25,121.43 | 13,243.21 | **0.007** | 37,247.62 | 14,632.67 | **0.071** |
| Accumulated Days (Perry) or Times (Abecedarian) in Jail or Prison | 1,326.71 | -380.83 | 0.237 | 0.14 | -0.12 | **0.027** |
| Never Arrested (Perry) or Accumulated Arrests (Abecedarian) | 0.46 | 0.18 | **0.039** | 0.61 | 0.26 | 0.151 |
| Physical Health | 0.00 | -0.02 | 0.553 | 0.00 | 0.28 | **0.096** |
| Mental Health | 0.00 | 0.31 | **0.072** | 0.00 | 0.20 | 0.111 |
| ***Panel f.*** **Sample Sizes** | | | | | | |
| Original Participants at Baseline | 65 | -7 | 57 | | | 2 |
| Original Participants at Midlife Follow-up | 50 | 2 | 45 | | | 6 |

**Source:** García and Heckman (2023). **Note:** Panels (a) and (e) present the control-group mean and treatment-control mean difference for the outcome in the label for the Perry Preschool (Perry) and Carolina Abecedarian (Abecedarian) projects. For each treatment-control mean difference (MD), we present the permutation *p*-value associated with the null hypothesis that such mean difference equals 0. We bold *p*-values when they are less than 0.10. Panels (b) to (d) are analogous in format to Panels (a) and (d). The null hypothesis in these latter panels is that the mean difference is less than or equal to 0. ¶The difference between treatment-group mothers in Abecedarian and Perry is that Abecedarian provided full-day childcare and Perry did not. ‡Based on identical variables of completed years of education for both Perry and Abecedarian. ∗For Perry, marriage is the fraction of years married between ages 20 and 40; labor income is the average earnings from labor income between ages 20 and 40; household labor income is the previous variable in addition to average spouse's labor income between ages 20 and 40 (if married); accumulated days in prison and never arrested are observed up to age 54. For Abecedarian, marriage is an indicator of whether an individual is married at age 45; labor income is measured at age 45; household income is the previous variable in addition to spouse's labor income at age 45 (if married); times in jail and accumulated arrests are measured at age 30. For Perry, physical health is a latent variable of measures describing prevalence and intensity of diabetes, stroke, heart disease, self-rated health, body-mass index, and waste-to-hip ratio at age 54. For Abecedarian, an analogous variable is constructed using information at age 34. For Perry, mental health is a latent variable of measures describing depression and anti-social behavior at age 54. For Abecedarian, an analogous variable is constructed using the information at age 45.

Methodological contributions at CEHD strengthened the rigor of Perry Preschool and related analyses. Innovations include accounting for small samples, correcting for compromised randomization, correcting for control contamination, and harmonizing measures to satisfy metric invariance. One frequent criticism of the study is that sample sizes are too small. This line of attack ignores the application of exact small-sample inference to the main Perry results.[6] Our studies use empirically validated dynamic models to forecast experimental outcomes outside of sample, conduct mediation and moderation analyses to control for endogeneity, disaggregated treatment effects by gender,[7] and examined intergenerational effects. Another frequent criticism of Perry is regarding fadeout of IQ.[8] However, the treatment group at the latest follow-up demonstrated higher cognitive and noncognitive skills than the control group (Garcia et al., 2023).

### *3.3. Additional Detail on the Perry Studies*

Follow-ups with the original Perry participants were conducted when they were 19, 27, 40, and 54 and for Abecedarian at ages 21, 30, 34 (health outcomes), and 45. Multiple studies analyze the data from these follow-ups, listed in Table 2.[9]

[6] Heckman and Karapakula (2021); Heckman et al. (2010, 2024).
[7] Elango et al. (2016); García et al. (2018).
[8] See Bailey et al. (2020); Burchinal et al. (2024); Duncan et al. (2023).
[9] Studies are evaluations based on the computation of treatment effects with methodologies developed for analyzing small samples; cost-benefit analyses, including those developing strategies to forecast long-term treatment effects when follow-ups are unavailable and those based on annual observations up to late midlife (age 54) and intergenerational benefits; and mediation analyses aiming to understand which early life skills are building blocks of treatment effects on outcomes during adulthood.

**Table 2. Studies of the Perry Preschool by the Center for the Economics of Human Development (University of Chicago) and Collaborators**

| *Study* | Authors | Outlet | Year of Publication |
|---|---|---|---|
| *The rate of return to the HighScope Perry Preschool Program* | Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz | Journal of Public Economics | 2010 |
| *Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program* | Heckman, James, Seong Hyeok Moon, Rodrigo Pinto, Peter Savelyev, and Adam Yavitz | Quantitative Economics | 2010 |
| *Understanding the mechanisms through which an influential early childhood program boosted adult outcomes* | Heckman, James, Rodrigo Pinto, and Peter Savelyev | American Economic Review | 2013 |
| *The effects of two influential early childhood interventions on health and healthy behaviour.* | Conti, Gabriella, James J. Heckman, and Rodrigo Pinto | Economic Journal | 2016 |
| *Using a satisficing model of experimenter decision-making to guide finite-sample inference for compromised experiments* | Heckman, James J., and Ganesh Karapakula | Econometrics Journal | 2021 |
| *The lasting effects of early-childhood education on promoting the skills and social mobility of disadvantaged African Americans and their children* | García, Jorge Luis, James J. Heckman, and Victor Ronda | Journal of Political Economy | 2023 |
| *Parenting promotes social mobility within and across generations* | García, Jorge Luis, and James J. Heckman | Annual Review of Economics | 2023 |
| *Dealing with imperfect randomization: Inference for the highscope perry preschool program* | Heckman, James, Rodrigo Pinto, and Azeem M. Shaikh | Journal of Econometrics | 2024 |
| *The dynastic benefits of early-childhood education: participant benefits and family spillovers* | Bennhoff, Frederik H., Jorge Luis García, and Duncan Ermini Leaf | Journal of Human Capital | 2024 |

García et al. (2023) analyze the latest follow-up data from participants of Perry in midlife (age 54). They combine observations from retrospective surveys observed throughout adulthood and administrative data to construct longitudinal life-cycle education, labor income, and crime outcomes up to late midlife, shown in Table 1. Both follow-up studies asked original participants about their children, enabling scholars to analyze intergenerational outcomes and to construct similar outcomes for participants and their children across the two programs.

Current collaborations with HighScope are active as these rich data have not been exhausted; daily classroom lesson plans, notes from the home visits, qualitative interview data with elementary school teachers can provide a more vibrant picture of how early childhood programs affect children, families and communities.

#### *3.3.1. Fadeout*

Previous research claims that the impacts of early childhood education disappears (fades out) shortly after the interventions. These studies are primarily based on short-run follow-ups.

Both Perry and Abecedarian programs have a long-run impact on cognition, as measured by well-established cognitive tests (Raven & Stroop). These long-lasting impacts on cognition for both programs contradict the frequently repeated refrain about "fade-out" in the treatment effects on skills—specifically cognition. The evidence presented in Table 1 refutes the fade-out claim for comparable (across programs) cognitive and non-cognitive skill measures.

Panels c. and d. show that the long-term impact of the programs goes well beyond just enhancing cognitive and non-cognitive skills. We observe completed education at midlife. Both programs significantly increase the high-school graduation rate by 20 percentage points from a control-group base rate of around 50%. They also considerably increase earnings from labor income during adulthood and decrease criminal behavior. Perry decreases the likelihood of ever being arrested by 18 percentage points ($p$-value = 0.04) from a control-group rate of 46%. Abecedarian decreases the average number of times in jail or prison by 0.12 ($p$-value = 0.03) from a control-group average of 0.14.

We also have measures of physical and mental health. Perry improves mental health by 0.31 standard deviations ($p$-value = 0.07) from a control-group mean of 0. Abecedarian improves physical health by 0.28 standard deviations ($p$-value = 0.096) from a control-group mean of 0. The outcomes in Table 1 represent broad categories. They show sustained program impacts generating marriage, labor, and law-abiding stability across the life cycle.

### *3.4. Mechanisms*

Our work expands the understanding of mechanisms producing treatment effects by investigating how home environments were shaped by Perry and Abecedarian and how this shaping impacted the development of skills fostered by the programs.

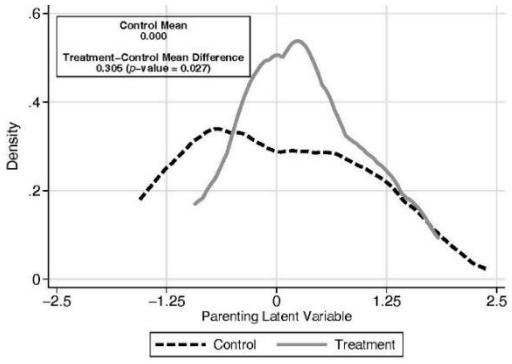
### *3.4.1. Parenting*

For Perry, we dug deeper into the original data archive and recovered the Parental Attitude Research Instrument.[10] These data had never previously been analyzed. PARI was collected from the mothers of the original participants by staff conducting home visits when the children were between three and five years old. It measures the quality of parental investment and parenting.

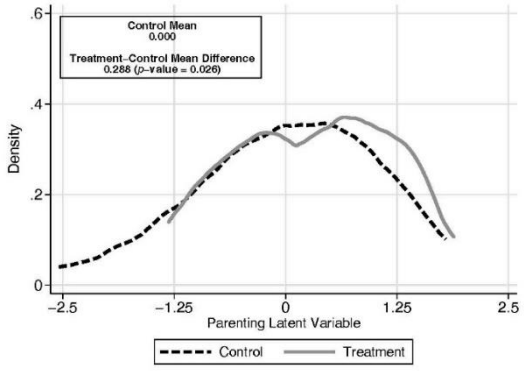
For each program, we construct a measure based on items of PARI or Home Observation Measurement of the Environment (HOME), in the case of Abecedarian. In Figure 2, panel (a) and (b) show the distributions of our measures by treatment status for each program. Perry and Abecedarian enhance parenting or parental investment by an average of 0.3 (*p*-value = 0.027) and 0.3 (*p*-value = 0.026), respectively. These findings bolster our interpretation of Perry and Abecedarian as policies targeting disadvantaged families. They support Weikart's initial observations made some 60 years ago. Perry and Abecedarian enhance home environments and improve the interactions of child participants with their caretakers, which last long after the program ends. When examined with research on home visiting-only programs, it is clear that this is a critical piece of how outcomes in the Perry and Abecedarian programs were achieved.

**Figure 2.** Parenting Received by the Original Participants of the Perry and Abecedarian

**(a)** Parenting and Distribution, Perry




**(b)** Parenting and Distribution, Abecedarian



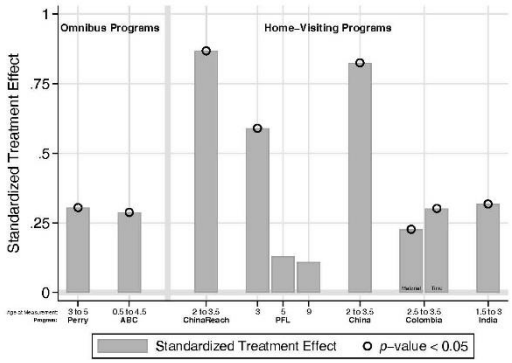

**Source:** Garcia and Heckman (2023). **Note:** Panel (a) shows the probability density function of a latent variable describing the parental investment (parenting) received by the original participants of the Perry Preschool Project (Perry) by treatment status. We also display the control-group mean and the treatment-control mean difference in the index and the permutation *p*-value for this difference. The null hypothesis for the difference is that it is less than or equal to 0. Panel (b) is analogous in format to Panel (a) for the parental investment received by the original participants of the Carolina Abecedarian Project.

[10] PARI; Loewenstein (1973).

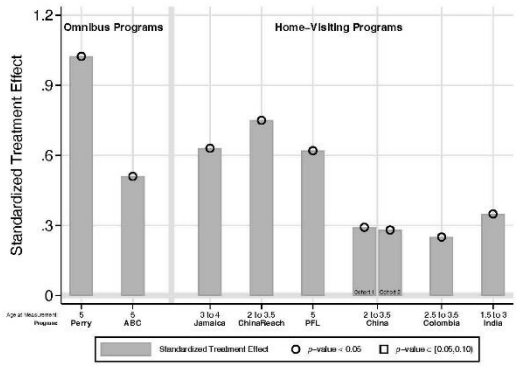
Additional work by García and Heckman (2023) investigates the extent to which the increase of positive parent-child interactions generated by the program are building blocks for impact on skills later in midlife. Remarkably, these measures are based on data collected fifty years apart. For the two programs, a one-standard-deviation increase in our post-treatment measure of parenting, collected when the original participants were at most five years old, is associated with an increase of half a standard deviation in the average of midlife skills. The improvement in parenting generated by the two programs is a clear mechanism for their long-term success.

**Figure 3.** Impacts on the Home Environment and Very Early-Life Skills, Omnibus and Home-Visiting Programs

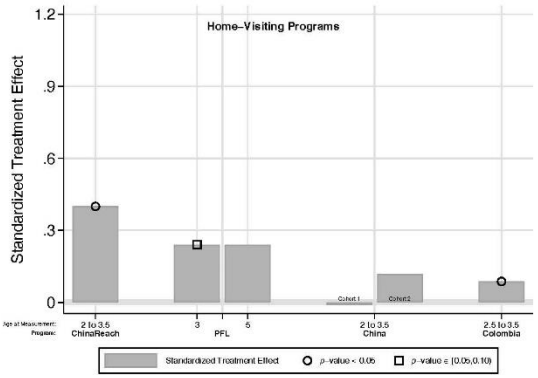
**(a)** Home Environment




**(b)** Cognitive Skills




**(c)** Non-Cognitive Skills




**Source:** Garcia and Heckman (2023). **Note:** Panel (a) displays program impacts on home environment measures. Both measures are standardized by subtracting the control-group mean and dividing by the control-group standard deviation. The measures for the rest of the programs are standardized similarly. We report treatment effects for all programs except for PFL (estimates of treatment-control mean differences). The impacts reported for PFL are effect sizes. We mark impacts when the *p*-value associated with the null hypothesis that they are less than or equal to 0 is less than 0.05. The measures are in Garcia and Heckman (2023). Panels (b) and (c) are analogous in format to Panel (a) for measures of cognitive and non-cognitive skills. *Cohorts* in China: For cognitive and non-cognitive skills, we display from Sylvia et al. (2021) who report separate results for two cohorts within their sample, while for the home environment they report results in pooling the two cohorts.

### 3.4.2. *Intergenerational Outcomes*

The long-lasting improvements in the original participants' skills, marriage stability, earnings, criminal behavior, and health of both Perry and Abecedarian are solid during their child-rearing years. These impacts translate into better family environments for their children. García et al. (2023) document that the more skilled and educated original treatment-group members of Perry who become parents are less likely to have children out of wedlock and cohabitate with new partners during the childrearing years relative to control-group counterparts. They are more likely to stay married while their children grow up and read more to them. Larger average labor incomes for the original treatment group translate into more resourced home environments. Lower incarceration rates translate into greater parental presence at home, especially for treatment-group fathers.

The Perry data analyzed by García et al. (2023) and newly available data on the children of the original participants of Abecedarian show no economically or statistically significant average differences across experimental groups on fertility variables, including whether participants have children or not, age at onset of fertility, number of children, and whether participants have more than five children. Overall, the data indicate that the program has minimal impacts on childbearing, making differences in experimentally induced fertility a secondary consideration.

Table 3 presents the expected outcomes and corresponding treatment effects. There are important intergenerational spillovers. Results vary by gender in the first generation. The same is true in the second generation. Both programs have economically and statistically sizable intergenerational impacts on employment, health status, and marriage rates of the average male child of the original treatment-group participants relative to the average child of their control-group counterparts. Perry also has a sizable intergenerational impact on reducing crime. Both programs sizably increase high-school graduation for girls. Abecedarian also increases college female graduation rates.

**Table 3.** Summary of Intergenerational Outcomes: Children of the Original Participants of Perry and Abecedarian

| | Male Children | | | Female Children | | |
|---|---|---|---|---|---|---|
| | Control Mean | Mean Difference (MD) | MD p-value | Control Mean | Mean Difference (MD) | MD p-value |
| ***Panel a.*** **Perry** | | | | | | |
| High School Graduate (Age 18 or older) | 0.67 | -0.01 | 0.582 | 0.74 | 0.13 | **0.026** |
| College Graduate (Age 23 or older) | 0.04 | 0.08 | **0.063** | 0.31 | -0.09 | 0.846 |
| Employed (Age 23 or older) | 0.48 | 0.19 | **0.040** | 0.41 | 0.09 | 0.218 |
| Never Arrested (Age 18 or older) | 0.37 | 0.14 | **0.089** | 0.78 | 0.06 | 0.210 |
| In Good Health (Age 18 or older) | 0.82 | 0.12 | **0.006** | 0.85 | 0.10 | **0.030** |
| Not a Parent (Ages 14 to 22) | 1.00 | 0.00 | 1.000 | 0.83 | 0.12 | 0.234 |
| Never Divorced (Age 23 or older) | 0.93 | 0.07 | **0.028** | 0.86 | 0.11 | **0.016** |
| ***Panel b.*** **Abecedarian** | | | | | | |
| High School Graduate (Age 18 or older) | 0.66 | -0.06 | 0.718 | 0.28 | 0.18 | **0.067** |
| College Graduate (Age 23 or older) | 0.55 | -0.08 | 0.683 | 0.18 | 0.25 | **0.068** |
| Not Idle (Age 15 or older)[†] | 0.91 | 0.06 | **0.083** | 0.98 | 0.00 | 0.572 |
| In Good Health (Age 18 or older) | 0.83 | 0.18 | **0.000** | 0.88 | 0.10 | 0.133 |
| Not a Parent (Ages 14 to 22) | 0.63 | 0.17 | **0.069** | 0.94 | -0.01 | 0.584 |

**Source:** García and Heckman (2023). **Note:** Panel (a) presents the control-group mean and treatment-control mean difference (MD) for the intergenerational outcome in the label for the Perry Preschool Project (Perry). Intergenerational outcomes are for the average child. We construct them by averaging within original program participants up to their five eldest children. We present the permutation *p*-value associated with the null hypothesis that the mean difference is less than or equal to 0 for each mean difference. We bold *p*-values when they are less than 0.10. Panel (b) is analogous in format to Panel (a) for the Carolina Abecedarian Project.

[†] Not Idle: enrolled in school or working.

### *3.4.3. Cost-Benefit Analyses*

The evidence shows that Perry and Abecedarian are effective at boosting the life-cycle outcomes of their original participants. Do the benefits of the programs outweigh their costs? Bennhoff et al. (2024) for Perry and García et al. (2020) for Abecedarian, conduct cost-benefit analyses accounting for the full social opportunity cost of public expenditure.

The cost per participant of Perry is $23,478 in 2021 dollars. The cost per participant of Abecedarian is $105,530 dollars. García et al. (2020) and Bennhoff et al. (2024) monetize the average treatment effects of these programs over the life cycle. These studies report that the program has statistically significant average net social benefit per participant (average total benefits less cost per participant) of $175,548 and $672,359, respectively. The estimates of the net social benefits account for the welfare cost of distorting taxation required to fund public programs. The corresponding statistically significant benefit-cost ratios are 6.0 and 5.2. The source studies show that the reported estimates are robust to extensive robustness checks of the assumptions underlying their estimation.

Bennhoff et al. (2024) report an additional intergenerational contribution to the net social benefits of $43,000 per average male child of original participants of Perry and $14,000 per average female child. They also report an additional intragenerational contribution to the net social benefits of $68,000 per average male sibling of the original participants and $13,000 per average female sibling. Although the latter estimates are imprecise, these results indicate that the program generates additional benefits without additional costs that cover the average cost per original participants.

## *4. In Summary*

The studies of the Perry Preschool Program show it was an effective program. Data from the Perry Preschool Study are a wellspring for which new evidence can still emerge, as previously unanalyzed data are brought to light and as fresh waves of data are collected. The benefits of this intervention are intergenerational and substantial in magnitude. Beyond IQ and education, it promoted social mobility and overall quality of life. A main mechanism of this preschool program was the impact on parenting, which creates a nourishing and supportive home environment conducive to learning. García et al. (2023) show that the single component of home visiting produces similar effects at a lower cost than the center-based preschool program. The role of parental engagement in early childhood programs is one major lesson from Perry that should stimulate further research and the design of future intervention studies.

## References


Bailey, D. H., Duncan, G. J., Cunha, F. B., Foorman, R., & Yeager, D. S. (2020). Persistence and fade-out of educational-intervention effects: Mechanisms and potential solutions. *Psychological Science in the Public Interest 21*(2), 55–97.

Beatty, B. (2013). Playing with numbers: Constance Kamii and reinventing arithmetic in early childhood education. In *The Hidden History of Early Childhood Education* (pp. 238-262). Routledge.

Bennhoff, F. H., García, J. L., & Leaf, D. E. (2024). The dynastic benefits of early-childhood education: Participant benefits and family spillovers. *Journal of Human Capital 18*(1), 44–73.

Burchinal, M., Whitaker, A., Jenkins, J., Bailey, D., Watts, T., Duncan, G., & Hart, E. (2024). Unsettled science on longer-run effects of early education. *Science 384*(6695), 506–508.

Conti, G., Heckman, J. J., & Pinto, R. (2016). The effects of two influential early childhood interventions on health and healthy behaviour. *The Economic Journal*, *126*(596), F28-F65.

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

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

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

Dehaene, S. (2020). *How we learn: The new science of education and the brain*. Penguin.

Derman-Sparks, L. & Moore, E. K. (2016). Our proud heritage: Two teachers look back—the Ypsilanti Perry Preschool, part I. *NAEYC*. https://www.naeyc.org/resources/pubs/yc/sep2016/ypsilanti-perry-part-1.

Derman-Sparks, L., & Moore, E. K. (2016). Our proud heritage: Lessons for today the Ypsilanti Perry Preschool, part II. *NAEYC*. https://www.naeyc.org/resources/pubs/yc/nov2016/ypsilanti-perry-part-2.

Derman-Sparks, L. (2016). What I learned from the Ypsilanti Perry Preschool Project: A teacher's reflections. *Journal of Pedagogy*, *7*(1), 93-106.

Duncan, G., Kalil, A., Mogstad, M., & Rege, M. (2023). Investing in early childhood development in preschool and at home. In *Handbook of the Economics of Education*, Volume 6, pp. 1–91. Elsevier.

Duncan, G. J. & Magnuson, K. (2013). Investing in preschool programs. *Journal of Economic Perspectives*, *27*(2), 109–132.

Elango, S., García, J. L., Heckman, J. J., & Hojman, A. (2016). Early childhood education. In *Economics of Means-Tested Transfer Programs in the United States*, Volume II. University of Chicago Press.

Ertem, I. O., Krishnamurthy, V., Mulaudzi, M. C., Sguassero, Y., Balta, H., Gulumser, O., Bilik, B., Srinivasan, R., Johnson, B., Gan, G., Calvocoressi, L., Shabanova, V., & Forsyth B. W. C. (2018). Similarities and differences in child development from birth to age 3 years by sex and across four countries: A cross-sectional, observational study. *The Lancet Global Health, 6*(3), e279–e291.

Fernald, L., Prado, E. L., Kariger, P., & Raiker, A. (2017). *A toolkit for measuring early childhood development in low and middle-income countries.* World Bank.

García, J. L. & Heckman, J. J. (2023). Parenting promotes social mobility within and across generations. *Annual Review of Economics, 15*, 349–388.

García, J. L., Heckman, J. J., Leaf, D. E., & Prados, M. J. (2020). Quantifying the life-cycle benefits of an influential early-childhood program. *Journal of Political Economy, 128*(7), 2502–2541.

García, J. L., Heckman, J. J., & Ziff, A. L. (2018). Gender differences in the benefits of an influential early childhood program. *European Economic Review, 109*, 9–22.

García, J. L., Heckman, J. J., & Ziff, A. L. (2019). Early childhood education and crime. *Infant Mental Health Journal*, 40(1), 141-151.

García, J. L., Heckman, J. J., & Ronda, V. (2023). The lasting effects of early-childhood education on promoting the skills and social mobility of disadvantaged African Americans and their children. *Journal of Political Economy*, *131*(6), 1477-1506.

Hechinger, F. M. (1986, April 22). About education: Preschool programs. *The New York Times*. https://www.nytimes.com/1986/04/22/science/about-education-preschool-programs.html.

Heckman, J. J. & Karapakula, G. (2021). Using a satisficing model of experimenter decision-making to guide finite-sample inference for compromised experiments. *Econometrics Journal, 24*(2), C1–C39.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review, 103*(6), 2052-2086.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics, 94*(1-2), 114-128.

Heckman, J. J., Pinto, R., Moon, S. H., Yavitz, A., & Savelyev, P. (2010). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool program. *Quantitative Economics, 1*(1), 1–46.

Heckman, J. J., Pinto, R., & Shaikh, A. M. (2024). Dealing with imperfect randomization: Inference for the HighScope Perry Preschool program. *Journal of Econometrics*.

Inhelder, B., & Piaget, J. (2013). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures*, 84. Routledge.

Leventhal, T., Martin, A., & Brooks-Gunn, J. (2004). The EC-HOME across five national data sets in the 3rd to 5th year of life. *Parenting, 4*(2-3), 161-188.

Linver, M. R., Martin, A., & Brooks-Gunn, J. (2004). Measuring infants' home environment: The IT-HOME for infants between birth and 12 months in four national data sets. *Parenting, 4*(2-3), 115-137.

Loewenstein, S. (1973). Conclusion of a study of PARI as a research tool. *Child Welfare, 52*(6), 400-402.

Nir, S. M. (2017, April 24). De Blasio's Pre-K Expansion Means Smaller Paychecks for Some Teachers. *The New York Times*. https://www.nytimes.com/2017/04/24/nyregion/de-blasio-pre-k-expansion.html

Ramey, C. T., & Campbell, F. A. (1984). Preventive education for high-risk children: Cognitive consequences of the Carolina Abecedarian project. *American Journal of Mental Deficiency, 88*(5), 515-523.

Sparling, J., Lewis, I., & Ramey, C. T. (2001). *LearningGames for threes and fours*. Teaching Strategies.

Schweinhart, L. L., Weikart, D. P., & Larner, M. B. (1986). Consequences of three preschool curriculum models through age 15. *Early Childhood Research Quarterly, 1*(1), 15-45.

Schweinhart, L. J., & Weikart, D. P. (1997). *Lasting differences: The HighScope preschool curriculum comparison study through age 23.* Monographs of the HighScope Educational Research Foundation, No. 12. HighScope Press

Stockstill, C. (2023). *False starts: The segregated lives of preschoolers*. NYU Press.

U.S. Department of Education, Office for Civil Rights (2023, November). *2020-21 Civil Rights Data Collection: Student Discipline and School Climate in U.S. Public Schools*.

Vygotsky, L. (1980). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Wakabayashi, T., Claxton, J., Melton-Grace, E., Delcamp, M., Schweinhart, L. J., & Gardner, S. (2024) What happened during home visits? A closer look at the family engagement component of HighScope's Perry preschool project. Unpublished manuscript.

Washington State Institute for Public Policy (2015). *What works and what does not? Benefit-cost findings from WSIPP*. Technical report.

Weikart, D. P. (2004). *How HighScope grew: A memoir*. HighScope Press.

Weikart, D. P., Karnii, C. K., Radin, N. L., Ernerich, P. H. (1964). *Perry Preschool Project progress report.* Ypsilanti Public Schools. Submitted to the U.S. Department of Health, Education, and Welfare, Office of Education. ERIC. [ED001894].

Weikart, D. P. (1966). Preschool programs: Preliminary findings. *Journal of Special Education, 1*(2), 163–181.

Weikart, D. P., Deloria, D. J., Lawser, S. A., Wiegerink, R. (1971). *Longitudinal results of the Ypsilanti Perry Preschool Project. Final Report.* HighScope Educational Research Foundation, Ypsilanti Public Schools. Office of Education (DHEW), Bureau of Research. Contract No. OFC-4-10. Document No. 40PP-P-0328. [ERIC 000638]