

The Chicago Child-Parent Center early-childhood program: A critical appraisal

Geert Driessen

Institute of Applied Social Sciences, Radboud University, Nijmegen, The Netherlands (Retired)

Correspondence to: Geert Driessen, Email: driessenresearch@gmail.com

Abstract: Early Childhood Education and Care (ECEC) interventions are often the core of policies to tackle educational disadvantage. They target socio-economically disadvantaged children and aim to improve their educational and life chances by providing compensatory stimulation activities in education institutions and at home. Many intervention programs have been developed and implemented since the 1960s. Three renowned so-called model programs are Perry Preschool, Abecedarian, and the Chicago Child-Parent Center (CPC). Though they date from the 1960s, 1970s, and 1980s, respectively, they are still often cited today as evidence that such programs are highly effective and that their effects can be generalized to later ECEC programs. This article focuses on the CPC and – if possible – makes comparisons with Perry and Abecedarian. The main question is whether this program really leads to positive effects, and if it does, whether these effects can be generalized to other, similar programs today. To find an answer to this question, a critical literature study was conducted, using the snowball method. On the basis of the hundreds of studies examined, it was concluded that, indeed, CPC yielded many positive effects in the domains of education, work, well-being, health, and crime. At the same time, several limitations were found, which seriously restricts its generalization.

Keywords: Early Childhood Education and Care, Child-Parent Center, Intervention program, Impact evaluation, Effectiveness, Educational disadvantages

1. Introduction

1.1 Equal opportunities

The realization of equal opportunities has been at the top of the political agendas since the 1960s. However, the extent to which this has resulted in corresponding concrete measures in practice depends largely on the political wind blowing at the time (Giudici et al., 2024; Lu et al., 2024). A critical part of this endeavor concerns education. Life chances start with children's educational opportunities and are largely determined by them (Aiston & Walraven,

2024; National Academies of Sciences, Engineering, and Medicine, 2023). Two factors are central to this: the inborn capacities of a child, and the conditions a child grows up under (Arpawong et al., 2023; Marks & O'Connell, 2024; Tan, 2024). To a large degree, both are topics of separate policies and practices (Cahill, 2021; Lenkeit et al., 2022).

Capacities are often measured in terms of IQ. In the past, children with a low IQ (and/or specific physical, behavioral and emotional qualities) were often referred to special education institutions; nowadays, they are more and more included in regular classes and schools under the banner of inclusive education (Mann et al., 2024; Shevlin & Banks, 2021). There they receive extra support in order to try and

Received: Dec.25, 2024; Revised: Jan.9, 2025; Accepted: Feb.17, 2025; Published: Feb.25, 2025

Copyright ©2025 Geert Driessen.

DOI: <https://doi.org/10.55976/rppe.32025134339-51>

This is an open-access article distributed under a CC BY license (Creative Commons Attribution 4.0 International License)

<https://creativecommons.org/licenses/by/4.0/>

create the same chances as children without such inborn detrimental characteristics.

Children growing up in disadvantaged families and environments are often lacking favorable conditions or support in their home environment, or are faced with external obstacles which hinder the full realization of their potential. Disadvantage generally is indicated in terms of low parental level of education, low parental professional level and low family income (Liu et al., 2022; Long & Renbarger, 2023), often in combination with immigrant and minority background. The attention to this type of disadvantage has resulted in numerous local and national initiatives, policies, and interventions (Driessen, 2022; Eurydice, 2020; Farquharson et al., 2022). The most common is providing schools that care for many disadvantaged children with extra subsidies to tackle educational delays (OECD, 2021; Shero & Hart, 2022). Such subsidies can then (often freely) be spent on a wide range of activities, for instance, individual help or class size reduction.

This article focuses on this latter factor, the socio-ethnic disadvantage. Despite the investment of billions of dollars to create equal opportunities in terms of social and ethnic backgrounds, the results are simply disappointing (Hanushek et al., 2019). In fact, the achievement gap between disadvantaged children and their more privileged fellow-students (or: the poor and the rich) has been on the rise for several years now (Betthäuser et al., 2023; Reardon, 2012). Spending huge amounts of money and developing and implementing a wide range of interventions obviously is not enough or not a simple and straightforward approach (Carter & Merry, 2021; Gorard et al., 2020). But the question as to what else to do is not easy to answer (Driessen, 2022; Maisuria & Lally, 2024; The Sutton Trust, 2024a).

1.2 Early Childhood Education and Care

Additional funding for schools with disadvantaged children can be seen as a general provision. A more specific measure is Early Childhood Education and Care (ECEC; also: Pre- and Early Childhood Education). ECEC comes in different appearances, especially in terms of its target groups. The European Commission (2014), for instance, applies a very broad definition of at-risk children, which includes both children from families belonging to a socially disadvantaged group (e.g., poverty, social or migrant background), and children with special educational needs (SEN; e.g., disabilities, mental health and behavioral problems). So does the OECD (2006), which states that children with physical, intellectual or sensory disabilities or from socio-economically disadvantaged environments are entitled to inclusive universal programs in the early educational stage. From this perspective, ECEC can be seen as universal or regular ECEC, which is intended for all children, regardless of their capacities or environment (Schmutz, 2024). However, ECEC is often part of a national

socio-ethnic disadvantage policy specifically focusing on children with sufficient abilities, but growing up under unfavorable conditions (The Sutton Trust, 2024b). Indeed, this type of ECEC is probably the most commonly practiced specific educational policy intervention for disadvantaged children today (Driessen, 2022). There are several reasons for this. The most important probably is that interventions at a later stage have led to disappointing results. It is felt that it is better to prevent delays at an early age than try to combat them at a later age when the children are already lagging behind considerably (Heckman Equation, 2013). Therefore, ECEC activities are increasingly being viewed as a means to better prepare both children and parents from disadvantaged socio-ethnic backgrounds for the formal school trajectory. In addition, it is believed that activities in the preschool period should be closely attuned to those in the early elementary school period.

ECEC most often aims at children from around 2 to 6 years of age, but in some countries from a few weeks after birth till their entry into elementary school and even up to the age of 9 years. Usually, a distinction between family-based and center-based programs is made. Family-based programs primarily aim at parents and, through them, their children. An important goal is to teach the parents the skills needed to stimulate the cognitive, linguistic, and motor development of their children. But parents may also receive actual support with respect to child-rearing and the social and emotional functioning of their children. Center-based programs are typically conducted in preschools and in the kindergarten groups of elementary schools. The emphasis can differ, but is mostly on stimulating cognitive development, and then especially language development. There are also integrated programs, focusing on various developmental domains at the same time. In addition to (separate) family-based programs and center-based programs, there are also combinations of the two, thus focusing both on the children in the institutions and their parents at home.

Over the years, thousands of evaluative studies and hundreds of review studies and meta-analyses have been conducted. The results can be summarized as follows: The effects of ECEC programs are generally small at best, and habitually disappear ("fade-out") once the program has stopped. In a few cases, small effects reappear after some years ("sleeper effect"). In addition, more recent studies show smaller effects than earlier studies. However, it is not easy to interpret these results correctly. A critical reading of the reports and articles reveals many severe methodological flaws, and that cherry-picking and neglecting the many null and even negative effects is quite common. And there is another problem. ECEC is a broad and complex field with many differences. The target groups can differ, as can the children's ages, the content of the intervention, the specific topics addressed, the duration and intensity of the program, the quality of the program and the teachers (education, experience), the scientific support provided by the developers, the time period in which the program is

offered, available alternatives, financial costs and support, and parental background (education, work, income, family composition, neighborhood) (Phillips et al., 2017). The decisive factor for success appears to be the high quality of the ECEC program, and the continuation of high quality in subsequent educational stages (National Academies of Sciences, Engineering, and Medicine, 2023).

In many synthesizing studies and statistical meta-analyses, such differences are ignored and lumped together, resulting in an amalgam of findings of low validity: it is not clear at all whether the findings can be generalized, and in case they can, to what populations and degrees (e.g., Bruhn & Emick, 2023; Burchinal et al., 2024; Cook & Wong, 2007; DeAngelis et al., 2020; Duncan et al., 2022; Eady et al., 2024; Fukkink et al., 2017; Gilliam & Zigler, 2001; Hahn & Barnett, 2023; Joo et al., 2020; Melhuish et al., 2015; Von Suchodoletz et al., 2023; Whitaker et al., 2023).

1.3 The Big Three

Probably the best-known ECEC programs are three early so-called model programs ("The Big Three"): Perry Preschool, Abecedarian, and Chicago Child-Parent Center (American Public Media, 2018; Besharov et al., 2011; Isaacs, 2008). The main reasons for their fame are that two of them constituted (true) experiments, that is, randomized controlled trials (RCTs – the gold standard of effect studies), and that the participating children have been followed for several decades. Many advocates of ECEC programs are of the opinion that these model programs have laid the foundations for later programs (Heckman, 2000). They claim that they are highly effective and have yielded many positive effects on both cognitive and non-cognitive outcomes. According to them, these programs may serve as an example ("a cornerstone" or "a prototype") and their effects can be generalized to later stimulation and compensatory programs (Garcia et al., 2020; Heckman, 2011; Heckman et al., 2010).

In a recent paper, Perry and Abecedarian publications were critically re-examined, especially regarding their validity and generalizability of the outcomes (Driessen, 2025). The results are often summarized as follows: Significant effects on educational achievement, employment, and other important life outcomes, such as health and criminal behavior occurred, and these effects persisted into adulthood. However, there were many limitations to the effect studies (which were regularly not mentioned or only partly mentioned): the programs were from the 1960s and 1970s, a completely different era; the experiments were carried out at just one site; the samples were very small and actually became smaller and smaller with each round of measurement due to the longitudinal nature of the studies; many of the evaluations were carried out by the developer/project leader; the target group was exceptional: low-IQ children ("mentally retarded," according to Perry's project leader) from African-American families (many single mothers) with no or little education and often no job, living

in high-poverty neighborhoods; parental participation constituted an essential element; the teachers were highly overqualified and received continuous coaching; the teacher-child ration was very low; the duration and intensity of participation was highly unusual (1,000 and 12,500 hours, respectively), and so were the costs (\$43,500 and \$120,000 per child, respectively). Taken together, these limitations raise serious doubts regarding the reliability and validity of the findings as reported in a series of publications and the possibility to scientifically soundly and realistically generalize them to other programs.

2. Research question and method

In the remainder of this article, the focus is on the last program of The Big Three, the Chicago Child-Parent Center (which is being mentioned less in the literature than the other two). The main question here is whether this program really leads to positive effects, and if it does, whether these effects can be generalized to other, similar programs today.

For this literature review, the so-called snowball method was used. The author brings over forty years of expertise in research on educational disadvantage and specializes in Early Childhood Education and Care. His (digital) bookcase holds many hundreds of reports and articles focusing on the effectiveness of pre- and early-school programs. This formed the starting point for an intensive literature search. The references mentioned in the relevant publications were also followed up. In addition, Google Scholar was searched for following the citations mentioned there. And lastly, Google was searched using keywords such as "Child-Parent Center"; "model program", "early childhood education", "effectiveness", and "criticism"; in addition, the name of the evaluation program's director, Dr. Arthur Reynolds, was added as a keyword. This resulted in an impressive number of relevant publications, mainly with Reynolds as the author or as a co-author (for an overview see <https://innovation.umn.edu/cls/selected-publications-and-reports/> and <https://clstudy.org/Publication.htm>). The ensuing publications were then critically examined, and notes were made. This eventually resulted in an overview of distinct outcomes and limitations. These will be discussed in the next section. But first, the main features of the Child-Parent Center will be described.

3. The Chicago Child-Parent Center

3.1 The program

The Chicago Child-Parent Center (CPC) program is a center-based early intervention program that provides comprehensive educational and family-support services to young, economically disadvantaged children living in

Chicago's poorest neighborhoods. Children begin preschool when they are 3 or 4 years old and are served in elementary school through third grade with an extended curriculum. Thus, it aims at 3- to 9-year-olds (P-3). (Duration of participation ranged from 0 to 6 years, with an average of 3.7 years; Phillips et al., 2007.) In addition to activities for the children, there is also a parent component, which includes parent resource rooms and a parent resource teacher who oversees parent activities at the center and within the community. More specifically, besides educational and intensive enrichment for the children, the program included a variety of family services, home visits, health and nutrition workshops and services, and also community outreach. Main components include small classes, activity-based learning (a mix of teacher-directed and child-initiated instructional approaches), collaborative leadership, family engagement and participation, and alignment of instruction across ages (preschool – kindergarten – elementary school). The program was founded in 1967 by Dr. Lorraine Sullivan and federally funded from Title I. Since then, the program has been modified and enhanced to reflect the changing socio-economic diversity. Moreover, it has also been reconceptualized as a school reform model. At present, 26 centers operate in the Chicago region, and 9 in Minnesota and Wisconsin. The overall goal is to promote educational success, leading to higher educational attainment and, eventually, to greater economic well-being and health.

3.2 The evaluation

The actual evaluation of CPC started in 1986, when the Chicago Longitudinal Study (CLS) was launched to investigate the effects of CPC participation on kindergarten graduates of that year. The researchers opted for a matched-group quasi-experimental design in which the performance of one complete cohort of CPC kindergartners was compared with that of children of the same age and living in similar socio-economic circumstances (home and neighborhood), who until that moment had not participated in CPC. Both groups were also well-matched on the basis of eligibility for the intervention and demographic factors such as gender, race, and ethnicity. There were 974 CPC children who had been enrolled in 1983 and finished kindergarten in 1986, i.e., the treatment or intervention group; 60 percent of these children attended full-day kindergarten, and the others attended half-day kindergarten. The matched comparison group actually consisted of two subgroups: (a) 389 children from randomly selected schools who also graduated from kindergarten in 1986, but who participated in other all-day kindergarten programs (which was a common intervention for at-risk children, i.e., "the treatment as usual"); (b) 176 children who started CPC services in kindergarten, but who had not attended the preschool component. These three comparison groups allow for many comparisons, such as any CPC vs. no CPC; preschool CPC vs. no preschool CPC; school-age CPC vs. no school-age CPC; extended CPC vs. nonextended CPC. In addition to differences in terms of

CPC participation, therefore, there also were differences in terms of duration and intensity of participation (e.g., full-day vs. half-day kindergarten). In total, the entire sample consisted of 1,539 children who all lived in low-income neighborhoods; 93 percent of them were African-American, 7 percent Latino; 76 percent grew up in a single-parent family; 54 percent of the mothers had not finished high school, and 63 percent were unemployed. Data from a variety of birth records, K-12 school records, surveys, interviews and administrative records from education, health and earnings have been collected since the start of the program. In the CLS, the children have been followed for many years. Data collection took place at ages 10, 15 to 18, 18 to 24, 26 to 28, 32 to 37, 37 to 39, and 40. (Sources for the above are, among others: Besharov et al., 2011; Eddy, 2012; Ou et al., 2020, 2021; Reynolds, 2000; Reynolds et al., 2001, 2010, 2016, 2021, 2023.)

3.3 Findings

The CPC research team has produced an abundance of scientific evaluation publications (more than 300), mostly distinguished by stage, subgroup and dimension. It would be almost impossible and probably unnecessary to present a detailed overview of all of their findings in this article. Nearly all the results point to the same – positive – conclusion. Therefore, only a very concise summary will be given here.

The CPC program aimed at improving outcomes in the realms of education, physical and mental health, economic well-being, and crime reduction. A wealth of positive intervention effects have been reported pertaining to, among others, parent involvement; kindergarten readiness; reading and math achievement; classroom adjustment; special education placement; grade retention; school dropout; high school completion; college degree attainment; social-emotional learning; employment; earning a living wage; income; economic well-being; problem behaviors; juvenile arrests; criminal justice involvement; incarceration; conviction for felonies; well-being in midlife; physical health; mental health; child maltreatment and neglect; health behaviors; obesity; smoking; diabetes; substance use; drug and alcohol abuse; cardio-vascular disease; and health insurance coverage (sources: see the References with "Reynolds").

What should be kept in mind is, firstly, that the rather one-sided focus of the CPC research team on positive effects, distracts from the many null and even negative effects. Relevant questions that arise are, for instance: How strong are they? Are they systematic: are there differences between domains, stages, and subgroups? Why do they occur? Secondly, the many positive effects vary quite a bit depending on participation phase (e.g., preschool CPC participation vs. no participation), background characteristics (e.g., boys vs. girls), and life stage (e.g., youth vs. midlife). As a result, there are many inconsistencies which require an explanation.

4. Limitations & criticism

4.1 Big three features

ECEC programs differ across many dimensions. In the current article, the Big Three stand central, with a focus

on CPC. In this section, an overview will be presented of criticism and limitations of CPC; where possible and if relevant, comparisons will be made with Perry and Abecedarian. Firstly, Table 1 presents a number of structural features of the Big Three.

Table 1. Comparison features "Big Three"

	Perry	Abecedarian	CPC
Period (original program)	1962-1967	1972-1977	1983-1986
Target group	Socially disadvantaged children; at-risk children; children from low-income families	Disadvantaged children; at-risk children; children from under-resourced families	At-risk children; low-income children
Ethnicity/Race	100% African-American	99% African-American	93% African-American; 7% Latino
Intelligence	Children: "mentally retarded"; IQ 61-80	Mothers: IQ 84 (avg.)	?
Ages	3-4y	6w-5y	3-9y
Duration (maximum)	2y	5y	6y
Intensity (total)	1,000h	12,500h	PS+KG: 1,890-2,520h; EIS: 2-3y full day*
Costs per child (2023 \$)	\$43,500	\$120,000	\$18,082 (PS-G3)
Teacher certificate	PS; EIS; SE	PS; EIS	PS (bachelor)
Teacher-child ratio	PS: 1 : 5-6	PS: 1 : 3; KG: 1 : 6	PS: 1 : 8.5; KG: 1 : 12.5
Parent component	Home visits by staff 1.5h/w	Home visits by school staff 1x/2w	Parents help at centers 3h/w; Home visits by School-Community Representative
Sites (treatment group)	1	1	20
Sample (at start)	T: 58 / C: 65	T: 57 / C: 54	T: 989 / C: 550
Cohorts	5	4	1
Start main data collection	1962	1972	1986
Data collection (ages)	3-54y	6w-45y	3-40y
Research method	Random experimental; Prospective	Random experimental; Prospective	Matched-group quasi-experimental; Retrospective / Prospective**
Dependency evaluator	Developer = Evaluator	Developer = Evaluator	Developer ≠ Evaluator
Effects: / Significance ES (mainly)	<i>p</i>	<i>p</i>	<i>p</i> + ES
Effects on crime	Yes	No	Yes

Note. y=year, w=week, h=hours; T=treatment group, C=comparison group; p=probability (p- value), ES=Effect Size; PS=preschool, KG=kindergarten, EIS=elementary school, SE=special education, G=grade; * actual number depends on number of years, and half-day or full-day participation; ** program / impact study (CLS)

4.2 Target group

The CPC target group formally was the same as that of Perry and Abecedarian, viz. young children living in socio-economically disadvantaged families and high-poverty neighborhoods. However, in practice, this implied that virtually all children in the three intervention programs were of African-American descent (CPC also included 7% Latinos). Evidently, this seriously restricts the

generalization of the studies' findings (which often is not mentioned; e.g., [Arteaga, 2014](#)). There is no information about, for instance, Whites, Asians, and (especially relevant for Europe) labor migrants, asylum-seekers and refugees from North Africa, the Middle East and Eastern Europe. In addition to the latter groups' problems related to their reasons for migration, they also most often have considerable culture and language difficulties. Therefore, there is no evidence, whatsoever, that these programs are

effective for children other than African-Americans living in one city in the USA. Eddy (2012, p. 9) comments: "Yet unknown are what trends and effects the CPC program will have across different populations in different environments, who are vulnerable in different ways, to different degrees." (In 2012/13 a new cohort study started, the Midwest Longitudinal Study, which was based on the new CPC version, was geographically and ethnically more diverse, and included a larger sample; Mondí et al., 2021; Reynolds, 2019.)

ECEC programs standard aim at young children growing up in socio-economically disadvantaged families. The target group consists of children who indeed have the capacities, but because of their home situation are hindered to fully realize them. The Big Three in practice have another implicit selection criterion, viz. ethnicity/race. But there is another implicit criterion which plays a role, namely intelligence. In Perry, the children's IQ was between 61 and 80 points, and they were labeled by Perry's director as "culturally deprived Negroes, diagnosed as mentally retarded" (Weikart, 1966, p. 173). Obviously, this typically is not a selection criterion, because disadvantage policies aim at children with normal capacities. This thus seriously restricts generalization of the Perry findings. In the Abecedarian program, intelligence played a role, too, as a selection criterion, but not that of the children but that of their mothers. On average, the mothers had an IQ of 84 points, which can be considered as "borderline mental disability". However, 13 mothers (thus more than 10%) had an IQ of 70 points or lower, and therefore – using the same language – were "mentally retarded" as well. This implies, that just like Perry, Abecedarian targeted an exceptional category of children, and therefore its findings cannot be generalized to other "normal" ECEC programs and populations. (Because this would imply that all socio-economically disadvantaged African-American youngsters as well as their mothers would have subnormal intelligence.) In the CPC program, intelligence did not play a (formal) role, that is, no information about intelligence is presented in its publications.

4.3 Sample size

The CLS sample has several advantages over the Perry and Abecedarian samples. In the first place, it is much larger (a total of 1539, 123 and 111 participants, respectively). In principle, this means that it provides a more precise estimate of the treatment effect and makes it easier to assess the representativeness of the sample and generalize the results (Biau et al., 2008). In addition, if attrition occurs, which is almost always the case in longitudinal studies, enough participants may remain. According to Besharov et al. (2011), Perry and Abecedarian were small model programs with highly qualified staff and ample resources. CPC, on the other hand, is a public program with a large sample and thus is more representative of what programs really look like in the field, and this, according to Reynolds

(2000, p. 63), implies that its "generalizability extends to low-income children in many public programs in central cities." Whether this is really true is the question, however, as there are many restrictions, as will be shown below.

CPC has another advantage over the Perry and Abecedarian programs, as the latter were implemented at just one site each. This evidently restricts the programs' generalization. The CPC program, on the other hand, was carried out in 20 CPC centers, which – in principle – gives its evaluation a sounder and more credible foundation. Furthermore, in the statistical analyses, the centers actually form a second, higher, measurement level. This not only offers the opportunity to estimate differences between the sites, but also their joint contribution in explaining effects. It, of course, is interesting and relevant to know whether one center functions better than the other, and, if this actually is the case, what causes this difference? To adequately examine this, multi-level analyses are preferred (Leyland & Groenewegen, 2020). Unfortunately, this is hardly ever done in the vast number of CLS studies.

Though the CLS sample was large, the children came from just one cohort. On the other hand, Perry and Abecedarian were very small, but their samples contained 5 and 4 cohorts, respectively. Seen from this perspective, the generalization of CPC is much more limited than that of the other two programs.

4.4 Significance & effect size

When working with large samples, researchers have to consider what level of significance is appropriate, because there is a strong direct relation between them (The Statsig Team, 2024). If you opt for the often-chosen 5 percent level ($p < 0.05$), the chance of finding significant effects is much greater with large samples than with small samples. Therefore, the level should be adapted to the sample size. With 1,539 participants, the 1 percent level ($p < 0.01$) is probably more appropriate than the 5 percent level. Many of the CLS publications by Reynolds and colleagues use $p < 0.05$, which means that they may find more significant effects than is warranted. In other words, the number of effects may be overestimated. As these researchers often only report effects with $p < 0.05$, it would have been much more informative if they had reported exact p -values.

For several decades, most researchers have agreed that presenting p -values is not enough, and that it is imperative to also present some sort of effect size (ES), which gives an idea of the magnitude of an effect (Sullivan & Feinn, 2012). Nevertheless, nearly all the Perry and Abecedarian studies rely on significance only. In the CLS studies, presenting effect sizes is more common, but unfortunately not universal. (Reynolds et al. assume that an ES of ≥ 0.20 is educationally meaningful.) In fact, many "effects" are expressed in terms of percentage differences, which are not really meaningful. These practices seriously hamper the way effects can be interpreted. In other words: not only may the number of significant effects be overestimated, it

also remains ambiguous how strong they are.

4.5 Selection bias & non-random attrition

The main threats to the validity of findings are selection bias in the program and non-random attrition from the sample. Longitudinal studies almost per definition suffer from attrition, and, as a rule, the longer participants are followed, the higher the attrition rate. In the CLS, participants have been surveyed for more than four decades and, therefore, it is inevitable that the sample is faced with a considerable and increasing number of dropouts. Compared with Perry and Abecedarian, the CLS has one major advantage, however. While the former studies only had very small samples right from the start, the CPC sample is considerably bigger and more robust. As a consequence, it is much easier to create a new sample with the same properties as the original.

If deemed necessary, Reynolds and colleagues used several sophisticated statistical techniques to correct for biases in program selection and attrition rates. Among others, Reynolds et al. (2001, 2023) used linear regression analysis with inverse propensity score weighting to adjust for selection and attrition bias and included well-established baseline covariates (for an explanation, see Arteage et al., 2014). Also, data for the covariates were imputed due to the large amount of missing data. (To give an impression of the problem: 25 percent of CPC preschoolers and 30 percent of comparison children had missing data on parental education or meal eligibility.) According to them, potential problems were adequately addressed. Reynolds et al. (2016) add that results have been consistent across a wide range of robustness testing by model specifications, assumptions, and alternative comparison groups. Morency et al. (2024) report that assessment of key covariates and baseline attributes showed equivalence on nearly all factors, which has been confirmed in many CPC studies.

Besharov et al. (2011) and Phillips et al. (2007), however, are not entirely convinced by these reassurances. Though the statistical approaches used to address possible selection bias may have led to consistency in program impact estimates, Besharov et al. point to the fact that only those unmeasured variables that are associated with measurable proxies can be modeled, and that some uncertainty remains given the limited data available. They also refer to a critical review by Cook and Wong (2007, p. 14), who are of the opinion that the CPC evaluation "depends on an opaque matching procedure and on data analyses (Heckman-type selection models and propensity scores) that have routinely failed to recreate similar effect sizes to an experiment on the same topic. This implies the possibility of a selection confound not fully controlled." Therefore, Besharov et al. (2011, p. 32) persist that "high levels of missing data, the absence of a true randomized experiment, and various self-selection biases raise considerable uncertainty about the findings." As a consequence, the many analyses by Reynolds et al. have resulted in rather large differences in estimated effects depending on the specific data sets used and their

completeness, and the statistical methods used. Wong et al. (2008) add to this that the CLS is a quasi-experimental longitudinal study, and not a real experiment (as Perry and Abecedarian are). Therefore, it has much less evidential value than a true experiment with random assignment, which is generally regarded as "the gold standard" in effect research. Shadiz et al. (2002, p. 503) even go further and label such quasi-experiments as "generally uninterpretable" causally.

4.6 Static & dynamic backgrounds

Besharov et al. (2011) also point out a problem inherently associated with the longitudinal character of the study. At the beginning of the CLS study, key background characteristics were collected, such as pertaining to the socio-economic status of the family and the neighborhood. (But none was collected at the children's enrollment in the program.) However, such variables are not static per se but can change in the course of decades and thus influence the effects of the intervention. One problem, then, is that developments like these do not apply to all participants in the same way and to the same degree, and that some sort of correction is therefore necessary in the analyses. Though Besharov et al. admit that Reynolds et al. attempted to statistically control for site variation, they are not confident that this captured the many differences and changes over time. Making adjusting for neighborhood differences even more difficult is that, depending on the comparison group, children did not always remain in their original neighborhoods. There was a lot of mixing, possibly owing to participating in the program.

Related to the above-mentioned point is the question of when the data collection started. In the Perry and Abecedarian evaluations, data collection began at the moment the programs were implemented. This was not the case for the CPC evaluation, however. The program study started in 1983, but it was not until 1986, that, with the start of the CLS cohort, the main background variables were retrospectively collected. This may influence data reliability negatively (Gertler et al., 2017).

4.7 Evaluator independence

The actual evaluation of the CPC program is based on data from the CLS, which is led by Arthur Reynolds and who was not involved in the development of the original CPC program. Therefore, developer and researcher are independent of each other. This is not the case for Perry and Abecedarian, however, where both programs' directors were also the main evaluators (Phillips et al., 2017). From a methodological point of view, this implies that the reported findings of CPC (in principle) are more objective.

4.8 Intervention alternatives

Part of the comparison group consisted of children who

had not participated in the preschool CPC program, but in another intervention program (like Head Start) or had not participated in any program. Reynolds et al. (2016) are of the opinion that by comparing these groups, the analyses test the impact of CPC over and beyond the other childhood services, and that this results in a conservative bias compared with previous studies. This may be so, yet, it would be relevant to check what the children in this comparison group exactly had done until they entered CPC, that is, in what alternative program, or what activities in case they didn't participate in any program (which was seldom the case), and during how many years and with what intensity. And whether there have been systematic differences related to these alternatives. As with the issue of the possible differential impact of the 20 CPC sites (see above), multilevel analyses would have been adequate here.

4.9 Ethics

In addition to criticism regarding methodology, there are also comments from a more ethical/moral perspective. The National Academies of Sciences, Engineering, and Medicine (2023) summarize some of them. They criticize the terminology used in the original CPC, the deficit perspective (instead of building on the children's strengths), the definitions of quality, and the lack of attention to the racial backgrounds of the program's participants (all of them were either African-American or Latino). Furthermore, they also challenge the traditional ECEC programs' emphasis on adapting the individual child and his/her parents to prevailing social structures, without tackling these structural barriers that are responsible for the existence and the preservation of unequal opportunities in the first place.

4.10 Comparisons

Phillips et al. (2007) add another problem to the list, which is the diversity of the comparisons. The sample consists of one treatment group and two comparison groups, which allows for numerous comparisons. Though Reynolds et al. show that for several comparisons the background characteristics do not differ (regularly after corrections are applied), they do not present this proof for all comparisons. Furthermore, there was considerable movement among the groups, such as children who participated in CPC but who nevertheless were included in the comparison group. Yet another point is that in the numerous analyses different numbers of participants were involved, which makes it difficult to compare the results. It is also important that one of the main goals of ECEC programs is to give children an even start when they enter elementary school. The comparison group then would be non-disadvantaged children. Regrettably, such comparisons are hardly ever made in the many studies (Reynolds, 2019, compares them to the national average). Therefore, it's not sure whether

this goal is achieved.

4.11 Costs & benefits

The CPC program was relatively cheap. The total costs per child were \$18,802 (in 2023), while those of Perry were three times as high, \$43,500, and those of Abecedarian amounted to a staggering \$120,000 (Driessen, 2025; Reynolds et al., 2016). Therefore, choosing one of the three, all other things being equal, is simple.

Computing the benefits of an ECEC intervention is a perilous undertaking, as the various publications of Reynolds prove. For instance, there are differences in outcomes between the studies (Besharov et al., 2011). Furthermore, there is criticism of the reference points and categories chosen for the estimates; for example, the national average instead of a specific (local) subgroup's average. The result is a lot of ambiguity as to what the financial benefits really are (Cook & Wong, 2007). In addition, there is also some confusion regarding differences found between the Big Three. Barnett (2007), specifically, points to huge differences in crime benefits across the three. Perry had enormous benefits from crime reduction (in fact, they were the program's largest benefits by far; Driessen, 2025), CPC had much smaller but still considerable benefits, and Abecedarian had, surprisingly, none at all. Crime cost savings (in 2023) were \$300,744, \$62,199, and \$0, respectively. The intriguing question then is, what could have caused these differences? Barnett (2007) suggests that differences in populations and neighborhoods could have played a role, but also program differences.

Cook and Wong (2007, p. 14-15), in general have little confidence in cost-benefit calculations: "These findings are all the more limited because of a temporal mismatch built into almost all the long-term benefit-cost calculations now available", and: "an indirect case has to be cobbled together from long-term studies implemented in a past that does not match even today, let alone any realistically imaginable future."

4.12 Then & now

The fact that at present, nearly sixty years after the first four centers were established, 26 centers are still in operation, albeit after considerable revisions, indicates that there is a high level of trust in their success. However, it is important to remember that all CLS analyses were conducted using the original CPC version of the 1983-1986 cohort. Over the years, the original program has been modified, enhanced and reconceptualized. This means that any effects reported on the basis of the CLS data are unlikely to be related to the updated CPC version. Therefore, there is no proof that the CPC/CLS findings can be generalized to the "new" CPC and other ECEC intervention programs of today. This thus substantially limits the usefulness of all of these publications.

The original, evaluated program was carried out some

forty years ago. Much has changed in society since then. The question, therefore, is how current and relevant the CLS's findings still are today. Society is not a static entity, and neither are its citizens. The same applies to pedagogy and didactics. Phillips et al. (2017, p. 3), for instance, point to the fact that nowadays the CPC target group's parents have had more years of schooling and have fewer children (and therefore can give these children more attention). They also have had greater access than in the past to publicly funded ECEC. "As a result of these differences in design, scope, characteristics of participants, and access to alternative early education programs, the bar that pre-k must exceed in order to be judged effective has been rising over time."

5. Discussion

In particular, this study shows that while many ECEC programs have been implemented, only a few were evaluated using an experimental design. Most ECEC evaluations have a retrospective quasi-experimental design, which in itself raises many methodological questions. Though neither RCT designs are perfect (Para & Edwards, 2024), they are still considered as "the gold standard". This raises the question of why there are no more experimental studies that are also large enough to inspire confidence.

Ethical objections are often raised to carry out experiments with a (normally small and selected) group of children. But what these opponents curiously forget is that the alternative, the reigning "business as usual" approach, is considerably more objectionable. Currently, millions of young children are being subjected to all sorts of programs and interventions without any scientific evidence that they are effective.

Reynolds and colleagues have written hundreds of publications. Unfortunately, there is no quantitative overview of the effects of all those studies, not only positive effects, but also null and negative effects, and preferably not just in terms of significance (p -values), but also in terms of effect sizes. This will undoubtedly be a hell of a job, but at the same time very rewarding as it will provide developers, researchers and policymakers with much needed support to make substantiated decisions.

The CPC program that formed the input for the many CLS impact studies has been updated several times. The number of sites using this new version has expanded in the course of years. As there are significant differences between the two versions, findings cannot simply be transferred from the old to the new version. Therefore, more studies focusing on the effects of the new version are urgently needed (Reynolds, 2019). Furthermore, until now, most studies have been on the effects of the program. What is equally important, however, is to learn what it exactly is that causes the program's positive outcomes (Ou et al., 2020; Phillips et al., 2017; Reynolds et al., 2019). What happened in this "black box" and how can we use this in

current and future programs?

6. Conclusion

ECEC programs are the core of many educational disadvantage policies. Though they were implemented many decades ago, the Big Three, Perry, Abecedarian, and CPC, nowadays are often still being put forward as the ultimate proof that such programs are highly effective in preventing educational delays. In addition, advocates are of the opinion that this effectiveness can be generalized to similar target groups and situations today. On the basis of a thorough literature review, this article critically appraises both claims. While the focus is on CPC, comparisons are made with Perry and Abecedarian.

Compared to the other two model programs, the CPC research team, led by Arthur Reynolds, has published an overwhelming number of studies. Their conclusions are generally very confident and optimistic. They report numerous positive effects in terms of education, physical and mental health, economic well-being, and crime reduction. Though the researchers admit that there are some methodological issues, they repeatedly state that these have adequately and sufficiently addressed. However, this has not convinced all critics.

Undoubtedly, CPC has many advantages over Perry and Abecedarian, such as the sample, sites and time period, and also the presentation of effect sizes instead of only p -values, but there are also some problems that have not been solved convincingly as well. What remains are discussions regarding, among others, the non-experimental design of the study, selection bias, and non-random attrition. One insurmountable problem that all three programs have in common concerns the target group. While all three claim that socio-economically disadvantaged children are their target group, in practice this implied that the participants in the programs were nearly all African-American. This means that all results can solely be generalized to African-American disadvantaged children in American poor neighborhoods. There is no proof, whatsoever, that the findings of the many studies can simply be generalized to other target groups, places, and times. With this in mind, it is probably justifiable to give "benefit of the doubt" as the final verdict. With regard to the two research questions of this article, the first answer is: Yes, there is ample evidence that the CPC program has many positive effects on, e.g., education, economic position, well-being, health, and crime. The second answer, however, is a more cautious one, namely that, insofar as effects occur, they can only be generalized to a specific part of the ECEC's target group, viz. African-American disadvantaged children living in high-poverty American cities. Therefore, nothing can be concluded with regard to, e.g., Whites, Asians, labor migrants, asylum-seekers and refugees.

Conflict of interest

The author declares no conflicts of interest.

Funding

The author received no funding.

References

- Aiston, S., & Walraven, G. (2024). A re-view of educational inequalities. *Educational Review*, 76(1), 1-12. <https://doi.org/10.1080/00131911.2023.2286849>.
- American Public Media (2018). *Early lessons*. December 3, 2024. <https://americanradioworks.publicradio.org/features/preschool/bigthree.html>.
- Arpawong, T., Gatz, M., Zavala, C., Gruenewald, T., Walters, E., & Prescott, C. (2023). Nature, nurture, and the meaning of educational attainment: Differences by sex and socioeconomic status. *Twin Research and Human Genetics*, 26(1), 1-9. <https://doi.org/10.1017/thg.2023.6>.
- Arteaga, I., Humpage, S., Reynolds, A., & Temple, J. (2014). One year of preschool or two – is it important for adult outcomes? Results from the Chicago Longitudinal Study of the Child-Parent Centers. *Economics of Education Review*, 40, 221-237. <https://doi.org/10.1016/j.econedurev.2013.07.009>.
- Barnett, W. (2007). Commentary. Benefit-cost analysis of early childhood programs. *Social Policy Report*, 21(3), 12-13. <https://doi.org/10.1002/j.2379-3988.2007.tb00052.x>.
- Besharov, D., Germans, P., Higney, C., & Call, D. (2011). *Assessments of twenty-six early childhood evaluations*. Welfare Reform Academy. December 12, 2024. https://welfareacademy.umd.edu/pubs/early_education/index.html.
- Bethäuser, B., Bach-Mortensen, A., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behavior*, 7, 375-385. <https://doi.org/10.1038/s41562-022-01506-4>.
- Biau, D., Kernéis, S., & Porcher, R. (2008). Statistics in brief: The importance of sample size in the planning and interpretation of medical research. *Clinical Orthopaedics and Related Research*, 466(9), 2282-2288. <https://doi.org/10.1007/s11999-008-0346-9>.
- Bruhn, J., & Emick, E. (2023). *Lottery evidence on the impact of preschool in the United States: A review and meta-analysis*. MIT Department of Economics. December 2, 2024. <https://blueprintlabs.mit.edu/research/lottery-evidence-on-the-impact-of-preschool-in-the-united-states-a-review-and-meta-analysis/lyEmick>.
- Burchinal, M., Auger, 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. <https://doi.org/10.1126/science.adn2141>.
- Cahill, K. (2021). Intersections of social class and special educational needs in a DEIS post-primary school: School choice and identity. *International Journal of Inclusive Education*, 28(7), 977-991. <https://doi.org/10.1080/13603116.2021.1968519>.
- Carter, P., & Merry, M. (2021). *Wall to wall: Examining the ecology of racial and educational inequality with research*. Spencer Foundation. December 6, 2024. <https://spencerfoundation.s3.us-east-2.amazonaws.com/store/339fb3782c08228282dec49f5c12433a.pdf>.
- Cook, T., & Wong, V. (2007). The warrant for universal Pre-K: Can several thin reeds make a strong policy boat? *Social Policy Report*, 21(3), 14-15. December 2, 2024. http://www.srcd.org/documents/publications/spr/21-3_early_childhood_education.pdf.
- DeAngelis, C., Holmes Erickson, H., & Ritter, G. (2020). What's the state of the evidence on pre-K programmes in the United States? A systematic review. *Educational Review*, 72(4), 495-519. <https://doi.org/10.1080/00131911.2018.1520688>.
- Driessen, G. (2022). *The many facets of educational disadvantage. Policies, interventions, effects*. Eliva Press. <https://doi.org/10.5281/zenodo.6641161>.
- Driessen, G. (2025). The fragile foundation of pre- and early-school programs for disadvantaged children. *Forum for Education Studies*, 3(1), [1-16]. <https://doi.org/10.59400/fes1869>.
- Duncan, G., Kalil, A., Mogstad, M., & Rege, M. (2022). *Investing in early childhood development in preschool and at home*. NBER Working Paper 29985. National Bureau of Economic Research. December 2, 2024. <http://www.nber.org/papers/w29985>.
- Eadie, P., Page, J., Levickis, P., Elek, C., Murray, L., Wang, L., & Lloyd-Johnsen, C. (2024). Domains of quality in early childhood education and care: A scoping review of the extent and consistency of the literature. *Educational Review*, 76(4), 1057-1086. <https://doi.org/10.1080/00131911.2022.2077704>.
- Eddy, S. (2012). *Developing an early childhood and education program: Implications from a family impact analysis of the child-parent center program*. Policy Institute for Family Impact Seminars. December 4, 2024. https://evidence2impact.psu.edu/wp-content/uploads/2023/04/fia_analyses_decep.pdf.
- European Commission (2014). *Key data on early childhood education and care in Europe. 2014 Edition*. Publications Office of the European Union. <https://doi.org/10.2797/75270>.
- Eurydice (2020). *Equity in school education in Europe – Structures, policies and student performance*. Publications Office of the European Union. <https://doi.org/10.2797/658266>.

- Farquharson, C., McNally, S., & Tahir, I. (2022). *Education inequalities*. Institute for Fiscal Studies. December 23, 2024. <https://ifs.org.uk/inequality/wp-content/uploads/2022/08/Education-inequalities.pdf>.
- Fukink, R., Jilink, L., & Oostdam, R. (2017). A meta-analysis of the impact of early childhood interventions on the development of children in the Netherlands: An inconvenient truth? *European Early Childhood Education Research Journal*, 25(5), 656-666. <https://doi.org/10.1080/1350293X.2017.135657>.
- García, J., Heckman, J., Leaf, D., & Prados, M. (2020). Quantifying the life-cycle benefits of an influential early-childhood program. *Journal of Political Economy*, 128(7), 2502-2541. <https://doi.org/10.1086/705718>.
- Gertler, P., Martinez, S., Premand, P., Rawlings, L., & Vermeersch, C. (2016). *Impact evaluation in practice*. Inter-American Development Bank and World Bank. <https://doi.org/10.1596/978-1-4648-0779-4>.
- Gilliam, W., & Zigler, E. (2001). A critical meta-analysis of all evaluations of all state-funded pre-school from 1977 to 1998: Implications for policy, service delivery and program evaluation. *Early Childhood Research Quarterly*, 15(4), 441-473. [https://doi.org/10.1016/S0885-2006\(01\)00073-4](https://doi.org/10.1016/S0885-2006(01)00073-4).
- Giudici, A., Gruber, O., Schnell, P., & Pultar, A. (2024): Far-right parties and the politics of education in Europe. *Journal of Contemporary European Studies*, 1-14. <https://doi.org/10.1080/14782804.2024.2352518>.
- Gorard, S., See, B., & Siddiqui, N. (2020). What is the evidence on the best way to get evidence into use in education? *Review of Education*, 8(2), 570-610. <https://doi.org/10.1002/rev3.3200>.
- Hahn, R., & Barnett, W. (2023). Early childhood education: Health, equity, and economics. *Annual Review of Public Health*, 44, 75-92. <https://doi.org/10.1146/annurev-publhealth-071321-032337>.
- Hanushek, E., Peterson, P., Talpey, L., & Woessmann, L. (2019). The achievement gap fails to close: Half century of testing shows persistent divide between haves and have-nots. *Education Next*, 9(3), 8-17. December 10, 2024, from <https://www.educationnext.org/achievement-gap-fails-close-half-century-testing-shows-persistent-divide/>.
- Heckman, J. (2000). *Policies to foster human capital*. Research in Economics, 54(1), 3-56. <https://doi.org/10.1006/reec.1999.0225>.
- Heckman, J. (2011). The economics of inequality. The value of early childhood education. *American Educator*, 35(1), 31-47. December 2, 2024. <https://files.eric.ed.gov/fulltext/EJ920516.pdf>.
- Heckman Equation (2013). *James Heckman changes the equation for American prosperity*. December 23, 2024. https://heckmanequation.org/wp-content/uploads/2014/05/F_Heckman_Brochure_041515.pdf.
- Heckman, J., Moon, S., Pinto, R., Savelyev, P., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94(1-2), 114-128. <https://doi.org/10.1016/j.jpubeco.2009.11.001>.
- Isaacs, J. (2008). *Model early childhood programs. Impacts of early childhood programs*. Brookings. December 23, 2024. https://www.brookings.edu/wp-content/uploads/2016/07/09_early_programs_brief4.pdf.
- Joo, Y., Magnuson, K., Duncan, G., Schindler, H., Yoshikawa, H., & Ziol-Guest, K. (2020). What works in early childhood education programs? A meta-analysis of preschool enhancement programs. *Early Education and Development*, 31(1), 1-26. <https://doi.org/10.1080/10409289.2019.1624146>.
- Lenkeit, J., Hartmann, A., Ehlert, A., Knigge, M., & Spörer, N. (2022). Effects of special educational needs and socioeconomic status on academic achievement. Separate or confounded? *International Journal of Educational Research*, 113(1), 101957. <https://doi.org/10.1016/j.ijer.2022.101957>.
- Leyland, A., & Groenewegen, P. (2020). *Multilevel modelling for public health and health services research*. Springer. https://doi.org/10.1007/978-3-030-34801-4_3.
- Liu, J., Peng, P., Zhao, B., & Luo, L. (2022). Socioeconomic status and academic achievement in primary and secondary education: A meta-analytic review. *Educational Psychology Review*, 34, 2867-2896. <https://doi.org/10.1007/s10648-022-09689-y>.
- Long, K., & Renbarger, R. (2023). Persistence of poverty: How measures of socioeconomic status have changed over time. *Educational Researcher*, 52(3), 144-154. <https://doi.org/10.3102/0013189X221141409>.
- Lu, Y., Yan, K., & Wang, C. (2024). Political orientation and education investment: An OECD perspective. *Business and Politics*, 26(4), 489-504. <https://doi.org/10.1017/bap.2024.2>.
- Maisuria, A., & Lally, C. (2024). *Education: Inequalities and attainment gaps*. Post UK Parliament. <https://doi.org/10.58248/HS85>.
- Mann, G., Carrington, S., Lassig, C., Mavropoulou, S., Sagers, B., & Poed, S. (2023). Closing special schools: Lessons from Canada. *The Australian Educational Researcher*, 51, 1729-1747. <https://doi.org/10.1007/s13384-023-00661-5>.
- Marks, G., & O'Connell, M. (2024). Student achievement is much more about cognitive ability and genetics than SES: A response to Debowere. *Review of Education*, 12, e3483. <https://doi.org/10.1002/rev3.3483>.
- Melhuish, E., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K., Tawell, A., Slot, P., Broekhuizen, M., & Leseman, P. (2015). *A review of research on the effects of Early Childhood Education and Care (ECEC) upon child development*. European Commission. December 5, 2024. <http://ecec-care.org/resources/publications>.
- Mondi, C., & Reynolds, A. (2021). Socio-emotional learning among low-income prekindergartners: The

- role of individual factors and early intervention. *Early Education and Development*, 32(3), 360-384. <https://doi.org/10.1080/10409289.2020.1778989>.
- Morency, M., Reynolds, A., Loveman-Brown, M., Kritzik, R., & Ou, S. (2024). Structural equality and support index in early childhood education. *JAMA Network Open*, 7(8), e2432050. <https://doi.org/10.1001/jamanetworkopen.2024.32050>.
- National Academies of Sciences, Engineering, and Medicine (2023). *Closing the opportunity gap for young children*. The National Academies Press. <https://doi.org/10.17226/26743>.
- OECD (2006). *Starting strong II: Early childhood education and care*. OECD Publishing. <https://doi.org/10.1787/9789264035461-en>.
- OECD (2021). Towards equity in school funding policies. *OECD Education Policy Perspectives*, No. 41. OECD Publishing. <https://doi.org/10.1787/6a3d127a-en>.
- Ou, S., & Reynolds, A. (2020). Preschool education and school completion. *Encyclopedia on Early Childhood Development [online]*. December 3, 2024. <https://www.child-encyclopedia.com/school-success/according-experts/preschool-education-and-school-completion>.
- Ou, S., Yoo, S., & Reynolds, A. (2021). Educational growth trajectories in adulthood: Findings from an inner-city cohort. *Developmental Psychology*, 57(7), 1163-1178. <https://doi.org/10.1037/dev0001198>.
- Ou, S., Mondì, C., Yoo, S., Park, K., Warren, B., & Reynolds, A. (2020). Thirty years later: Locating and interviewing participants of the Chicago Longitudinal Study. *Early Childhood Research Quarterly*, 51, 1-13. <https://doi.org/10.1016/j.ecresq.2019.08.002>.
- Parra, J., & Edwards, D. (2024). Challenging the gold standard consensus: Randomised controlled trials (RCTs) and their pitfalls in evidence-based education. *Critical Studies in Education*, 65(5), 513-530. <https://doi.org/10.1080/17508487.2024.2314118>.
- Phillips, D., Lipsey, M., Dodge, K., Haskins, R., Bassok, D., Burchinal, M., Duncan, G., Dynarski, M., Magnuson, K., & Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge on pre-kindergarten effects*. A consensus statement. Brookings. December 23, 2024, from https://www.brookings.edu/wp-content/uploads/2017/04/consensus-statement_final.pdf.
- Reardon, S. (2012). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. *Community Investments*, 24(2), 19-39. December 3, 2024. https://www.frbsf.org/wp-content/uploads/CI_Summer2012_Reardon.pdf.
- Reynolds, A. (2019). The power of P-3 school reform. *Phi Delta Kappan*, 100(6), 27-33. <https://doi.org/10.1177/0031721719834025>.
- Reynolds, A. (2000). *Success in early intervention: The Chicago Child-Parent Centers*. University of Nebraska Press. December 3, 2024. <https://www.nebraskapress.unl.edu/nebraska-paperback/9780803245426/success-in-early-intervention/>.
- Reynolds, A., Magnuson, K., & Ou, S. (2010). Preschool-to-third grade programs and practices: A review of research. *Children and Youth Services Review*, 32(8): 1121-1131. <https://doi.org/10.1016/j.childyouth.2009.10.017>.
- Reynolds, A., Hayakawa, M., Candee, A., & Englund, M. (2016). *CPC P-3 Program Manual: Child-Parent Center Preschool-3rd Grade Program*. University of Minnesota. December 2, 2024. <https://cpcp3.org/wp-content/uploads/2024/08/Final-CPC-P-3-manual-8-27-20.pdf>.
- Reynolds, A., Ou, S., Mondì, C., & Giovanelli, A. (2019). Reducing poverty and inequality through preschool-to-third-grade prevention services. *American Psychologist*, 74(6), 653-672. <https://doi.org/10.1037/amp0000537>.
- Reynolds, A., Temple, J., Robertson, D., & Mann, E. (2001). Long-term effects of an early childhood intervention on educational achievement and juvenile arrest: A 15-year follow-up of low-income children in public schools. *JAMA*, 285(18): 2339-2346. <https://doi.org/10.1001/jama.285.18.2339>.
- Reynolds, A., Ou, S., Eales, L., Mondì, C., & Giovanelli, A. (2021). Assessment of a comprehensive early childhood education program and cardiovascular disease risk in midlife. *JAMA Network Open*, 4(8), e2120752. <https://doi.org/10.1001/jamanetworkopen.2021.20752>.
- Reynolds, A., Varshney, N., Ou, S., Kritzik, R., & Loveman-Brown, M. (2023). Early childhood education and midlife ideal cardiovascular health in a prospective urban cohort. *JAMA Pediatrics*, 177(12), 1350-1352. <https://doi.org/10.1001/jamapediatrics.2023.4010>.
- Schmutz, R. (2024). Is universal early childhood education and care an equalizer? A systematic review and meta-analysis of evidence. *Research in Social Stratification and Mobility*, 89(1), 1-15. <https://doi.org/10.31235/osf.io/2vjba>.
- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Company. December 4, 2024. <https://iaes.cgiar.org/sites/default/files/pdf/147.pdf>.
- Shero, J., & Hart, S. (2022). Methodological decisions and their impacts on the perceived relations between school funding and educational achievement. *Frontiers in Education*, 7, 1043471. <https://doi.org/10.3389/educ.2022.1043471>.
- Shevlin, M., & Banks, J. (2021). Inclusion at a crossroads: Dismantling Ireland's system of special education. *Education Sciences*, 11(4), 161. <https://doi.org/10.3390/educsci11040161>.
- Sullivan, G., & Feinn, R. (2012). Using effect size – or why the p value is not enough. *Journal of Graduate Medical Education*, 4(3), 279-282. <https://doi.org/10.4300/JGME-D-12-00156.1>.
- Tan, C. (2024). Socioeconomic status and student learning:

- Insights from an umbrella review. *Educational Psychology Review*, 36, 100. <https://doi.org/10.1007/s10648-024-09929-3>.
- The Statsig Team (2024). *5 Key factors in determining significance levels*. December 2, 2024. <https://www.statsig.com/perspectives/-key-factors-in-determining-significance-levels>.
- The Sutton Trust (2024a). *Closing the attainment gap*. The Sutton Trust. December 2, 2024. <https://www.suttontrust.com/wp-content/uploads/2024/02/Closing-the-attainment-gap.pdf>.
- The Sutton Trust (2024b). *Inequality in early years education*. The Sutton Trust. December 2, 2024. <https://www.suttontrust.com/our-research/inequality-in-early-years-education/>.
- Von Suchodoletz, A., Lee, D., Henry, J., Tamang, S., Premachandra, B., & Yoshikawa, H. (2023). Early childhood education and care quality and associations with child outcomes: A meta-analysis. *PLoS ONE*, 18(5), 1-26. <https://doi.org/10.1371/journal.pone.0285985>.
- Weikart, D. (1966). Preschool programs: Preliminary findings. *Journal of Special Education*, 1(2), 163-181. <https://doi.org/10.1177/002246696600100208>.
- Whitaker, A., Burchinal, M., Jenkins, J., Watts, T., Duncan, G., Hart, E., & Peisner-Feinberg, E. (2023). *Why are preschool programs becoming less effective?* EdWorkingPaper: 23-885. Annenberg Institute at Brown University. <https://doi.org/10.26300/smqa-n695>.
- Wong, V., Cook, T., Barnett, W., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management*, 27(1), 122-154. <https://doi.org/10.1002/pam.20310>.