

Abecedarian: An Impossible Model Preschool Program

Subjects: Education & Educational Research

Contributor: Geert Driessen

The primary goal of pre- and early-school programs is to prevent young children from socioeconomically disadvantaged backgrounds to start school already with educational delays. The programs offer compensatory stimulation activities which are supposed to be not available in the home situation; the focus is on language development. Proponents claim that such programs can be effective, provided they are of high quality. The belief in their success is very much based on the outcomes of a few so-called model programs from the 1960s and 1970s. One of these programs is the Carolina Abecedarian Project, a small single-site project started in 1972. Four cohorts of in total 111 children and their poor, Black parents participated in this experiment with a random allocated treatment and control group. The children were followed from 6 weeks after birth to 6 years of age, that is, when they entered school. They were regularly tested and observed, and then after the program had ended again until they were 40 years of age.

Keywords: Pre- and Early-School Education ; Early Childhood Education ; educational disadvantage ; effects ; internal validity ; external validity ; experiment ; Abecedarian ; Perry Preschool

1. Programs for Young Disadvantaged Children

Pre- and early-school programs are meant for young children from socioeconomically disadvantaged backgrounds to compensate for specific educational stimulation which is not available for them in their home environment, but which is seen as more or less essential for a successful school career. In contrast to middle- and upper-class environments, such stimulation is often lacking in lower-class and (linguistically) minority families.^{[1][2]} In this context, the European Commission (2023)^[3] is of the opinion that “Educational attainment should be decoupled from social, economic and cultural status”.

Such programs exist in many capacities. Their common aim is to reduce or – even better – prevent educational disadvantage in young children, generally between birth and the age of six (infants, toddlers, preschoolers), that is, till the children start their formal schooling. The underlying idea is that combating disadvantage at an early age is much more efficient and effective than trying to catch up at a later point in the school career.^[4] “High-quality early childhood education and care lays the foundations for later success in life in terms of education, well-being, employability and social integration.”^[3]

Programs for disadvantaged children are a special case of Early Childhood Education and Care (ECEC), which is a general provision and open to *all* children. According to the European Commission^[3] every child in the European Union has the right to affordable and high-quality ECEC. Because of their specific home situation, disadvantaged children need extra stimulation, instead of or over and above ECEC.

Pre- and early-school programs are provided in institutions (various types of childcare and the early/kindergarten years of elementary schools), sometimes combined with support for parents in the home situation.^{[5][6]} Through play, the children learn the skills they need to make a good start at elementary school. According to Slot (2014: 8),^[7] “The quality of young children’s environments is critical to the start children make in their lives, with quality referring to the emotional support, secure social relationships, cognitive stimulation, exposure to language models, and opportunities to gain control over activities and to develop self-regulation that are provided on a regular basis and in a consistent manner across the different contexts of child development throughout early childhood.”

2. The Effectiveness of Stimulation Programs

Since the early 1960s, a great diversity of programs have been developed and implemented, mostly in the USA, but a few decades later also in many European and Asian countries.^{[8][9]} Some programs have been accompanied from the beginning by rigorous evaluations, but most lack adequate effect studies.^{[1][10][11]} For instance, in the Netherlands, where programs are provided and funded by the national and local government for almost 25 years now, only a handful of “effect

studies” have been conducted, and, furthermore, none of them meets the so-called gold research standard, that is, a randomized controlled trial (RCT) or true experiment.^{[12][13]} As a consequence, every year some 70,000 Dutch disadvantaged children participate in preschool and 35,000 in early-school compensation programs (2015 data)^[13] for none of which there is any evidence that they are effective. And this probably applies to most other countries as well. A lot of controversy thus exists regarding the effects and effectiveness of the pre- and early-school programs.^{[1][7][14][15][16][17][18][19]} According to some, there is ample evidence that the programs work; others, however, argue that most program evaluations show serious methodological flaws and – apart from that – no or even negative effects.

Two early programs, so-called model programs or demonstration programs, are often mentioned and put forward as examples of how adequate programs should be developed, implemented, and evaluated, that is, as true experiments, and that these programs generate significant positive effects, even in the long run.^{[1][7]} The first program, the HighScope Perry Preschool Program (or for short Perry) was a small single-site project started in 1962, the second, the Carolina Abecedarian Project, also was a limited single-site project, which was started ten years later in 1972. Both interventions show many resemblances. In another paper, Perry and its evaluation and results were critically examined.^{[20][21][22]} The main conclusion was that Perry, despite the claims made by, especially, Nobel laureate James Heckman, is not a model program at all, mainly because its participants were unique, specifically as they had been selected because of their exceptionally low IQ-scores (“culturally deprived Negroes, diagnosed as mentally retarded”, according to Perry’s founder David Weikart (1966: 173).^[23] This selection criterion deviates fundamentally from that employed in regular pre- and early-school programs, where parental socioeconomic status is the criterion. In the present paper Abecedarian stands central; one of its chief objectives was to demonstrate that sociocultural retardation can be prevented. The main question here is whether Abecedarian really constitutes a “model” project/program.

3. The Abecedarian Project

3.1. Sample

The Abecedarian Approach was applied in three successive interventions; this paper focuses in the first, the single-site Abecedarian Project (or for short Abecedarian). (For the following description, in particular use was made of several publications.)^{[24][25][26][27][28]} The sample included a total of 111 poor, high-risk families in Chapel Hill, NC, whose children were born between 1972 and 1977 (four waves). The children were randomly assigned to an experimental treatment group (57 children) or a control group (54 children). The criteria for inclusion were a High Risk Index score of 11 or greater (out of 13), and being healthy at birth (i.e., free of biological conditions associated with mental, sensory, or motor disabilities). The index included, among others: a very low income (well below 50% of the federal poverty line); very low levels of parental education (approx. 10 years); single parenthood (76% fathers absent); unemployment of mothers (almost all). A typical mother was not married, had no income, and lived in with her parents. All these characteristics are well-known indicators of socioeconomic disadvantage and are applied universally in educational funding policies.^[15] However, the criterion of being healthy at birth is a rather unique one (because there are hardly any programs starting so early). Also exceptional is the low IQ of mothers (avg. 84; 13 mothers had an IQ of 70 or less and were labeled as “retarded”). The core of educational disadvantage policies is that the children have the capacities, but that the circumstances (“educational capital” or “cultural capital”) in their home situation are unfavorable;^{[29][30][31]} the latter does not include a low IQ of the mothers (if only that this would suggest that disadvantaged children have less intelligent mothers). Other unique features of the participating parents are: the mothers were rather young (on avg. 20 years of age; but several even younger than 16); 98 percent were African-American (thus no other ethnicities); and all spoke English at home. The latter also is quite remarkable. The problem typically is that when disadvantaged children start (pre-)school, they do not speak English. In the USA, more than half of the school-age Spanish and Chinese children speak another language than English at home.^{[32][33]} In the Netherlands, two-thirds of the preschool immigrant children do not speak the national language (which is Dutch), but a minority language (such as Turkish or Berber).^[34] Learning English or Dutch in these two countries therefore is the primary goal of pre- and early-school programs. All in all, the Abecedarian Project participants constituted a pretty unique sample, which by no means is representative for disadvantaged populations normally participating in pre- and early-school programs.

3.2. Program

But there is more that made the Abecedarian Project a truly exceptional endeavor. The treatment group was being enrolled in (just) one child development center starting as early as 6 weeks of age, and lasting until they entered public kindergarten, that is, a total of five years. This is extremely early and long; as a reference point: the Perry Program was for 3- and 4-year-olds, and Head Start programs serve children between 3 and 5 year old. In addition to the deviating duration of the program, its intensity is even more notable: it was provided full day (7:30 a.m. to 5:30 p.m.; up to 10 hours

per day), 5 days a week, 50 weeks per year (a total of 12,500 hours). Children even attended when they were ill; they then received medical help. In fact, the children got their primary medical care on site through staff pediatricians, nurses, and physical therapists. In addition, home – center transportation was provided to ensure attendance. Almost all children participated fully. For comparison: Perry's treatment consisted of 2.5 hour educational preschool on weekdays during two school years (a total of 1,000 hours). The typical part-day preschool program in the US provides 450–540 hours per year (2.5–3 h/day, 180 days).^[24] In the Netherlands, funded preschool programs are being offered to disadvantaged children between the ages of 2.5 and 4 (1.5 year, a total of 960 hours).

The treatment the experimental group received at the center is considered as extremely high-quality. The program had strong supervision, a well-designed curriculum, well-compensated staff, and was accompanied by an ongoing evaluation. The intervention focused on the domains of knowledge, language, and behavior. The educational activities themselves were game-based and emphasized language development; practices were designed to be highly engaging, fun, and active, with learning occurring throughout the day. In addition to activities at the daycare center school, the Abecedarian approach included activities for the mothers at home. In the school phase, a so-called resource teacher was assigned to each child and mother, and prepared an individualized set of home activities to supplement the school's curriculum in reading and math, taught the mothers how to use these activities with their children, tutored children directly, met regularly with classroom teachers to ensure that home activities aligned with the ones being taught in school, served as a consultant for the classroom teacher when problems arose, and advocated for the child and family within the school and community. The resource teachers made some 15 home visits per year; in addition, they offered children a variety of summertime supports, including help with summer camp, trips to the public library, and tutoring in reading skills.

Summarizing, the main features of the program were:

- an education program that started in early infancy;
- full-day, full-week, year-round, five years;
- a very high-quality program designed and conducted by a large multidisciplinary team of investigators;
- a structured curriculum grounded in developmental theory and research findings;
- a highly trained and actively monitored teaching staff (the majority of which were African-American), that documented each child's progress with bi-weekly summaries;
- very low teacher : child ratios (1 : 3 for babies; 1 : 4 for toddlers and age 2; 1 : 6 for ages 3 and 4);
- ongoing professional development, with weekly meetings and frequent monitoring and in-classroom supports, active instruction and supervision;
- stable and stimulating adult – child interactions to ensure that each child engaged daily in many rich and varied language and learning activities;
- individualization of pace and attention to the special needs of the child and family;
- the use of varied, complex, and informative language (Standard English) by the teachers throughout the day;
- referral of children to specialists if needed, based on systematic and frequent assessments in language, cognition, social-emotional progress, and health;
- provision of high-quality health and safety practices;
- provision of home – center transportation;
- supportive social services provided to families facing problems with housing and food;
- parent (mostly mothers) engagement: group meetings, special topic sessions, and teacher – parents meetings;
- provision of a book- and toy-lending library for families.

3.3. Effects

In the Abecedarian Project, many aspects of the children's growth and development were measured at frequent intervals during their first five years. These included cognitive, linguistic, and social-emotional measurements for the children and educational and employment status of their mothers. Abecedarian families were visited and observed at home when the children were 6, 18, 30, 42, and 54 months of age. Data were also collected at ages 8, 12, 15, 21, 30, 35, and 40, that is, during the school years and adulthood. Young adult outcomes were collected through means of semi-structured interviews and included educational attainment, employment, teen parenthood, and criminal behavior; the use of illegal drugs was also measured.

Ramey and Ramey^[35] present a summary of Abecedarian's effects; some highlights follow.

From birth to kindergarten entry:

- at 7 out of 9 measurements Abecedarian children scores significantly $p < 0.05$) higher on an IQ test than control group children (effect sizes 0.39 – 1.45).

During the school years, Abecedarian children had significantly:

- higher achievement scores in reading and math than the control group (ages 8, 12, and 15);
- lower rates of grade retention;
- lower rates of placement in special education (by age 15).

At ages 21, 30, 35, and 40, the Abecedarian children showed significantly more favorable outcomes, such as:

- higher rates of holding a skilled job;
- being employed full-time and/or enrolled in higher education;
- much lower rates of becoming a teen parent;
- few reports of depression and illegal substance use;
- higher academic and cognitive achievement scores;
- much lower reliance on welfare or public assistance programs and better (self-rated) overall health;
- fewer risk factors for cardiovascular and metabolic disease, especially for males;
- an increased size of the whole brain, the cortex, and 4 of the 5 regions of interest, more for males than females;
- a brighter future, healthier, more productive, and stronger promoters of the norms on which our society is built.

The study found no statistically significant effects on high school graduation rates, income, type of employment, marital status, mental or physical health, criminal activity, or substance use. Contrary to the Perry findings, the incidence of youth crime at the age of 21 showed no statistically significant differences. Early results also indicated that the program children were more aggressive (e.g., kicking or hitting). In general, the Abecedarian program did not produce the gains in social and emotional development reported in other preschool projects to account for a very large portion of potential benefits.

3.4 Costs

Abecedarian thus produced considerable scholastic effects; however, the question is, at what costs?^[1] The total costs per child participating in the Abecedarian Project accumulated to a staggering \$120,000 (in 2023 dollars).^[36] For comparison: the total costs for a Perry child were \$27,085,^[24] the yearly costs for a Head Start child are \$11,392. A child participating in a preschool program in the Netherlands (ages 2,5 - 4) costs some \$5,500 per year.^[13] According to Duncan et al.,^[1] the benefit/cost ratio for Abecedarian was 2.5 : 1; for Perry it was 9 : 1. One important difference between these two “model” programs was the absence of any crime benefits for Abecedarian, compared to nearly \$175,000 in crime “savings” for Perry (largely caused by the societal costs of murders in Perry; see Driessen^[20]). A possible explanation is the difference in crime incidences in the respective areas: while the Perry children lived in an extremely poor, high-risk area, the Abecedarian children lived in a relatively affluent area. Garcia et al.^[37] arrive at another, more favorable benefit/cost ratio. Using a complete dataset and refined statistical analyses, they estimate the benefit/cost ratio to be 7.3 : 1 with an annualized rate of return of 13.7% for the cost of participating 5 years in the Abecedarian project.

3.5. Relevance

Is the Abecedarian Project, which was carried out in the early 1970s, that is, half a century ago, still relevant today with respect to populations and circumstances? According to Pages et al.^[27] it is. Despite the fact that between then and now there have been significant changes in, for instance, terms of populations, maternal employment, and economy, Abecedarian is still relevant because we now know that learning and development occur early and are cumulative and consequential, and that therefore, a systematic and comprehensive approach that begins early is needed. In this interpretation, the relevance lies a general realization that starting early can be important, but not in the Abecedarian approach or project itself.

Farran and Lipsy^[38] are of another opinion. According to them, an often-mentioned argument pertains to savings generated by Abecedarian (and two other model programs, Perry, and Chicago Parent Child Centers). Later and contemporary preschool programs, however, differ greatly from these intensive small-scale demonstration programs. Suggesting that comparable outcomes can be achieved by preschool programs that cost less and differ hugely is unsupported by any available evidence, on the contrary, it suggests just the opposite.

Whitehurst^[39] concludes that the best available evidence raises serious doubts that a large public investment in the preschool expansion will have the long-term effects that advocates advertise. In another article, Whitehurst^[40] summarizes his objections: “There is nothing now available to parents called childcare or daycare that is even grossly similar to Abecedarian in the program that is delivered, the characteristics and social circumstances of the children and families that are served, the teachers and staff who are employed, the age at which children are initially

enrolled (6 weeks), the continuity of enrollment from infancy to 5 years, the delivery of on-site primary health care, program leadership and management, or costs. As a program from which one can generalize results with confidence to present day public policies on childcare, Abecedarian fails abysmally.”

The very unusual duration of the Abecedarian project makes it incomparable with other large-scale pre- and early-school programs. However, there is more to this: the children not only received 5 years of pre-school stimulation, but half of children, both in the experimental and in the control group received 3 more years (grades 1 – 3) of educational assistance thereafter.^[41] The total of (intensive) extra stimulation and help thus amounts to an incredible 8 years.

Earlier in this paper, the exceptionally low IQ of the mothers was mentioned, which is seldom put forward as a problem. Besharov et al.^[42] however, pointed to the fact that, though the Abecedarian researchers concluded that the program resulted in positive and lasting gains on a wide range of cognitive and school-related outcomes, these gains were concentrated among the subgroup of children whose mothers had IQs below 70, the “retarded” (the terminology used at the time) mothers. Furthermore, these gains did not lead to many improved outcomes when the children were age 21. Hu,^[43] in this connection, refers to a study by psychologist Arthur Jensen, who commented on the unexpected and unexplained mother IQ - child IQ (aged 3 years) correlation, which was -0.05 in the experimental group and 0.43 in the control group. A possible, but unconfirmed, explanation could be the children’s test’s validity.

The project children have been tested regularly for their IQ development, at 3, 6, 12, 18, 24, 36, 42, 48, 54, 60 months of age (data pertaining to the month 3 measurement have not been published). Four different tests were used: the Bayley Scales of Infant Intelligence; the Stanford-Binet Intelligence Scale, Form L-M; the McCarthy Scales of Children’s Ability; and the Wechsler Preschool and Primary Scale of Intelligence. Each of the tests measures different domains and included mother – child observations and questionnaires. There are several problems with this, such as reliability, validity, and objectivity. Spitz,^[44] for instance, cautions that the project staff were aware that frequent retesting on the same instrument may influence the results, especially when the child’s mother is also present at every test taking. In addition, there is the possibility that mothers will provide their infants with practice. Thus, both staff and mothers may have inadvertently been “teaching to the test”. Furthermore, there are direct effects of repeated testing.

Olsen and Snell (2006: 23)^[45] refer to Spitz, who was of the opinion that Abecedarian authors presented some of their findings in a biased way, favoring Abecedarian. In the Abecedarian publications and analyses, generally a distinction is made between an experimental and a control group. However, there actually were four cohorts involved, and thus four by two combinations. By combining the findings regarding the children’s IQ development, the researchers concluded that Abecedarian raised IQ. “However, they neglected to report that scores improved only for two of the four groups. In fact, for the third and fourth cohorts, the experimental group actually lost 3.68 IQ points more than did the control group, providing no support for the efficacy of the intervention on this measure.”

Spitz^{[44][46][47]} also noted that the IQ difference between the experimental and the control group at the age of 5 was already present at age six months, suggesting that 4.5 years of very intensive stimulation activities ended with practically no effect. According to Spitz this could mean that the IQ difference between both groups may have been latently present right from the beginning due to randomization gone wrong. Several of Spitz’s criticisms were contested by Abecedarian director Ramey,^[48] however.

Ponniah^[49] doubts that it would be possible “to replicate the experiment and achieve the same results in different cultural contexts, and to provide ECE of the same type and with the same very high standards consistently to a larger number of children. The results therefore should not be generalised by making claims that ECE has long term benefits for children beyond the context of the Abecedarian Project experiment.”

The Abecedarian Project also has raised methodological criticism, which undermines, restricts and reduces the effects as reported by several scholars. According to Whitehurst,^{[39][40]} the study has serious problems in terms of internal and external validity. Abecedarian was designed as a randomized trial, but there were compromises in the random assignment protocol that are likely to have pushed upwards the effect estimates. Nobel laureate James Heckman wrote a series of papers analyzing the Abecedarian data and thereby claiming large positive effects,^{[37][50]} his conclusions depend on the integrity of that random assignment. Which is especially important, since the sample was quite small. Therefore, the question is what the implications are of all this for Heckman’s conclusions. “It is a hothouse university-based program from nearly a half century ago for a few dozen children from very challenging circumstances who were deemed to be at risk of mental retardation. Its relevance to present day policies on childcare for the general population is uncertain, at best. Even ignoring this failure of external validity, the reported results from Abecedarian favoring the treatment participants are in doubt because the evaluation of Abecedarian’s impacts on participants is seriously compromised by a large imbalance in decline-to-participate rates by those assigned to the treatment vs. the control condition; and by the presence in the treatment group of an appreciable proportion children who were not randomized into that condition.” (Whitehurst, 2017).

Abecedarian is one of the early small-scale experimental projects. Bruhn and Emick^[14] point to differences with later large-scale programs. While estimates of the former generate convincing causal inferences and therefore contain a high degree of internal validity, it is not always possible to translate them to other times, places, and populations. As a consequence, such evidence may reduce the external validity.

The small sample of participating children and parents also poses a problem because of attrition, which makes the sample even smaller. Attrition increased during the various measurement rounds. Though the differences in attrition often were comparable for both experimental as control group, for several variables attrition was substantial. For instance, attrition for welfare receipt at age 30 amounted to 31%, which means that per group only 37 subjects remained. At follow-up measurements, attrition increased even more and, furthermore, differed per group, and thereby weakening the reliability of the study's finding considerably.^[51]

Whitehurst^[33] points to yet another problem: several hundred outcome variables were analyzed, but the researchers did not properly adjust for the likelihood that 5% of those outcomes would appear to be statistically significant simply on the basis of chance. When the data are adequately analyzed, most of the differences disappear. For this argument he refers to the reanalyses of Anderson (2008),^[52] which focuses on two issues: effect heterogeneity by gender and over-rejection of the null hypothesis due to multiple inference. The primary findings are that girls gained substantial short- and long-term benefits from the intervention, but there were no significant long-term benefits for boys. These conclusions have thus far appeared ambiguous when using "naive" estimators that fail to adjust for multiple testing. Anderson therefore advises that in such situations to either combine measures, or report both adjusted and unadjusted p-values. The small sample poses a problem, not only for level of significance, but also for the effect sizes. Slavin and Smith^[53] show that in such experiments with small samples any effects will be dramatically inflated. For instance, in a sample of less than 50 children the average effect is 0.44, but in a sample of more than 2,000 children this effect is not more than 0.09 (also see Duncan & Magnuson, 2013^[11]). Also relevant here is the commonly advised rule of thumb that for many analyses you need a minimum of 100 subjects or 10 subjects per variable.^{[54][55]} This poses a problem here as the Abecedarian sample is very small, indeed, and, in addition, there is attrition, especially at follow-up time points, though testing for baseline variables differences across groups, at each time-point, was never statistically significant.^[27]

Bruno and Iruka (2022: 165)^[56] point to a more fundamental problem. They argue that "the study's 'colorblind' approach and deficit-model orientation limit its ability to inform practice and policy for promoting positive developmental outcomes in the population of Black children represented in the sample." The latter suggests "that people of color in general, and Black people in particular, are deficient because they lack the education, resources, and experiences the dominant White culture has determined promote positive outcomes." (Bruno & Iruka, 2022: 170).^[56]

Another fundamental issue is ethical in nature. The question is why these mother actually have brought children into the world. During the first five years of their lives, the children almost permanently were at the center. Early in the morning, the children were even picked up and delivered at the center, and in the evening vice versa. This means that, from a pedagogical perspective, during these crucial years, the mothers saw their children only in the weekends. The question is what this has done to the mother – child bond and attachment. And then there is the question what the mothers did all day, while their baby's, infants and toddlers were away at the center. It has been suggested that because of this freedom, the mothers now were able to follow a study. But there is no information on this in the project's publications.

3.6. Developments

Several researchers report a downward development in the effectiveness of pre- and early-school programs. Pages et al.,^[27] for instance, find that evaluations of more recent programs have shown smaller short-term impacts than early programs, and then especially the so-called model programs. Reasons for this may be a combination of scaling difficulties and improvements of counterfactual conditions. Another possible explanation is that the Abecedarian project provided an educational setting from birth to kindergarten where most of the teachers were Black. There is some evidence showing that teacher - child racial match can improve a teacher's perception of a child's behaviors and abilities, and as a consequence can lead to positive longer-term outcomes. In this way, the combination of the project's duration and organizational context might have contributed to a supportive socialization and to the children's development of relevant skills.

Whitaker et al. ^[35] find that while model preschool programs from the 1960s and 1970s reported positive effects, on the short and long run, evaluations of recent programs produce ambiguous results, including divergent, weaker and even negative effects. Potential explanations for this development focus on changes in instructional practices and counterfactual conditions (improvement of early-childhood enrichment opportunities as a consequence of increased quality of children's home environments and the expansion of community-based preschool).

On the basis of a review of the literature, Slot^[17] concludes that the strongest evidence for the compensatory effects comes from a few small experiments, model programs, targeting socioeconomically disadvantaged children. These report moderate to large effects on cognitive and non-cognitive outcomes, in the short and in the long run, in some instances persisting far into adulthood. However, effects in general population studies focusing on universal large-scale ECEC systems are much smaller. According to her, the smaller effects probably point to lower overall quality in large-scale ECEC systems compared to the small-scale experimental studies evaluating model programs.

Sawhill and Welch^[17] elaborate on the same point as raised by Pages et al.^[27] The children in the early programs were born over half a century ago and much has changed since then, such as the mothers' educational level, the number of young children in out-of-home care, the ages of mothers when they had their first child, and the availability of a safety net. They expect that with higher-educated mothers, children are likely to also receive higher-quality care at home. In addition: "With more children enrolled in out-of-home care, those in pre-K may not be experiencing a very different level of care than the children to whom they are being compared in various studies, diminishing any possible treatment effects".

Duncan and Magnuson^[14] show that programs beginning before 1980 produced significantly larger effect sizes (0.33) than those that started later (0.16). One explanation is that the distinction between early education and other kinds of center-based childcare programs have blurred. Another is that family sizes have decreased, which means that the quality of parental care, as an alternative to early childhood education, is likely to have increased as the time parents are available has been spread across fewer children. Therefore, "whatever the internal validity of the research on high-quality programs, such as Perry and Abecedarian, their external validity remains an issue."

According to Bruhn and Emick,^[14] variations in effect sizes between the model programs and modern programs can be explained by, firstly, differences in the quality and intensity of the interventions, and, secondly, the expansion of preschool, which changes the nature of the counterfactual for children not in preschool. In contrast to the earlier, small-scale model programs, more recent studies focus on the effects of preschool attendance on real-world programs serving more representative large-scale samples.

But, is the answer to the question of considerably smaller recent effects not much simpler? This paper on Abecedarian and its companion paper on Perry^[20] have shown that both experiments and evaluations have been plagued with several methodological problems. Isn't it possible that because of these critical issues the effects of both programs have always been greatly overestimated? And that thus the "real" effects have always been small at best, right from the beginning?

4. Conclusion

The Abecedarian project combined a truly exceptional group of very young high-risk children and their mothers, with an extremely high-quality program, in terms of staff, content, intensity, and duration, and, above all, a staggering budget. In addition, though the project always is being presented as a true experiment, quite a number of methodological flaws seriously hamper the program's internal and external validity. Therefore, the only conclusion can be that it is unrealistic to qualify Abecedarian as a model program.

On the basis of their literature review, Duncan et al. (2022: 2)^[1] conclude: "This suggests that early childhood programs might play a significant role in helping children realize their potential in life. Nevertheless, our review of early childhood programs demonstrates that the evidence is mixed – some programs are successful in fostering lasting skill development, but many are not. We conclude that existing research on early childhood education falls short of sufficiently answering fundamental questions about what works for whom and why. A tighter link between theory, econometric methods and data is essential to compare and reconcile the mixed and sometimes conflicting empirical results across studies, and to understand when and why the impacts of home environment and pre-school interventions fadeout."

References

1. Greg Duncan; Ariel Kalil; Magne Mogstad; Mari Rege. Investing in Early Childhood Development in Preschool and at Home; National Bureau of Economic Research: Cambridge, MA, United States, 2022; pp. x.
2. OECD. Equity, diversity and inclusion in early childhood education and care; Organisation for Economic Co-Operation and Development (OECD): Paris, France, 2023; pp. x.
3. Early childhood education and care. European Commission . Retrieved 2024-1-10
4. James Heckman. Policies to Foster Human Capital; National Bureau of Economic Research: Cambridge, MA, United States, 1999; pp. x.

5. Geert Driessen. Early Childhood Education Intervention Programs in the Netherlands: Still Searching for Empirical Evidence. *Educ. Sci.* **2017**, 8, 3.
6. Geert Driessen. The evidence for the effectiveness of family- and center-based early childhood education programs. *Int. J. Pedagog. Innov. New Technol.* **2020**, 7, 106-115.
7. Pauline Slot. Early childhood education and care in the Netherlands: Quality, Curriculum, and Relations with Child Development; Utrecht University: Utrecht, 2014; pp. x.
8. Cadima, J.; Nata, G.; Barros, S.; Coelho, V.; Barat, C. Literature Review on Early Childhood Education and Care for Children under the Age of 3; OECD Publishing: Paris, 2020; pp. x.
9. European Commission. Key Data on Early Childhood Education and Care in Europe – 2019 Edition; Office of the European Union: Luxembourg, 2019; pp. x.
10. Ackerman, D. Reviewing the Research Base from 1985-2020: Infant and Toddler Child Care and At-Risk Children's School Readiness; Infant and Toddler Policy Research Center at NIEER: New Brunswick, NJ, 2021; pp. x.
11. Greg J. Duncan; Katherine Magnuson. Investing in Preschool Programs. *J. Econ. Perspect.* **2013**, 27, 109-132.
12. Ruben Fukkink; Lisanne Jilink; Ron Oostdam. A meta-analysis of the impact of early childhood interventions on the development of children in the Netherlands: an inconvenient truth?. *Eur. Early Child. Educ. Res. J.* **2016**, 25, 656-666.
13. CPB. Kansrijk Onderwijsbeleid; Centraal Planbureau: Den Haag, 2016; pp. x.
14. Bruhn, J.; Emick, E. Lottery evidence on the impact of preschool in the United States: A review and meta-analysis; MIT Department of Economics: Cambridge, MA, 2023; pp. x.
15. Driessen, G. The Many Facets of Educational Disadvantage. Policies, Interventions, Effects; Eliva Press: Chişinău, Moldova, 2022; pp. x.
16. Amy E. Lowenstein. Early Care and Education as Educational Panacea: What Do We Really Know About Its Effectiveness?. *Educ. Policy.* **2011**, 25, 92-114.
17. Sawhill, I.; Welch, M. Early Childhood Education. What Do We Know And What Should We Do? ; The Brookings Institution: Washington, DC, 2023; pp. x.
18. Schweinhart, L.; Montie, J.; Xiang, Z.; Barnett, W.; Belfield, C.; Nores, M. The High/Scope Perry Preschool Study Through Age 40. Summary, Conclusions, and Frequently Asked Questions; High/Scope Press: Ypsilanti, MI, 2005; pp. x.
19. Straight Talk on Evidence: Large randomized trial finds state pre-k program has adverse effects on academic achievement. Reform is needed to increase effectiveness. Straight Talk on Evidence. Retrieved 2024-1-11
20. Geert Driessen: The Perry High/Scope Preschool Program: A Critique. Encyclopedia. Retrieved 2024-1-10
21. Geert Driessen. The Perry Preschool Program: Model or Atypical? ; D. Pérez-Jorge, Eds.; Nova Science Publishers: Hauppauge, NY, 2024; pp. x.
22. Geert Driessen. De generaliseerbaarheid van een VVE-modelprogramma Heckmans dubieuze claims. *Orthopedagogiek: Onderzoek en Praktijk.* **2024**, 63, x.
23. Weikart, D. Preschool Programs: Preliminary Findings. *Journal of Special Education* . **1966**, 1, 163-181.
24. Barnett, W.; Masse, L. Comparative benefit–cost analysis of the Abecedarian program and its policy implications. *Economics of Education Review* . **2007**, 26, 113-125.
25. Campbell, F.; Ramey, C. Cognitive and School Outcomes for High-Risk African-American Students at Middle Adolescence: Positive Effects of Early Intervention. *American Educational Research Journal.* **1995**, 32, 743-772.
26. Kilburn, M.; Cannon, J.; Mattox, T.; Shaw, R. Programs That Work, from the Promising Practices Network on Children, Families and Communities; RAND Corporation: Santa Monica, CA, 2014; pp. x.
27. Remy Pages; Drew H. Bailey; Greg J. Duncan. The impacts of Abecedarian and Head Start on educational attainment: Reasoning about unobserved mechanisms from temporal patterns of indirect effects. *Early Child. Res. Q.* **2023**, 65, 261-274.
28. Craig T. Ramey; The Abecedarian Approach to Social, Educational, and Health Disparities. *Clin. Child Fam. Psychol. Rev.* **2018**, 21, 527-544.
29. Doolaard, S.; Leseman, P. Versterking van het fundament. Integrerende studie n.a.v. de opbrengsten van de onderzoekslijn; GION: Groningen, 2008; pp. x.
30. Twinkl: What is Cultural Capital (EYFS)?. Twinkl. Retrieved 2024-1-11

31. Geert Driessen. Ethnicity, forms of capital, and educational achievement. *International Review of Education* . **2001**, 47, 513-538.
32. Federal Interagency Forum on Child and Family Statistics. America's children: Key national indicators of well-being, 2023; U.S. Government Printing Office: U.S. Government Printing Office, 2023; pp. x.
33. G. Whitehurst: Testimony on Early Childhood Education to the House Committee on Education and the Workforce. Brookings. Retrieved 2024-1-10
34. Geert Driessen. Language choice and language proficiency of immigrant parents and children. *Journal of Education Research* . **2016**, 10, 97-108.
35. Ramey, C.; Ramey, S. Early Childhood Education that Promotes Lifelong Learning, Health, and Social Well-being: The Abecedarian Project and Its Replications. *Medical Research Archives* . **2023**, 11, 1-17.
36. Whitaker, A.; Burchinal, M.; Jenkins, J.; Bailey, D.; Watts, T.; Duncan, G.; Hart, E.; Peisner-Feinberg, E. Why are Preschool Programs Becoming Less Effective?; Brown University: Providence, RI, 2023; pp. x.
37. Jorge Luis García; James J. Heckman; Duncan Ermini Leaf; María José Prados; Quantifying the Life-Cycle Benefits of an Influential Early-Childhood Program. *J. Politi- Econ.* **2020**, 128, 2502-2541.
38. Farran, D.; Lipsey, M.: Misrepresented evidence doesn't serve pre-K programs well. Brookings. Retrieved 2024-1-10
39. Whitehurst, G.: Does Pre-k Work? It Depends How Picky You Are. Brookings. Retrieved 2024-1-10
40. Whitehurst, G.: Rigorous preschool research illuminates policy (and why the Heckman Equation may not compute). Brookings. Retrieved 2024-1-10
41. Schaeffer, A.: Pre-K Pushers Peddling Patent Prevarications. The Cato Institute. Retrieved 2024-1-10
42. Besharov, D.; Germanis, P.; Higney, C.; Call, D. 2 The Abecedarian Project; Besharov, D., P. Germanis, C. Higney, and D. Call, Eds.; University of Maryland: Washington, DC, 2011; pp. x.
43. Hu, M.: What happened with the Abecedarian study ? IQ-malleability theories in danger. Human Varieties. Retrieved 2024-1-10
44. Spitz, H. Does the Carolina Abecedarian Early Intervention Project Prevent Sociocultural Mental Retardation?. *Intelligence*. **1992**, 16, 225-237.
45. Olsen, D.; Snell, L. Assessing Proposals for Preschool and Kindergarten: Essential Information for Parents, Taxpayers and Policymakers; Reason Public Policy Institute: Los Angeles, CA, 2006; pp. x.
46. Spitz, H. When Prophecy Fails: On Ramey's Response to Spitz's Critique of the Abecedarian Project. *Intelligence*. **1993**, 17, 17-23.
47. Spitz, H. Some Questions About the Results of the Abecedarian Early Intervention Project Cited by the APA Task Force on Intelligence. *American Psychologist* . **1997**, 52, 72.
48. Ramey, C. Rejoinder to Spitz's Critique of the Abecedarian Experiment. *Intelligence*. **1993**, 17, 25-30.
49. Ponniah, R.: The Abecedarian Project. Office of Early Childhood Education. Retrieved 2024-1-10
50. Heckman, J.; Moon, S.; Pinto, R.; Savelyev, P.; Yavitz, A. Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quant Econom*. **2010**, 1, 1-46.
51. Social Programs That Work Review: Evidence Summary for the Abecedarian Project. Social Programs that Work. Retrieved 2024-1-10
52. Anderson, M. Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*. **2008**, 103, 1481-1495.
53. Slavin, R.; Smith, D. The Relationship Between Sample Sizes and Effect Sizes in Systematic Reviews in Education. *Educ. Evaluation Policy Anal.* **2009**, 31, 500-506.
54. Grace-Martin, K.: How big of a sample size do you need for factor analysis? . The Analysis Factor. Retrieved 2024-1-10
55. Kyriazos, T. Applied Psychometrics: Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General. *Psychology*. **2018**, 9, 2207-2230.
56. Bruno, E.; Iruka, I. Reexamining the Carolina Abecedarian Project using an antiracist perspective: Implications for early care and education research. *Early Childhood Research Quarterly* . **2022**, 58, 165-176.

