

# evidence for the benefits of state prekindergarten programs: myth & misrepresentation

Dale C. Farran & Mark W. Lipsey

## abstract

In 2014, New York City launched the historic Pre-K for All program, which massively expanded early education for children in the city. The state followed suit with a budgetary pledge of some \$1.5 billion over 5 years to implement full-day universal prekindergarten (pre-K) statewide. Many other states have implemented or expanded state-funded pre-K programs in the last decade, encouraged by claims about expected benefits. However, there is remarkably little scientifically rigorous evidence for these benefits. Claims of pre-K success rely largely on small, experimental, model programs run 50 or more years ago—programs that bear little resemblance to current pre-K implementations. Evidence for program effectiveness is crucial given the current interest in and expansion of state-funded pre-K programs. This review of the evidence raises serious questions about the presumed benefits of state pre-K programs.

## Evidence for the benefits of state prekindergarten programs: Myth & misrepresentation

Dale C. Farran & Mark W. Lipsey

**abstract.** In 2014, New York City launched the historic Pre-K for All program, which massively expanded early education for children in the city. The state followed suit with a budgetary pledge of some \$1.5 billion over 5 years to implement full-day universal prekindergarten (pre-K) statewide. Many other states have implemented or expanded state-funded pre-K programs in the last decade, encouraged by claims about expected benefits. However, there is remarkably little scientifically rigorous evidence for these benefits. Claims of pre-K success rely largely on small, experimental, model programs run 50 or more years ago—programs that bear little resemblance to current pre-K implementations. Evidence for program effectiveness is crucial given the current interest in and expansion of state-funded pre-K programs. This review of the evidence raises serious questions about the presumed benefits of state pre-K programs.

The momentum to create or expand state-funded prekindergarten (pre-K) programs increased dramatically in the mid-2000s, encouraged in large part by a perspective article published in *Science* by James J. Heckman that called for investments in early childhood education for disadvantaged children.<sup>1</sup> This charge was quickly adopted by education advocacy groups such as ReadyNation. However, Heckman's conclusions about the benefits of such investments were based on research from the 1960s and 1970s on the effects of two well-funded experimental programs: the Perry

Preschool Project and the Abecedarian Project. The Perry Preschool Project was a 2-year intensive intervention that began when children were 3 years old and required substantial parental involvement—researchers visited parents in their homes once a week for an hour and a half. The Abecedarian Project began when children were 6 weeks old and lasted seamlessly until they entered kindergarten. The care covered 8–10 hours a day for 50 weeks of the year. These programs, because of their intense interventions and high cost, have been dubbed “Cadillac programs,” and few feel that they are sustainable on a national or state stage.

More recent support for statewide early education programs came in the form of a 2014 White House report titled *The Economics of Early Childhood*

---

Farran, D. C., & Lipsey, M. W. (2016). Evidence for the benefits of state prekindergarten programs: Myth & misrepresentation. *Behavioral Science & Policy*, 2(1), pp. 9–18.

*Investments*.<sup>2</sup> This report referred to a “deep” research base supporting pre-K and cited the same 1960s and 1970s studies that Heckman referenced. The report went on to state that “dozens of other programs have been rigorously examined since the 1960s” and asserted that high-quality early education will narrow the achievement gap, boost adult earnings, and result in savings of \$8.60 for every \$1 spent on the program.

Despite various correction efforts by independent watchdogs<sup>3,4</sup>—such as the *Washington Post*’s Fact Checker,<sup>5</sup> which called into question the administration’s pre-K cost–benefit claims—belief in the long-term payoff of pre-K programs is so well entrenched in the thinking of policymakers that the state sponsors of a relatively new program, the Great Start Readiness Program in Michigan, were willing to claim long-term, adult benefits even though the program has not existed long enough for such outcomes to be observed.<sup>6</sup> Similarly, in a Texas-commissioned report, researchers Robert Pianta and Catherine Wolcott of the University of Virginia<sup>7</sup> assured policymakers that they need not spend money on Cadillac programs with unsustainable costs because other states have demonstrated that lower cost programs can achieve the same results. The programs they cited do not have adequate research to justify these claims, but they are cheaper.

### Claims for Effectiveness Must Align with Reality for Sustained, Appropriate Long-Term Investments in Pre-K Education

The problem that states and society at large face is a genuine one: Children from low-income backgrounds underachieve in school compared with children from higher income families, and this poverty gap now eclipses racial achievement gaps.<sup>8</sup> The presumption that poor children can be made more ready for school by kindergarten entry in a way that will then propel them forward to achievement equal to that of their more privileged peers has been present in this country for 50 years, at least since the creation of Head Start.<sup>9</sup> In this article, we do not question whether this is an important goal; it is.

Our concern is the substantial discrepancy between the actual evidence and the expected benefits of contemporary scaled-up, statewide pre-K programs that have been adopted on the basis of strong claims about how compelling the supporting research is. Indeed,

long-term funding of early education is in considerable danger if the public is overpromised on what it can and should expect from pre-K programs. When a governmental study on the effects of the Head Start early education program was reported in 2010, the results were so dismal that *TIME Magazine* called for an end to Head Start.<sup>10</sup>

In this article, we place the recently reported results<sup>11</sup> (extending to Grade 3) from our study of Tennessee’s pre-K program within the spectrum of studies—of varying methodological quality—on similar state-wide programs. First, however, we challenge anyone to define with any specificity what a statewide pre-K program actually is or should be. Next, we consider the outcomes such programs are expected to bring about and examine the research evidence on state pre-K programs. Last, we highlight the difficulties associated with scaling up an effective program even when the intended outcomes are well defined.

### Highly Divergent State Pre-K Programs

A quote from a Pew Charitable Trusts analysis sums things up: “While there’s a growing consensus on the value of preschool, states disagree on where the programs should be based, who should run them, or how the government should support them.”<sup>12</sup> States are doing remarkably different things under the rubric of pre-K. And this situation is continually in flux—many states have recently passed legislation substantially changing or expanding their programs. Nonetheless, most, although not all, state programs more or less follow a public school model in setting parameters, meaning that pre-K classrooms live in public elementary schools, a full day is a standard school day (typically 6 hours), and a full year is a standard 9-month school year.

States differ on whether lead teachers in pre-K classrooms should be licensed and/or have bachelor’s degrees. Class size requirements differ as well; for example, New Jersey limits class size to 15 students, but Texas has no specified limit. Most states set the adult-to-child ratio at 1 to 10, but in Texas, it may go as high as 1 to 22, a stark difference.

And whereas most states target their programs to children from low-income families, financial eligibility cutoffs vary widely. For example, many states use the eligibility criteria for the free and reduced-price lunch

(FRPL) programs—135% and 185% of the federal poverty level, respectively. However, North Carolina uses 75% of the state’s median income as the cutoff, and Michigan sets it at 250% of the federal poverty level.<sup>13</sup> In the pre-K expansion grants newly funded by the federal government, the eligibility requirement is 200% of the federal poverty level.<sup>14</sup>

However, state pre-K programs are quite similar in one regard: No state requires implementation of a program with components that match the Perry Preschool Project or Abecedarian Project models, even though both have shown long-term benefits for participants. More commonly, programs are required to implement an early childhood curriculum, typically on a state approved list, but those curricula lack distinct evidence of effectiveness and do not share a common vision. In New Jersey, for example, programs are limited to choosing among Tools of the Mind, Creative Curriculum, Curiosity Corner, and HighScope—curricula that differ greatly from each other in content and approach. Research on three of these has found no additional benefits compared with generic practice, according to reviews by the What Works Clearinghouse,<sup>15</sup> and the fourth (HighScope) has not been reviewed (for more information on the What Works Clearinghouse, see the online Supplemental Material).

Very few states invest in monitoring the quality of pre-K programs once funds are awarded to school systems. Monitoring is left up to local school systems. A few states, such as Tennessee and Louisiana, require self-assessments using measurement tools such as the Early Childhood Environmental Rating Scale, but they do not regulate the rigor with which those assessments are done or how the results are used. Only rarely do states require that their public school pre-K programs be rated by the state’s Quality Rating and Improvement System (for more on the Quality Rating and Improvement System, see the online Supplemental Material), even though many states do require such an evaluation of community child-care programs serving children of the same age.

### Diverging Goals for Pre-K

Not only do pre-K programs differ across states, but also the declared goals of the programs vary—even within a state—and too often those goals are ambiguous. If there were statewide consensus about the

goals for pre-K, even with disagreement on the best way to attain them, at least relevant outcomes could be identified and progress toward achieving them could be evaluated. The goals for state pre-K programs are generally more implicit than explicit and include the following variations.

*School readiness.* Kindergarten readiness frequently appears as a primary objective of public pre-K—especially in the most recent request for proposals for pre-K expansion from the Obama administration.<sup>14</sup> Yet school readiness is often vaguely defined; it usually means some mix of literacy and numeracy skills, proper school behavior, and perhaps motor development skills and health objectives. Strictly speaking, school readiness as a goal includes no assumption of sustained effects beyond kindergarten entry, although the expectation certainly is that reducing the gap at school entry for disadvantaged children will allow them to progress more effectively in later grades.

*School achievement.* This goal focuses on learned academic content and skills that go beyond simple school readiness and instead are manifest most clearly in performance on achievement tests, grades, and grade retention. The general research literature on the effects of pre-K does not provide much encouragement for the expectation that gains in academic achievement will be sustained for very long. Any beneficial effects have typically been found to diminish well before high school, even for the Perry Preschool Project and the Abecedarian Project.<sup>16</sup> Nonetheless, sponsors and stakeholders, including the White House, expect that state pre-K programs will help close school achievement gaps for minority and economically disadvantaged children in the long run. As such, whatever initial positive effects pre-K may have on achievement are expected by policymakers to carry through into the later grades.

*Behavioral outcomes.* These are the “noncognitive” outcomes, such as attendance, disciplinary infractions, graduation, employment, and criminal behavior. Many proponents cite beneficial effects on these behaviors as a primary justification for scaling up pre-K programs. These behaviors are not directly reflective of academic achievement but may speak to broader effects on children’s persistence and goal orientation. Expectations of such effects are based on the classic

longitudinal studies of the Abecedarian Project and the Perry Preschool Project that found long-term behavioral benefits in some areas, although not consistently on the same outcomes across the two programs. The long-term benefits found in these early programs are also the primary basis for the claims of the cost effectiveness of pre-K.<sup>17</sup>

## Outcome Studies: Limited Research and Weak Methods

Relatively few studies of the effects of state-funded, scaled-up pre-K programs have been reported, and they rarely appear in peer-reviewed outlets. We find the most striking features of these studies to be weak methods and limited outcome assessments. Here, we summarize the study designs, outcome variables, and findings from this research—first for effects found at the end of the pre-K year and then for sustained effects through the school years following pre-K.

### Short-Term Effects

The methodological rigor of research on the immediate effects of state pre-K programs is not generally strong. Nonetheless, there is an overall pattern of positive effect estimates of sufficient magnitude to suggest likely benefits, particularly on achievement outcomes. Moreover, that conclusion is consistent with the broader research literature on the short-term cognitive effects of early childhood education.<sup>18</sup>

*Randomized studies.* The most methodologically rigorous design for studying program effects is a randomized control trial (RCT). In an RCT, participants are assigned to treatment and control groups via a chance process that results in no systematic differences between the groups at the beginning of the study. The only RCT of a state pre-K program is the one we have conducted on the Tennessee program, and it is not yet complete.<sup>19</sup> The study was done in 58 school-based pre-K programs that agreed to admit children in random order until the available seats were filled; the remaining applicants served as the control group. For a subsample of one-third of the total sample, outcome data were collected annually to track program effects through the third grade, when the state achievement tests are administered. The outcome measures included

Woodcock-Johnson III achievement tests for literacy, language, and math as well as teacher ratings of classroom behavior. The pre-K participants and nonparticipants in this comparison were virtually identical on baseline variables such as achievement pretests, demographics, and family characteristics.

We found statistically significant positive effects for the pre-K participants at the end of the pre-K year on all of the achievement measures except one (oral comprehension), as well as on the composite achievement score. In addition, benefits were observed on teacher ratings obtained at kindergarten entry on three measures: preparedness for grade, work-related skills, and social behavior.

*Regression discontinuity design studies.* The short-term effects of state pre-K programs have most often been studied using a research design that takes advantage of age thresholds for enrollment in state pre-K programs. Children entering kindergarten who attended a pre-K program in the previous year (treatment group) are assessed in the fall of their kindergarten year, and their results are compared with the results of children who, because of their later birthdays, are just entering the pre-K program at that time (control group). With statistical adjustments for the age difference, estimates of the pre-K effects can be derived. This method is called an *age-cutoff regression discontinuity design* (RDD); because the treatment and control groups are not created via random assignment, it is considered a quasi-experimental design.

The RDD is widely acknowledged to be one of the more rigorous quasi-experimental designs.<sup>20</sup> What those using this design typically do not acknowledge is that the age-cutoff version is not a true RDD but, rather, an approximation that is vulnerable to biases that can compromise the effect estimates.<sup>21</sup> For example, children from economically disadvantaged families who attend pre-K are more likely to move<sup>22</sup> and thus be omitted from outcome assessments conducted at the beginning of kindergarten the next year. Also, the children in the control group who are assessed at the beginning of pre-K were 3 years old the previous year. In that regard, they are not comparable to the pre-K participants assessed at the beginning of kindergarten who were 4 years old the previous year. A year lived as a 3-year-old provides different background experiences than a year in the life of a comparable 4-year-old would have.

The majority of pre-K age-cutoff RDD studies have been conducted by researchers associated with the National Institute for Early Education Research and include state-funded pre-K programs in Arkansas,<sup>23</sup> California,<sup>24</sup> Michigan,<sup>25</sup> New Jersey,<sup>26</sup> New Mexico,<sup>27</sup> Oklahoma,<sup>28</sup> South Carolina,<sup>29</sup> and West Virginia.<sup>20,30</sup> The validity of the results reported for these studies is difficult to assess—they provide little detail about the equivalence of the children in the successive pre-K cohorts that are compared, the data used in the analyses, or the nature of the analyses. Other age-cutoff studies conducted in Georgia,<sup>31</sup> North Carolina,<sup>32</sup> and Tennessee,<sup>33</sup> as well as the earliest of such studies done in Oklahoma,<sup>34</sup> have used statistical controls to better match the comparison groups on such characteristics as gender, race, and English language proficiency.

The outcomes measured in the age-cutoff RDD studies have almost exclusively been cognitive measures, for example, for emerging literacy, language, and math knowledge. Overall, the results on such measures show wide variation but almost always indicate positive pre-K effects. Only the Georgia study included behavioral outcomes, in particular for social skills, problem behavior, and social awareness, and positive effects were found on those as well.

*Post hoc matched studies.* A number of studies use a design that compares the results of outcome measures taken at the beginning of the kindergarten year by children who participated in the state pre-K program the previous year with those for children who did not attend the program. These children are typically matched on a few demographic variables, such as gender, race or ethnicity, age, and FRPL eligibility. However, these post hoc matched studies lack critical information about how comparable the children were on such factors as initial cognitive skills, relevant behavioral dispositions, and family background prior to their differential exposure to pre-K.

Without “before” measures, we cannot know if differences in the outcomes were a result of pre-K participation or preexisting differences between the groups. An inherent difference between these groups is that the parents of pre-K participants chose to send their children to a voluntary educational program, whereas the parents of nonparticipants did not. The greater motivation or ability of the parents who enrolled their children in pre-K, compared with that of the parents who did

not, likely reflects family differences that could easily be related to children’s academic performance.

Post hoc matched studies have been reported for Arkansas,<sup>23</sup> California,<sup>24</sup> Michigan,<sup>25</sup> New Jersey,<sup>35,36</sup> and Virginia.<sup>37</sup> Virtually all of the effects reported in these studies favor the pre-K participants, although they are generally smaller than those found in the RDD studies, even in studies of the same state programs conducted by the same researchers (e.g., in Arkansas and New Jersey). These studies are easy to do, but this posttest-only design is weak and has long been recognized as a preexperimental design that is incapable of supporting causal inferences about intervention effects.<sup>38</sup>

Albeit based largely on methodically weak study designs, rather consistent evidence of short-term benefits has been found for state pre-K programs, as described above. Moreover, that conclusion is consistent with the broader research literature on the short-term cognitive effects of early childhood education.<sup>18</sup> However, the available research on the extent to which those positive effects last past the end of the pre-K year is far shakier.

## Long-Term Effects

*Randomized studies.* Our RCT study in Tennessee is, again, the only one of its kind that reports longer term outcomes of pre-K program participation.<sup>21</sup> We have followed the subsample of children from our study through third grade and found that, despite the positive achievement gains made by the pre-K participants relative to the gains of the nonparticipants during the pre-K year, there was no longer any difference in those outcomes by the end of kindergarten. Moreover, by the end of third grade, the control group of children who did not attend pre-K actually scored higher than the pre-K participants did on some achievement measures. That is, we observed an acceleration in achievement for pre-K nonparticipants such that their performance soon caught up with, and in some cases surpassed, that of the pre-K participants. By the end of first grade, teacher ratings of preparedness for grade, work-related skills, social behavior, and related noncognitive outcomes similarly showed no differences between the pre-K participants and nonparticipants.

*Post hoc matched studies.* The largest number of studies of longer term effects of state pre-K programs

use post hoc matched designs to compare outcomes for former pre-K participants and nonparticipants some years later. These designs have the same sources of potential bias described earlier, plus the added possibility of differential attrition from the sample as time goes on. Such comparisons have been reported for state pre-K programs in Arkansas,<sup>23,39</sup> Colorado,<sup>40</sup> Louisiana,<sup>41</sup> Michigan,<sup>25,42</sup> North Carolina,<sup>43</sup> New Jersey,<sup>35,44</sup> South Carolina,<sup>45</sup> Tennessee,<sup>46</sup> Texas,<sup>47,48</sup> Virginia,<sup>37</sup> and Washington.<sup>49</sup> The outcomes examined—mostly achievement test scores and grade retention—are almost exclusively drawn from state databases.

Overall, the results vary considerably in magnitude and statistical significance, or how likely the findings are due to chance, but they overwhelmingly favor the children who attended the state pre-K programs. Further, those differences are typically described in the language of causal attribution—for example, as *pre-K effects*—when that interpretation is not justified as a conclusion from such methodologically deficient research designs.

Consider, for example, the report prepared for the Louisiana Department of Education by a team of researchers from the University of Louisiana, the University of Alabama, and Georgetown University.<sup>50</sup> The state achievement test scores of third through eighth grade children eligible for the FRPL programs who had attended the state pre-K program were compared with the scores of FRPL-eligible students in the same grades who had not attended the state pre-K program. The higher scores for pre-K participants were interpreted as proof of pre-K having a “positive impact” on achievement. Yet no data were presented to establish that economically disadvantaged pre-K participants and nonparticipants were equivalent at the beginning of the pre-K year or, indeed, in any year on other factors that might have influenced their test performance.

Some of the post hoc matched studies obtained more data about student characteristics than did the Louisiana study, and those data were then used for matching or statistical control. However, only a few of these studies went beyond FRPL status, age, gender, and race or ethnicity, and none included any true baseline data beyond static demographics. Although these studies have been cited as evidence of sustained effects from state pre-K programs, they fall well below the most minimal methodological standards required to support even a tentative claim of that sort.

*Difference-in-difference studies.* Another method that researchers have used to study the effects of state pre-K programs on school achievement is known as a *difference-in-difference (DD) design*. These studies examine the differences in state- or county-level student scores on measures administered after a pre-K program is rolled out and compare them with differences in test scores seen over a comparable period for another area in which there was no analogous pre-K implementation or expansion. The challenge for researchers using this design is to develop statistical analyses that isolate the difference made in the target outcomes by pre-K implementation from all of the other influential factors co-occurring over that same time period that are not necessarily also occurring in the same way in the comparison area.

For example, a DD design was used to investigate the effects of the Georgia universal pre-K program that grew from participation rates of 14% in 1995 to 55% in 2008.<sup>51</sup> The analysis included statistical controls for a range of factors other than the introduction of pre-K that could have influenced student outcomes. Initial analyses indicated positive pre-K effects on achievement scores from the National Assessment of Educational Progress, but further analyses exploring control group variants and different statistical models did not yield a robust conclusion. Similar sensitivity to including different kinds of information in the analyses and the use of alternate statistical models was found in the results for another DD study of the Georgia program as well as for the Oklahoma program.<sup>52</sup>

More robust findings emerged from a DD study of two early childhood programs in North Carolina, of which one (More at Four) was a pre-K program.<sup>53</sup> This study focused on differences across counties in the timing and magnitude of funding for these programs and was distinctive in at least two ways. First, a high proportion of the pre-K programs were not school based. Second, by drawing on birth records, researchers found an unusually rich set of statistical control variables for the analysis. Positive effects were found on third grade state achievement test scores for both reading and math in counties implementing pre-K early. The strong statistical control notwithstanding, the authors acknowledged that the validity of the findings rested on the assumption that there were no uncontrolled factors capable of influencing test scores that

were coincident with the increases in funding for the pre-K program.

The difficulty of drawing firm conclusions from DD analyses in the dynamic context of state pre-K expansion is further illustrated by an ambitious study conducted by Rosinsky.<sup>54</sup> She compared the 2007, 2009, and 2011 fourth grade National Assessment of Educational Progress math scores with program enrollment of 4-year-olds 6 years previously in Head Start, state-funded pre-K, or special education preschools across multiple states. Surprisingly, she found a negative association between math scores and higher enrollment in state-funded pre-K.

Although DD studies in general are methodologically stronger than the post hoc matching studies, the inconsistency in their findings makes it difficult to draw confident conclusions from them about the long-term effects of state pre-K programs.

*Concluding observations about long-term pre-K effects.* Overall, the methodological quality of research on the effects of state pre-K programs is poor, showing little improvement since a 2001 review<sup>55</sup> by Gilliam and Zigler that concluded that the knowledge base was dangerously weak just as states were ramping up their programs. The evidence for favorable immediate effects of pre-K participation is consistent across a range of research designs, despite their respective limitations. By contrast, not only is there little consistency in the findings on long-term effects, but few studies address that issue with strong research designs.

The limited range of outcome variables represented across all of the studies of state pre-K programs is also noteworthy. The measures used are almost exclusively indices of academic achievement, primarily achievement tests. Very few studies have examined behavioral outcomes, despite the fact that those are the pre-K effects that are supposed to be sustained in the long term and generate the cost savings that have been claimed for pre-K. Investigation of adult behavioral outcomes may be beyond the reach of time-limited state pre-K studies, but that is not the case for potential bridging variables that may connect pre-K experiences with adult outcomes, for example, self-regulation, engagement with school, and grit. Outcomes of this sort are strikingly absent from the research on state pre-K effects.

## Problems in Taking Programs to Scale

The research community has developed and validated many promising programs and practices, but few of these have been taken to scale while maintaining the same level of effectiveness.<sup>56</sup> In the case of pre-K programs implemented statewide, the situation is especially problematic. What is being scaled up is not a well-defined practice but rather an idea, a concept—the notion that some kind of school-like intervention provided to poor children prior to kindergarten entry will change their developmental trajectories in positive ways that will last well into adulthood. The grounding for this expectation comes from research conducted on the Cadillac programs implemented half a century ago that bear little resemblance to today's state programs.

In this regard, the recent scale-up effort for Response to Intervention (RTI; see the online Supplemental Material for more information) is instructive. RTI was developed and favorably evaluated in small experimental studies of a closely monitored small-group reading program for students in early grades. On that evidence, the program was written into the reauthorization of the Individuals with Disabilities Education Act in 2004. However, RTI is essentially the set of concepts that guided the original experimental intervention rather than a structured program, and school districts were free to implement it however they saw fit. A multistate evaluation of RTI implemented at scale actually found negative effects.<sup>57</sup>

The pre-K concept is much like the situation with RTI: There is no coherent vision for what the program should look like other than it should somehow be like the Perry Preschool Project and the Abecedarian Project—but not so much so that it would require as much time, effort, staff, or money. Viewed realistically, policymakers should not expect a scaled-up version of this rather open-ended concept to have the dramatic, sustained effects that are forecast by advocates and proponents.<sup>58</sup> As the research we have reviewed above indicates, little evidence supports the expectation of such effects.

## Conclusion

The importance of helping young children from impoverished backgrounds advance in school learning and behavior cannot be overstated. The idea that a year

of pre-K can have effects that will endure well into adulthood is appealing to policymakers, school administrators, businessmen,<sup>59</sup> and law enforcement officials.<sup>60</sup> If the achievement gap for poor children can be closed and life outcomes improved with a year of preschool, then state policymakers would be foolish not to implement programs making pre-K accessible to everyone. This idea and the children's needs are so very compelling that states have rushed to scale up pre-K programs without much attention to the question of how to design and support those programs so that they are effective.

Much of the research reviewed above was conducted or commissioned by the state departments of education that administer the respective pre-K programs. That the resulting reports do not mention the methodological weaknesses of these studies while they emphasize the positive effects the studies are purported to demonstrate raises a question about the purpose of this research. Many reports have the appearance of supporting state policies that have already been adopted. If the report writers adopted a more critical approach to describing such studies, the reports policymakers base their decisions on would not only be more forthright about the methodological limitations of the studies and less rosy about their conclusions, but they would also acknowledge the considerable difficulty of implementing an effective program at scale and avoid claiming or implying that scale-up had been successfully accomplished.

It is facile to call for more and better research, but the problems outlined here are fundamental and must be addressed. Even if good, long-lasting effects were found for some pre-K programs, researchers would be hard-pressed to specify which features of those programs were responsible for that success so they could be emulated elsewhere. What is lacking is a distinct and plausible *theory of change* that would articulate the expectations for what should happen in pre-K and how that would affect the desired immediate and long-term outcomes. Such a theory would help identify the outcomes states should monitor and the adjustments that might be tried if those outcomes are not satisfactory. Moreover, the expectations for state pre-K emphasize long-term effects, but policymakers are not prepared to wait 15 to 20 years to find out if their programs are successful. Researchers need to identify and validate the proximal outcomes in a theory

of change that are the precursors to the desired long-term outcomes.

Children are not well served by a perpetuation of magical thinking about the likelihood of profound effects resulting from poorly defined state-run pre-K programs. Moreover, researchers should not be aiding and abetting that thinking with weak and misleading research presented without acknowledgement of its serious limitations. Viewed with a critical eye, the currently available research raises real questions about whether most state pre-K programs do anything more than boost 4-year-olds' academic cognitive skills to where they would be by the end of kindergarten anyway.

### author affiliation

Farran, Department of Teaching and Learning, Vanderbilt University; Lipsey, Department of Human and Organizational Development, Vanderbilt University. Corresponding author's e-mail: Dale.Farran@Vanderbilt.edu

### supplemental material

- <https://behavioralpolicy.org/journal/>
- Supplemental Text

### References

1. Heckman, J. J. (2006, June 30). Skill formation and the economics of investing in disadvantaged children. *Science*, 312, 1900–1902. Retrieved from <http://www.jstor.org/stable/3846426>
2. Executive Office of the President of the United States. (2014, December). *The economics of early childhood investments*. Washington, DC: Author. Retrieved from [https://www.whitehouse.gov/sites/default/files/docs/early\\_childhood\\_report1.pdf](https://www.whitehouse.gov/sites/default/files/docs/early_childhood_report1.pdf)
3. Farley, R. (2013, February 20). *Obama's preschool stretch*. Retrieved from <http://www.factcheck.org/2013/02/obamas-preschool-stretch/>
4. Murray, C. (2013, February 20). *The shaky science behind Obama's universal pre-K*. Retrieved from <http://www.bloombergview.com/articles/2013-02-21/the-shaky-science-behind-obama-s-universal-pre-k>
5. Lee, M. Y. H. (2015, April 20). Obama's claim that every dollar spent on pre-kindergarten earns '\$7 back.' *Washington Post*. Retrieved from <http://www.bloombergview.com/articles/2013-02-21/the-shaky-science-behind-obama-s-universal-pre-k>
6. Michigan Department of Education. (n.d.). *Great Start Readiness Program*. Retrieved from [http://www.michigan.gov/documents/mde/GSRP-Overview\\_410757\\_7.pdf](http://www.michigan.gov/documents/mde/GSRP-Overview_410757_7.pdf)
7. Pianta, R., & Wolcott, C. (2014). *Pre-kindergarten for the modern age: A scalable, affordable, high-quality plan for*

- Texas. Austin, TX: Raise Your Hand Texas. Retrieved from [http://www.raiseyourhandtexas.org/wp-content/uploads/2015/02/RaiseYourHandTexas\\_PreK\\_ResearchReport\\_Aug2015.pdf](http://www.raiseyourhandtexas.org/wp-content/uploads/2015/02/RaiseYourHandTexas_PreK_ResearchReport_Aug2015.pdf)
8. Reardon, S. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. Duncan & R. Murnane (Eds.), *Whither opportunity: Rising inequality, schools, and children's life chances* (pp. 91–116). New York, NY: Russell Sage Foundation.
  9. Farran, D. C. (2007). Is education the way out of poverty? A reflection on the 40th anniversary of Head Start (with commentaries by James King and Bernard L. Charles). *Monographs of the Center for Research on Children's Development and Learning*, 3.
  10. Klein, J. (2011, July 7). Time to ax public programs that don't yield results. *TIME Magazine*. Retrieved from <http://content.time.com/time/nation/article/0,8599,2081778,00.html>
  11. Farran, D., & Lipsey, M. (2015, October 8). Expectations of sustained effects from scaled up pre-K: Challenges from the Tennessee study. *Evidence Speaks Reports*, 1(3). Washington, DC: Brookings Institution.
  12. Quinton, S. (2015, September 4). States agree on need for 'preschool,' differ on definition [Blog post]. Retrieved from <http://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2015/09/04/states-agree-on-need-for-preschool-differ-on-definition>
  13. Connors-Tadros, L., Brown, K., & Schilder, D. (2015). *Examples of state guidance to determine and verify income eligibility for prekindergarten programs*. Rutgers, NJ: Center for Enhancing Early Learning Outcomes.
  14. U.S. Department of Education & U.S. Department of Health and Human Services. (2014, August 18). *Applications for new awards; Preschool development grants—development grants*. *Federal Register*, 79(159), 48853–48872.
  15. What Works Clearinghouse. (2015). *Find what works/early childhood interventions*. Washington, DC: U.S. Department of Education, Institute of Education Sciences. Retrieved November 7, 2015, from [ies.ed.gov/ncee/wwc/findwhatworks.aspx#Social-emotionaldevelopment](http://ies.ed.gov/ncee/wwc/findwhatworks.aspx#Social-emotionaldevelopment)
  16. Duncan, G., & Magnuson, K. (2013). Investing in preschool programs. *Journal of Economic Perspectives*, 27, 109–131. doi:10.1257/jep.27.2.109
  17. Heckman, J., Moon, H., Pinto, R., Savelyev, P., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94, 114–128. doi:10.1016/j.jpubeco.2009.11.001
  18. Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, 112, 579–620.
  19. Lipsey, M. W., Farran, D. C., & Hofer, K. (2015). *A randomized control trial of the effects of a statewide voluntary prekindergarten program on children's skills and behaviors through third grade*. Retrieved from Peabody Research Institute website: [http://peabody.vanderbilt.edu/research/pri/VPKthrough3rd\\_final\\_withcover.pdf](http://peabody.vanderbilt.edu/research/pri/VPKthrough3rd_final_withcover.pdf)
  20. Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management*, 27, 122–154. doi:10.1002/pam.2031
  21. Lipsey, M., Weiland, C., Yoshikawa, H., Wilson, S., & Hofer, K. (2015). Prekindergarten age-cutoff regression-discontinuity design: Methodological issues and implications for application. *Educational Evaluation and Policy Analysis*, 37, 296–313. doi:10.3102/0162373714547266
  22. Coulton, C., Theodos, B., & Turner, M. (2012). Residential mobility and neighborhood change: Real neighborhoods under the microscope. *Cityscape: A Journal of Policy Development and Research*, 14, 55–89.
  23. Jung, K., Barnett, W. S., Hustedt, J. T., & Francis, J. (2013). *Longitudinal effects of the Arkansas Better Chance Program: Findings from first grade through fourth grade*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.
  24. Barnett, W. S., Howes, C., & Jung, K. (2009). *California's state preschool program: Quality and effects on children's cognitive abilities at kindergarten entry* [Report to the California Children and Families Commission]. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.
  25. Xiang, Z., & Schweinhart, L. J. (2002). *Effects five years later: The Michigan School Readiness Program evaluation through age 10* [Report to the Michigan State Board of Education]. Retrieved from HighScope Educational Research Foundation website: <http://www.highscope.org/file/Research/Effects%205%20Years%20Later.pdf>
  26. Lamy, C., Barnett, W. S., & Jung, K. (2005). *The effects of New Jersey's Abbott Preschool Program on young children's school readiness*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.
  27. Hustedt, J. T., Barnett, W. S., Jung, K., & Friedman, A. H. (2010). *The New Mexico pre-K evaluation: Impacts from the fourth year (2008–2009) of New Mexico's state-funded pre-K program*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.
  28. Lamy, C., Barnett, W. S., & Jung, K. (2005). *The effects of Oklahoma's early childhood four-year-old program on young children's school readiness*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.
  29. Lamy, C., Barnett, W. S., & Jung, K. (2005). *The effects of South Carolina's early childhood programs on young children's school readiness*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.
  30. Lamy, C., Barnett, W. S., & Jung, K. (2005). *The effects of West Virginia's early education program on young children's school readiness*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.
  31. Peisner-Feinberg, E. S., Schaaf, J. M., LaForett, D. R., Hildebrandt, L. M., & Sideris, J. (2014). *Effects of Georgia's pre-K program on children's school readiness skills: Findings from the 2012–2013 evaluation study*. Chapel Hill: University of North Carolina, Frank Porter Graham Child Development Institute.
  32. Peisner-Feinberg, E. S., & Schaaf, J. M. (2011). *Summary of key findings: Effects of the North Carolina More at Four prekindergarten program on children's school readiness skills*. Chapel Hill: University of North Carolina, Frank Porter Graham Child Development Institute.
  33. Lipsey, M. W., Farran, D. C., Bilbrey, C., Hofer, K. G., & Dong, N. (2011). *Initial results of the evaluation of the Tennessee Voluntary Pre-K program*. Nashville, TN: Vanderbilt University, Peabody Research Institute. Retrieved from [https://my.vanderbilt.edu/tnpreevaluation/files/2013/10/April2011\\_PRI\\_Initial\\_TN\\_VPK\\_ProjectResults.pdf](https://my.vanderbilt.edu/tnpreevaluation/files/2013/10/April2011_PRI_Initial_TN_VPK_ProjectResults.pdf)
  34. Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology*, 41, 872–884.
  35. Frede, E., Jung, K., Barnett, W. S., & Figueras, A. (2009). *The APPLES blossom: Abbott Preschool Program longitudinal effects study (APPLES). Preliminary results through 2nd grade*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.
  36. Frede, E., Jung, K., Barnett, W. S., Lamy, C. E., & Figueras, A. (2007). *The Abbott Preschool Program longitudinal effects study (APPLES): Interim report*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.

37. Huang, F. L., Invernizzi, M. A., & Drake, E. A. (2012). The differential effects of preschool: Evidence from Virginia. *Early Childhood Research Quarterly*, 27, 33–45.
38. Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.
39. Hustedt, J. T., Barnett, W. S., & Jung, K. (2008). *Longitudinal effects of the Arkansas Better Chance Program: Findings from kindergarten and first grade*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.
40. Colorado Department of Education. (2012). *Colorado preschool program: 2012 legislative report*. Denver, CO: Author.
41. Picard Center for Child Development and Lifelong Learning. (2013). *Eighth-grade outcomes for LA 4 cohort 1 students* [Technical brief]. Lafayette, LA: Author.
42. Malofeeva, E. V., Daniel-Echols, M., & Xiang, Z. (2007). *Findings from the Michigan School Readiness Program 6 to 8 follow up study*. Retrieved from HighScope Educational Research Foundation website: <http://www.highscope.org/file/Research/6%20-%208%20follow%20up%20report%20FINAL%2010-9-07.pdf>
43. Peisner-Feinberg, E. S., & Schaaf, J. M. (2010). *Long-term effects of the North Carolina More at Four prekindergarten program: Children's reading and math skills at third grade*. Chapel Hill: University of North Carolina, Frank Porter Graham Child Development Institute.
44. Barnett, W. S., Jung, K., Youn, M.-J., & Frede, M. C. (2013). *Abbott Preschool Program longitudinal effects study: Fifth grade follow-up*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.
45. Barnett, W. S., Frede, E. C., Mobasher, H., & Mohr, P. (1987). The efficacy of public preschool programs and the relationship of program quality to efficacy. *Educational Evaluation and Policy Analysis*, 10(1), 37–49.
46. Strategic Research Group. (2011). *Assessing the impact of Tennessee's pre-kindergarten program: Final report*. Columbus, OH: Author.
47. Andrews, R. J., Jargowsky, P., & Kuhne, K. (2012). *The effects of Texas's targeted pre-kindergarten program on academic performance* (CALDER Working Paper No. 84). Washington, DC: American Institutes for Research.
48. Huston, A., Gupta, A., & Schexnayder, D. (2012). *Study of early education in Texas: The relationship of pre-K attendance to 3rd grade test results*. Austin, TX: University of Texas, Ray Marshall Center for the Study of Human Resources.
49. Bania, N., Kay, N., Aos, S., & Pennucci, A. (2014). *Outcome evaluation of Washington State's Early Childhood Education and Assistance Program* (Document No. 14-12-2201). Olympia: Washington State Institute for Public Policy.
50. Picard Center for Child Development and Lifelong Learning. (2007–2008). *LA 4 & starting points: Prekindergarten program evaluation*. Lafayette, LA: Author
51. Fitzpatrick, M. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement. *The B.E. Journal of Economic Analysis & Policy*, 8, 1–38.
52. Cascio, E. U., & Schanzenbach, D. W. (2013, September). *The impacts of expanding access to high-quality preschool education*. Paper presented at the Fall 2013 Conference on the Brookings Papers on Economic Activity, Washington, DC. Retrieved from <http://www.brookings.edu/about/projects/bpea/papers/2013/fall-cascio-preschool-education>
53. Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2014). From birth to school: Early childhood initiatives and third-grade outcomes in North Carolina. *Journal of Policy Analysis and Management*, 33, 162–187.
54. Rosinsky, K. (2014). *The relationship between publicly funded preschool and fourth grade math test scores: A state-level analysis* (Master's thesis, Georgetown University). Retrieved from [https://m.repository.library.georgetown.edu/bitstream/handle/10822/709852/Rosinsky\\_georgetown\\_0076M\\_12517.pdf?sequence=1&isAllowed=y](https://m.repository.library.georgetown.edu/bitstream/handle/10822/709852/Rosinsky_georgetown_0076M_12517.pdf?sequence=1&isAllowed=y)
55. Gilliam, W. S., & Zigler, E. F. (2001). A critical meta-analysis of all evaluations of state-funded preschool from 1977 to 1998: Implications for policy, service delivery and program evaluation. *Early Childhood Research Quarterly*, 15, 441–473.
56. Granger, R. C. (2011, Winter). The big why: A learning agenda for the scale-up movement. *Pathways*, 28–32.
57. Balu, R., Zhu, P., Doolittle, F., Schiller, E., Jenkins, J., & Gersten, R. (2015). *Evaluation of Response to Intervention practices for elementary school reading* (NCEE 2016-4000). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
58. Kirp, D. (2015, October 3). Does pre-K make any difference? *The New York Times*. Retrieved from [http://www.nytimes.com/2015/10/04/opinion/sunday/does-Pre-K-make-any-difference.html?\\_r=0](http://www.nytimes.com/2015/10/04/opinion/sunday/does-Pre-K-make-any-difference.html?_r=0)
59. ReadyNation. (n.d.). *Business case for early childhood investments*. Retrieved from <http://readynation.s3.amazonaws.com/wp-content/uploads/ReadyNation-Business-Brief-Endnotes.pdf>
60. Christeson, W., Bishop-Joseph, S., O'Dell-Archer, N., Beakey, C., & Clifford, K. (n.d.). *I'm the guy you pay later*. Retrieved from Fight Crime: Invest in Kids website: [http://cdn.fightcrime.org/wp-content/uploads/I'm\\_The\\_Guy\\_Report.pdf](http://cdn.fightcrime.org/wp-content/uploads/I'm_The_Guy_Report.pdf)