

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/284542076>

# Evidence for the Benefits of State Pre-Kindergarten Programs: Myth and Misrepresentation

Research · November 2015

DOI: 10.13140/RG.2.1.5075.5924

---

CITATIONS

0

---

READS

361

1 author:



[Dale C. Farran](#)

Vanderbilt University

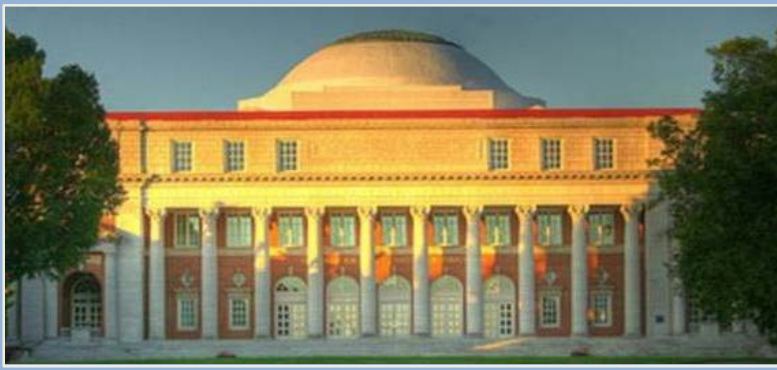
78 PUBLICATIONS 1,617 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Math Follow-Up Study [View project](#)



# Evidence for the Benefits of State Pre-kindergarten Programs: Myth and Misrepresentation

---

Dale C. Farran, PhD  
Mark Lipsey, PhD

11/9/2015

***DRAFT***

To appear in  
*Behavioral Science and Policy*  
Special Issue on Early  
Childhood and Preschool  
Programs

**Director**  
Mark W. Lipsey, Ph. D.

**Senior Associate Director**  
Dale C. Farran, Ph. D.

**Associate Director**  
Sandra Jo Wilson, Ph. D.

**Peabody Research Institute**  
Vanderbilt University

Our mission is to conduct research aimed at improving the effectiveness of programs for children, youth, and families. Using field research, program evaluation, and research synthesis (meta-analysis), our faculty and staff help determine which programs are actually making a difference in the lives of the people they serve. PRI research addresses many aspects of child and family programs, such as their implementation, costs, dissemination, and social or political support. But the main focus for all of our work is the effects of programs on children and families.

**Recommended Citation:**

Farran, D. C., & Lipsey, M. W. (2015). *Evidence for the Benefits of State Pre-kindergarten Programs: Myth and Misrepresentation*. Manuscript submitted for publication, Nashville, TN: Vanderbilt University, Peabody Research Institute.

**Funding Source:**

Development supported by Grant # R305E090009 from the Institute of Education Sciences, U.S. Department of Education for the study titled "Evaluating the Effectiveness of Tennessee's Voluntary Pre-kindergarten Program".

**Contact Us:**

Phone: 615.322.8540  
Fax: 615.322.0293

**Mailing Address:**

Peabody Research Institute  
230 Appleton Place  
PMB 181  
Nashville, TN 37203-5721

**Delivery Address:**

Peabody Research Institute  
1930 South Drive  
Room 410A  
Nashville, TN 37212

## **Evidence for the Benefits of State Pre-kindergarten Programs:**

### **Myth and Misrepresentation**

Dale C. Farran and Mark W. Lipsey  
Vanderbilt University

### **The Status of Statewide Pre-K Programs**

#### **Claims for Effectiveness**

With the publication in *Science* of Heckman's 2006 call for investments in early childhood education for disadvantaged children, the momentum within states for the creation or expansion of state-funded pre-k programs increased dramatically. Heckman's conclusions about the benefits of such investments were based on research about the effects of the Perry Preschool program, Abecedarian, and more recent studies of the Chicago Child Parent Center program. A similar assessment appears in a report released by the White House in 2014, *The Economics of Early Childhood Investments*, that makes reference to a "deep" research base supporting pre-k beginning with Perry and Abecedarian. This report goes on to state that "dozens of other programs have been rigorously examined since the 1960's" (p. 21), supporting its assertions 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.

This belief in the long term payoff of pre-k programs is so well entrenched in the thinking of policy makers that the state sponsors of a relatively new program, the Great Start Readiness Program (GSRP) 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:

Research shows students attending this successful program [GSRP] did better throughout their academic careers, had lower dropout rates, had higher incomes as adults and were imprisoned less frequently than children who didn't attend preschool (Michigan Department of Education, n.d., pg. 1).

In a report commissioned by the state of Texas, Pianta and Wolcott (2014) assure policy makers that other states have demonstrated that lower cost programs can achieve the same results as the earlier "Cadillac" programs (to use their term) that Pianta and Wolcott acknowledge cost more than any state would be willing to spend. They assert that a successful statewide program can operate at a per pupil cost no higher than the K-12 per pupil expenditure and still achieve strong, long lasting results into adulthood that will produce cost savings of \$3-\$7 per dollar spent on pre-k.

The problem these advocates are arguing pre-k will solve is a genuine one. Children from low-income backgrounds achieve less well in school than children from higher income families, and that gap has grown so that it now eclipses the racial achievement gaps that have long been of concern (Reardon, 2011). The presumption that poor children can be made more ready for school at 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 (Farran, 2007). The issue we wish to address in this paper is not whether this is a notable and critically important goal; one could hardly argue otherwise. Our concern is the substantial discrepancy between the available evidence for the effects of contemporary scaled up, statewide pre-k programs and the benefits expected from them because of the strong claims that have been made about how compelling the supporting research is.

In this paper, we will place our recently reported results through 3<sup>rd</sup> grade from a randomized control trial of the Tennessee pre-k program within the spectrum of studies of varying methodological quality on the effects of such statewide programs. First, however, we raise a question about the independent variable for such research—can we actually define with any specificity what a statewide prekindergarten program is? We then consider the relevant dependent variables—what are the outcomes such programs are expected to affect? After a look at the research evidence specific to scaled-up state pre-k programs, we consider the difficulties associated with scaling up an effective program even in circumstances where the intended program and outcomes are well defined.

### **Defining State Pre-K Programs**

As Quinton (2015) noted: “...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” (p. 2). A review of the pre-k programs states are implementing reveals the following. First, states are doing remarkably different things and calling all of it pre-k. Second, the situation is in flux—many states have recently passed legislation substantially changing and/or expanding their programs. Third, in most states the programs follow a public school model. The majority of pre-k classrooms are housed in public elementary schools, a “full day” is defined as a school day (or half a school day), and a full year is defined as a public school 9-month year.

States differ on whether lead teachers in pre-k classrooms should be licensed or have a bachelor’s degree or only an Associates (2 year) degree. States differ in class size requirements; e.g., New Jersey limits class size to 15, Texas, has no rule about class size.

Similarly, they differ on the adult-child ratio, with most setting it at 1:10, but in Texas it may go as high as 1:22, a stark difference.

Most states target their program to children from low-income families, but they vary on their income eligibility cutoffs and how those should be determined. For example, while many states use the eligibility criteria for free and reduced price lunch (135% and 185% of the federal poverty level respectively), North Carolina uses 75% of the state's median income as the cutoff while in Michigan it is 250% of the federal poverty level (Connors-Tadros, Brown, & Schilder, 2015). And, in the pre-k expansion grants newly funded by the federal government, the eligibility requirement is 200% of the federal poverty level (U. S. Department of Education, 2014).

State pre-k programs are quite similar in one regard, however. No state requires implementation of a program like either Perry or Abecedarian, each of which has research showing long term positive effects. The programs funded have few of the distinctive components of those widely cited programs. Common practice instead is to require that programs funded with state money implement an identified early childhood curriculum. In New Jersey, for example, programs are limited to choosing among High Scope, Tools of the Mind, Creative Curriculum and Curiosity Corner. Three of these have been evaluated by the What Works Clearinghouse and found to have effects no different from generic practice as usual (WWC, 2015), and the fourth (High Scope) has not yet been evaluated.

Almost none of the states invest money in monitoring their pre-k programs once contracts with school systems are awarded. Most states allocate some portion of the cost of the program itself from a central pool and then require local matching (though both the state amounts and the match requirements vary tremendously among states), but

monitoring the quality of the program as it is implemented is then left up to the local school systems. A few states such as Tennessee and Louisiana require “self assessment” on such measures as the Early Childhood Environmental Rating Scale (ECERS), but no one monitors the rigor with which those are done or how the results are used. Only rarely do states even require their public school pre-k programs to be rated by the state quality system (QRIS) required of community childcare programs.

### **Setting the Goals for Pre-K**

Along with the programmatic differences listed above, a further area of diversity, indeed, ambiguity for many state programs relates to the goals of the program. If there was consensus about the goals for pre-k, even within a state, there might still be disagreement on the best way to attain them but 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.

***School readiness.*** Kindergarten readiness frequently appears as a primary objective of public pre-k, including in the most recent request for proposals for pre-k expansion from the Obama administration (U.S. Department of Education, 2014). School readiness has been variously and often rather vaguely defined, but includes some mix of threshold literacy and numeracy skills, proper school behavior, and sometimes motor development and health. Strictly speaking, school readiness as a goal includes no assumption of sustained effects beyond kindergarten entry (e.g., Huang, Invernizzi, & Drake, 2012).

***School achievement.*** This goal focuses on learned academic content and skills that go beyond simple school readiness and is 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 of sustained effects on academic achievement. Those effects diminished well before high school even for the Perry preschool and Abecedarian programs (Duncan & Magnuson, 2013). Nonetheless, it is evident that there is an expectation by sponsors and stakeholders, including the White House, that state pre-k will contribute to closing achievement gaps for minority and/or economically disadvantaged children. As such, whatever initial positive effects pre-k has on achievement are expected to be sustained into the later grades, and it is disappointing to those sponsors and stakeholders if they are not.

***Behavioral outcomes.*** These are the “non-cognitive” outcomes Heckman and others cite as a primary justification for scaling up pre-k programs (Heckman, 2006). They include attendance, disciplinary infractions, graduation, employment, criminal behavior, and the like. These behaviors are not directly reflective of academic achievement performance (unlike grade retention and some categories of special education placement). Expectations of effects in these areas come from the classic longitudinal studies and are the primary basis for the claims of cost effectiveness for pre-k (Heckman, Moon, Pinto, Savelyev, & Yavitz, 2010).

In the next section, we describe the state of research into these variously defined statewide pre-k programs.

### **Research on the Effects of State Pre-K Programs**

Well more than a dozen studies of the effects of state-funded scaled-up pre-k programs have been reported, although almost none of them appear in peer reviewed outlets. One of the most striking features of these studies is how limited and methodologically weak they are. In the sections that follow, we summarize the research

designs, outcome variables, and findings of these studies. We first review the evidence they provide about the immediate effects of state pre-k programs—effects found at the end of the pre-k year or the beginning of kindergarten. We then turn to the evidence about the extent to which those effects are sustained through the school years following pre-k.

### **Evidence on the Immediate Effects of State Pre-K Programs**

***Randomized studies.*** The only randomized study of a state pre-k program is the one we have conducted, and it is quite recent and not yet complete (Lipsey, Farran, & Hofer, 2015). This study is based on 58 schools that received more applicants for pre-k than they could accommodate and agreed to admit children in random order until the available seats were filled. The remaining applicants served as the control group (a design analogous to that used in the national Head Start Impact study; Puma et al., 2012). Outcome data at the end of the pre-k and beginning of the kindergarten year were obtained from a subsample of children with parental consent on a battery of Woodcock Johnson III achievement tests and a set of teacher ratings of classroom behavior. The pre-k participant and nonparticipant groups in this comparison were virtually identical on baseline variables that included pretests of the achievement measures, demographic characteristics of the children, and family background variables obtained from parent interviews. To provide further assurance of equivalence, propensity scores and selected covariates were used as statistical controls in the analysis (further details are available in Lipsey, Farran, & Hofer, 2015). As shown in Table 1, statistically significant positive pre-k effects were found on all but one of the Woodcock Johnson achievement measures (Oral Comprehension with  $p=.093$ ) and on three of the six teacher ratings obtained at kindergarten entry (Preparedness for Grade, Work-Related Skills, and Social Behavior).

***Regression-discontinuity studies.*** The immediate effects of state pre-k programs, are most frequently studied with an age-cutoff regression discontinuity design (RDD). This design capitalizes on the age threshold for enrollment enforced by most state pre-k programs that requires children's birthdays to be before a specified cutoff date. Younger children whose birthdays come after that date must wait until the next school year to be eligible for enrollment. Outcomes measured at the beginning of kindergarten for children who attended pre-k the previous year (treatment group) are then compared with those measured at the same time for children in the lagged cohort who are just beginning pre-k (control group). With the age relationship adequately modeled in the statistical analysis, estimates of the effects of pre-k participation can be derived from this comparison.

Reports of studies using the age-cutoff RDD almost universally include language about the inherent strength of the regression-discontinuity design, which is indeed widely acknowledged as the quasi-experimental design with the best claim on internal validity (Cook & Wong, 2008). What they fail to recognize is that the age-cutoff version of this design is not a true RDD but, rather, an approximation that is vulnerable to multiple biases that can compromise the effect estimates (Lipsey, Weiland, Yoshikawa, Wilson, & Hofer, 2015). A true RDD divides a single intent-to-treat sample into groups that experience the different conditions to be compared. The pre-k age cutoff version of this design, by contrast, does not start with a complete intent-to-treat sample but, rather, compares children in two successive cohorts who have elected to enroll in the pre-k program. There are many ways the children in those cohorts may differ other than the known factor of age, and any such differences related to an outcome will bias the estimates of the pre-k effect on that outcome.

In particular, there is good reason to believe that the age cutoff RDD will tend to yield upwardly biased effect estimates. First, the greater mobility of economically disadvantaged families may result in more of those children in the pre-k treatment group being unavailable for outcome assessment at the beginning of kindergarten the next year. In addition, assessing the outcomes of the control children at the beginning of pre-k means that the counterfactual condition is their experience during the prior year, i.e., their three-year old year. This is a different counterfactual than would be represented by four-year old pre-k eligible children who did not attend pre-k and likely involves educational experiences and options that are not as extensive as those available to comparable four-year olds.

The largest number of pre-k age-cutoff RDD studies has been conducted by researchers associated with the National Institute for Early Education Research (NIEER) and include state funded pre-k programs in Arkansas (Jung, Barnett, Hustedt, & Francis, 2013), California (Barnett, Howes, & Jung, 2009), Michigan (Lamy, Barnett, & Jung, 2005a), New Jersey (Lamy, Barnett, & Jung, 2005b), New Mexico (Hustedt, Barnett, Jung, & Friedman, 2010), Oklahoma (Lamy, Barnett, & Jung, 2005c), South Carolina (Lamy, Barnett, & Jung, 2005d), and West Virginia (Lamy, Barnett, & Jung, 2005e). These studies are poorly documented and 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 conducted. For some, there is no indication that any covariates were included beyond the age selection variable or that any exploration of the functional form of the relationship between age and the outcome variables was made to ensure that the analysis model was appropriate.

The outcomes measured and the effect sizes reported in these studies are summarized in Table 1. These outcomes include only cognitive measures, in particular, PPVT vocabulary, Woodcock Johnson Applied Problems, and CTOPPP Print Awareness. Effect sizes from the Wong, Cook, Barnett, and Jung (2008) reanalysis of data from five of these studies are reported where available; the effect sizes that resulted from that reanalysis were different from those originally reported for many outcomes. As shown in Table 1, the effect sizes from these studies showed wide variation but mostly indicated positive pre-k effects.

Other age-cutoff RDD studies include one implemented as an adjunct to the randomized Tennessee study we have underway (described above). Findings for the first phase of that RDD, implemented in central Tennessee, have been reported in Lipsey, Farran, Bilbrey, Hofer, & Dong (2011). Pre-k effects were assessed with a battery of Woodcock Johnson achievement tests (itemized in Table 1) and special care was taken in the analysis to ensure that the children from the cohorts providing the comparison groups were as comparable as possible with regard to their demographic characteristics, pre-k school and classroom, and mobility. Consistent with the expectation that age cutoff RDDs overestimate effects, the effect size estimates were generally larger than those from the randomized component of the Tennessee study (also shown in Table 1).

Additional age-cutoff RDD studies have been reported for the pre-k programs in Georgia (Peisner-Feinberg et al., 2014) and North Carolina (Peisner-Feinberg & Schaaf, 2011). Various literacy, language, and math outcomes were assessed in both studies (itemized in Table 1), but the Georgia study was distinctive for also including behavioral measures of social skills, problem behavior, and social awareness. The analysis in these

studies made use of a range of covariates to better ensure that the comparison groups were comparable and showed positive effect sizes on all end of treatment outcomes (also shown in Table 1).

***Post Hoc Matched Studies.*** A number of studies compare outcome measures taken at the beginning of the kindergarten year for a sample of children divided into those who participated in the state pre-k program the previous year and those who did not. To make these groups somewhat comparable, this comparison is usually restricted to students who are similar on one or more demographic variables, typically gender, race/ethnicity, age, and/or free and reduced price lunch eligibility in the kindergarten year. From the standpoint of internal validity, this posttest-only design is very weak and is generally recognized as a pre-experimental design incapable of supporting causal inferences about intervention effects (e.g., Campbell & Stanley, 1966).

The major deficiency of these designs is the absence of baseline data from the beginning of the respective pre-k school year that demonstrates equivalence on key variables predictive of later outcomes or allows for matching or statistical controls to enhance equivalence. Most critical, with no pretest data for the outcome measures, it is unknown whether the pre-k participant and nonparticipant children started in the same place on those variables. Similarity on broad demographic characteristics in kindergarten does not ensure similarity of initial cognitive skills, relevant behavioral dispositions, or family background and context.

Even if children who attended the state pre-k program the prior year and those who did not could be matched on a range of typical baseline variables, the obvious fact that some parents enrolled their children in the program and others did not would be cause for

concern. The difference between parents who enroll their children in the state pre-k program and those who do not may simply be a matter of irrelevant circumstantial variation, but it seems likely that it reflects some differences in parents' motivation, value for education and ability to support it, aspirations for their children, and the like that may well influence their children's school performance.

Post hoc matched studies of this sort have been reported for Arkansas (Jung, Barnett, Hustedt, & Francis, 2013), California (Barnett, Howes, & Jung, 2009), Michigan (Xiang & Schweinhart, 2002), New Jersey (Frede et al., 2007; Frede et al., 2009), and Virginia (Huang, Invernizzi, & Drake, 2012). The outcomes assessed and the differences found between pre-k participants and nonparticipants represented in effect size terms are shown in Table 1. Though none provide credible estimates of the immediate effects of pre-k for the reasons indicated, it is worth noting that these estimates are generally smaller than those found in the RDD studies, even in studies of the same states conducted by the same researchers (e.g., Arkansas and New Jersey).

### **Evidence about the Sustained Effects of State Pre-K Programs**

As described above, there is almost no methodologically rigorous research on the immediate effects of state pre-k programs, but there is an overall pattern of positive effect estimates of sufficient magnitude to suggest possible benefits, particularly on achievement outcomes, even if all those estimates were to be upwardly biased. Moreover, that conclusion is consistent with the broader research literature on the short-term cognitive effects of early childhood education (Camilli, Vargas, Ryan, & Barnett, 2010). The available research on the extent to which state pre-k programs produce effects that are sustained past the end of the pre-k year, however, is far less convincing.

The Tennessee study described earlier is, again, the only randomized study that reports post pre-k outcomes (Lipsey, Farran, & Hofer, 2015). The follow-up data in that study revealed that all of the positive achievement differences favoring the state pre-k participants at the end of the pre-k year were no longer statistically significant by the end of kindergarten. Moreover, by the end of the third grade year, the children in the control group of nonparticipants actually scored higher, on average, than the pre-k participants on some of the achievement measures. The pattern of these results shows an acceleration in achievement for the nonparticipants such that their performance quickly overtook, and in some cases then surpassed, that of the pre-k participants. Teacher ratings of preparedness for grade, work-related skills, and other such outcomes showed a similar pattern.

While the findings of one study of the sustained effects of a state pre-k program should not be over interpreted, the remaining research on this topic is so poor that it offers little additional information of sufficient credibility to support any firm conclusion. Age cutoff regression-discontinuity designs cannot be used to examine effects longitudinally—the control group completes pre-k within a year and no longer provides an informative comparison. Other than the Tennessee study, the largest group of available studies uses pre-experimental post hoc matched designs to compare outcomes for state pre-k participants and nonparticipants for some number of school years past the beginning of kindergarten. These designs have the sources of potential bias described earlier plus the added possibility of differential attrition from the sample as the time past the pre-k year increases.

Such post hoc comparisons with at least some crude degree of matching have been reported for state pre-k programs in Arkansas (Hustedt, Barnett, & Jung, 2008; Jung,



Barnett, Hustedt, & Francis, 2013), Colorado (Colorado Department of Education, 2012, 2015), Louisiana (Picard Center, 2007-08, 2013), Michigan (Malofeeva, Daniel-Echols, & Xiang, 2007; Xiang & Schweinhart, 2002), North Carolina (Peisner-Feinberg & Schaaf, 2010), New Jersey (Barnett, Jung, Youn, & Frede, 2013; Frede et al., 2007, 2009), South Carolina (Barnett, Frede, Mobasher, & Mohr, 1987), Tennessee (Strategic Research Group, 2011), Texas (Andrews, Jargowsky, & Kuhne, 2012; Huston, Gupta, & Schexnayder, 2012), Virginia (Huang, Invernizzi, & Drake, 2012), and Washington (Bania, Kay, Aos, & Pennucci, 2014). As with the studies of the immediate effects of state pre-k programs, the outcomes examined are almost exclusively measures of achievement, either tests or grade retention. The reported results vary considerably in magnitude and statistical significance, but overwhelmingly favor the children who attended the state pre-k program. Those reports typically use the language of causal attribution, interpreting performance differences in later grades between pre-k participants and nonparticipants as pre-k effects despite the fact that such claims are unwarranted on methodological grounds. Because they are so uninterpretable, the effect estimates from these studies are not itemized here, but it is instructive to review the general nature of these studies.

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 (Picard Center, 2007-08, 2013). State achievement test scores for 3<sup>rd</sup> through 8<sup>th</sup> grade children eligible for the free or reduced price lunch (FRPL) programs were compared for children who had attended the state pre-k program and those who had not attended (and, separately, for race/ethnic and gender subgroups). The higher scores for pre-k participants were described as an indication that pre-k

participation increased performance and showed “positive impact” on achievement. Other than being similar on FRPL status, 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 in any other year, on any of the myriad other individual and family variables potentially related to their test performance on which they could easily have differed.

Some of the post hoc matched studies managed to obtain more data about student characteristics that were then used for matching or statistical control. However, only a few of these went beyond FRPL status, age, gender, and race/ethnicity and none included any true baseline data beyond static demographics. Though 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.*** A few difference in difference (DD) studies have been conducted on the effects of state-level pre-k programs on student achievement in the later grades. These studies examine before and after differences in state or county level student scores as the pre-k program is rolled out in comparison to differences over a comparable period for another area in which there was no analogous pre-k implementation or expansion. The challenge for this design is to isolate the difference made in the target outcomes by pre-k implementation from all 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. These studies typically compare changes in NAEP scores for children in the state of interest to the rest of the U.S.

Fitzpatrick (2008), for example, used a DD design to investigate the effects of the Georgia pre-k program that grew from participation rates of 14% in 1995 to 55% in 2008. The analysis included control variables for a range of factors that could have changed across the years independently of the introduction of universal pre-k. Initial analyses indicated positive pre-k effects, but further analyses exploring control group variants and different methods of inference did not yield a robust conclusion that the relationship was statistically significant. Similar sensitivity in the results was found in the Cascio and Schanzenbach (2013) DD study of the Georgia and Oklahoma programs. When the comparison states were limited to others in the south rather than all other states in the U.S., the DD estimates became substantially smaller. In the end, the authors could only support a conclusion of marginal statewide effects from the introduction of universal pre-k.

The difficulty of drawing firm conclusions from DD analyses in the dynamic context of state pre-k expansion and the limited availability of the statistical control variables so essential to the analysis is further illustrated by an ambitious study conducted by Rosinsky (2014). She compared the 2007, 2009, and 2011 NAEP 4th grade math scores across multiple states to program enrollment by 4-year-olds six years previously in Head Start, state-funded pre-k, and special education preschools. Surprisingly she found a negative association between the NAEP math scores and enrollment in public pre-k, primarily the state funded pre-k.

**Concluding observations.** Overall, the methodological quality of research on the effects of state pre-k programs is very poor, showing little improvement since the Gilliam and Zigler review (2001) that reached the same conclusion. Correspondingly, the conclusions that can be supported about the effectiveness of scaled-up state-funded pre-k

programs are quite limited. On balance, the evidence for positive effects immediately after pre-k participation, albeit weak, is stronger than that for sustained effects. Favorable end of treatment effects were found in the one randomized study available and in the numerous age cutoff RDD studies that, despite their problems, are somewhat less vulnerable to bias than the remaining alternatives. Perhaps most encouraging, these findings are consistent with those from the larger body of research on the effects of early childhood intervention programs generally (Camilli et al., 2010).

By contrast, there is virtually no methodologically credible evidence that these immediate positive effects of state pre-k programs are sustained for any length of time past the end of the pre-k year. Sustained effects were not found in the one randomized study and the findings of the few difference-in-difference studies are ambiguous at best. The results of the longitudinal post hoc matched studies do favor pre-k participants, and have been cited by advocates as evidence of sustained effects, but are too methodologically flawed to support any attribution of that difference to pre-k participation. Moreover, the pattern of diminished achievement effects referred to as “fade out” in the broader body of early childhood research (Duncan & Magnuson, 2013) provides little basis for expecting strong sustained effects on achievement from state pre-k programs.

The range of outcome variables represented across all the studies of state pre-k is also noteworthy. The measures used are almost exclusively indices of academic achievement, primarily achievement tests and grade retention, demonstrating what seems to be a strong presumption that these are important to policy makers. Very few studies have examined behavioral outcomes despite the fact that those are the pre-k effects that are supposed be sustained long term and generate the cost savings that have been claimed

for pre-k. While investigation of long-term behavioral outcomes may be beyond the reach of time-limited state pre-k studies, that is not the case for potential mediators that may connect early pre-k experiences and later positive adult outcomes, e.g., self-regulation, engagement with school, grit, and the like. Those also are strikingly absent from the research on state pre-k effects.

### **Problems in Taking Programs to Scale**

Most of the research reviewed above was conducted by or commissioned by the state departments of education that administer the respective pre-k program. The lack of attention to the methodological weaknesses of these studies while emphasizing the positive effects they purport to demonstrate raises a question about the purpose of this research. Many of the reports of these studies do not read like genuine inquiries into the effectiveness of pre-k but, rather, have the appearance of offerings aimed at supporting the policies already adopted. A more critical approach in such studies would not only be more forthright about their methodological limitations and less rosy about the conclusions, but would acknowledge the considerable difficulty of implementing an effective program at scale and avoid implying that this had been successfully accomplished.

The research community has been effective at creating and validating many promising practices, but there is almost no evidence that any of these have been taken to scale while maintaining the same level of effectiveness (Granger, 2011). In the case of prekindergarten 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 major positive ways that last well into adulthood.

This big idea has come to states from the research conducted on the Cadillac programs implemented half a century ago that bear little resemblance to the programs states are implementing today. The Perry Preschool project was a two-year intervention begun when children were three years old with substantial parental involvement (parents were visited in their homes once a week for an hour and a half). The Abecedarian program 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. Even the somewhat more recent Chicago Child Parent program differs from state pre-k; the intervention continued into elementary school and had a substantial requirement for parent involvement.

Recent attempts to take other promising programs to scale highlight how difficult this enterprise is and why it is unlikely that statewide implementation of any pre-k model will find ready success. The U. S. Department of Education substantially funded (\$50 million for 5 years) several promising educational programs to allow them to be taken to scale, and reports on three of them have just been published: Success for All (Quint, Zhu, Balu, Rappaport, & Delaurentis, 2015); KIPP (Knechtel et al., 2015), and Teach for America (Clark, Isenberg, Liu, Makowshy, & Zukiewicz, 2015). Each of these endeavors required compromises as the programs were scaled up, not quite implementing the program model that had been experimentally validated. And each of them found that the effects of the scaled up versions of their programs were different from those expected on the basis of the supporting research.

While it is illuminating to examine the successes and failures of these efforts, it is important to recognize one critical way in which scaling up pre-k programs presents even greater challenges than those faced in these examples. Unlike pre-k, these programs are well defined in a way that supports a clear vision of what is to be implemented and what it will look like if that implementation is successful. The evaluations showed that even when programs are well defined, have a coherent vision, are sufficiently funded, and have recent evidence of effectiveness, there are problems going to scale.

Scaling up pre-k more closely resembles the recent scale up effort for Response to Intervention (RtI). RtI was developed and evaluated in small experimental studies that demonstrated the effectiveness of a closely monitored small group reading program to help early grades readers. On that evidence, the program was then moved into policy and written into the reauthorization of IDEA in 2004. RtI, like pre-k, is a set of concepts rather than a structured program, and districts implement it however they see fit without monitoring or oversight from the developers. A multistate evaluation of RtI implemented at scale actually found *negative* effects (Balu, Pei, Doolittle, Schiller, Jenkins, & Gersten, 2015).

The pre-k concept is much like RtI; there is no coherent vision for what the program should look like other than the idea that it should somehow be like the Perry, Abecedarian, and Chicago programs, but not so much so that it would require as much time, effort, staff, or cost. Viewed realistically, we 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 (e.g., Kirp, 2015). As the research we have reviewed above indicates, there is little evidence of such effects.

## **Conclusion**

As we noted earlier, the importance of intervening to help the school achievement of young children from impoverished backgrounds cannot be overestimated and is, in fact, growing. The idea that a year of pre-k can achieve miraculous results manifest well into adulthood is appealing to policy makers, school administrators, businessmen (ReadyNation, n.d.), and even law enforcement (Christeson, Bishop-Joseph, O'Dell-Archer, Beakey, & Clifford, n.d.). If the school achievement of poor children could be fixed with a year of preschool, states would be foolish not to implement it. The idea and the needs are so very compelling that states have rushed to scale up something called pre-k with inadequate attention to what the program should be or how it must be supported to be effective. But our concern is what will happen to the needs of poor children when the programs being implemented do not show the promised results.

The research on effects of statewide implementation of pre-k is clearly inadequate methodologically and, in part for that reason, does not support the claims made by advocates and proponents about the nature of the evidence for the positive effects of pre-k. While it would be easy to call for more and better research, the problems here seem to be more fundamental.

Program definition, as we elaborated earlier, is all over the place. States are doing various things and calling them pre-k. From a methodological standpoint, we have a huge black box for an independent variable such that even if good effects were found, researchers would be hard pressed to specify the key features of the program that caused them. Thus, there is scant help for states as they seek to alter their programs (as most are doing) to tell them which aspects are most likely to have the most payoff.



Both the programs being launched and the research to support them lack a distinct and plausible theory of change that would articulate the expectations for what happens (or should happen) in pre-k and how that will affect immediate and long-term outcomes. Such a theory would help identify the outcomes states should be concerned with and what they should look for to know if their program is achieving those outcomes. The classic studies point to long term behavioral outcomes, but current research is dominated by achievement (a possible effect of the increasing assimilation of pre-k programs into the public schools). Few legislators are prepared to wait for 15 to 20 years to determine if their programs are successful; the research field needs to step up and conduct credible research to identify more proximal measures related to these long term outcomes.

Children and education policy makers are not being well served by a perpetuation of magical thinking about poorly defined state run pre-k programs. Moreover, researchers should not be aiding and abetting that thinking with weak and misleading research, the not-very-credible results of which are given a gloss in the reporting that obscures the fact that there are real questions about whether state pre-k programs in fact do anything other than boost 4-year olds' academic cognitive skills to where they would be at the end of kindergarten anyway.

## References

- 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. <http://www.caldercenter.org>
- Balu, R., Zhu, P., Doolittle, F., Schiller, E., Jenkins, J., and Gersten, R. (2015). *Evaluation of Response to Intervention practices for elementary school reading* (NCEE 2016-4000). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- 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, WA: Washington State Institute for Public Policy.
- 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.
- 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.
- 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.
- 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.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Company.
- Cascio, E. U. & Schanzenbach, D. W. (2013). *The impacts of expanding access to high quality preschool education*. Paper presented at the Fall 2013 Brookings Panel on Economic Activity.
- Christeson, W., Bishop-Joseph, S., O'Dell-Archer, N., Beakey, C., & Clifford, K. (n.d.). *I'm the guy you pay later*. Washington, DC: Fight Crime/Invest in Kids.
- Connors-Tadros, L. Brown, K., & Schilder, D. (May, 2015). *Examples of state guidance to determine and verify income eligibility for prekindergarten programs*. Rutgers, NJ: Center for Enhancing Early Learning Outcomes.
- Clark, M., Isenberg, E., Liu, A., Makowsky, L., & Zukiewicz, M. (March 2015). *Impacts of the Teach for America investing in innovation scale up*. Washington, DC: Mathematica Policy Research.
- Colorado Department of Education (2012). *Colorado preschool program: 2012 legislative report*. Denver, CO: Author.
- Colorado Department of Education (2015). *Colorado preschool program: Legislative report 2015*. Denver, CO: Author.
- Cook, T. D., & Wong, V. C. (2008). Empirical tests of the validity of the regression discontinuity design. *Annals of Economics and Statistics*, 91/92, 127-150.
- Duncan, G. & Magnuson, K. (2013). Investing in preschool programs. *The Journal of Economic Perspectives*, 27, 109-131. [dx.doi.org/10.1257/jep.27.2.109](https://doi.org/10.1257/jep.27.2.109).
- Executive Office of the President of the United States (December 2014). *The economics of early childhood investments*. Washington, DC: The White House. 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)
- Farran, D. C. (2007). *Is education the way out of poverty? A Reflection on the 40<sup>th</sup> anniversary of Head Start* (with commentaries by James King and Bernard L. Charles), Center for Research on Child Development and Learning, No. 3 (50 pages – ISBN: 0-9727709-2-5).

- 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.
- 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* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.
- Frede, E., Jung, K., Barnett, W. S., Lamy, C. E., & Figueras, A. (2007). *The Abbott Preschool Program longitudinal effects study (APPLES): Interim report* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.
- 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(4), 441-473
- Granger, R. C. (2011). The big why: A learning agenda for the scale-up movement. *Pathways* (Winter): 28-32.
- Heckman, J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312, 1900-1902. <http://www.jstor.org/stable/3846426>
- Heckman, J., Moon, H., Pinto, R., Savelyev, P., & Yavitz, A. (2010). The rate of return to the High Scope Perry Preschool program. *Journal of Public Economics*, 94, 114-128. doi:10.1016/j.jpubeco.2009.11.001
- 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.
- Hustedt, J. T., Barnett, W. S., & Jung, K. (2008). *Longitudinal effects of the Arkansas Better Chance Program: Findings from kindergarten and first grade* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.
- 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* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.
- 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.
- 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* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.
- Kirp, D. (Oct. 3, 2015). Does pre-k make any difference? *New York Times, Sunday Review*. [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)
- Knechtel, V., Anderson, M., Burnett, A., Coen, T., Sullivan, M., Tuttle, C., & Gleason, P. (September 2015). *Understanding the effect of KIPP as it scales: Volume II, leadership practices at KIPP*. Washington, DC: Mathematica Policy Research.
- Lamy, C., Barnett, W. S., & Jung, K. (2005a). *The effects of the Michigan School Readiness Program on young children's abilities at kindergarten entry* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.
- Lamy, C., Barnett, W. S., & Jung, K. (2005b). *The effects of New Jersey's Abbott Preschool Program on young children's school readiness* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.
- Lamy, C., Barnett, W. S., & Jung, K. (2005c). *The effects of Oklahoma's early childhood four-year-old program on young children's school readiness* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.

- Lamy, C., Barnett, W. S., & Jung, K. (2005d). *The effects of South Carolina's early childhood programs on young children's school readiness* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.
- Lamy, C., Barnett, W. S., & Jung, K. (2005e). *The effects of West Virginia's early education program on young children's school readiness* (Research Report). New Brunswick, NJ. Rutgers University, National Institute for Early Education Research.
- 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* (Research Report). Nashville, TN Vanderbilt University, Peabody Research Institute. Available from: <http://peabody.vanderbilt.edu/research/pri/>
- 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* (Research Report). Nashville, TN Vanderbilt University, Peabody Research Institute. Available from: <http://peabody.vanderbilt.edu/research/pri/>
- 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
- Malofeeva, E. V., Daniel-Echols, M., & Xiang, Z. (2007). *Findings from the Michigan School Readiness Program 6 to 8 follow up study*. Ypsilanti, MI: High/Scope Educational Research Foundation. <http://www.highscope.org>
- Michigan Department of Education. (n.d.) *Great Start Readiness Program*. [http://www.michigan.gov/documents/mde/GSRP-Overview\\_410757\\_7.pdf](http://www.michigan.gov/documents/mde/GSRP-Overview_410757_7.pdf)
- 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: The University of North Carolina, FPG Child Development Institute.
- 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: The University of North Carolina, FPG Child Development Institute.
- 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: The University of North Carolina, FPG Child Development Institute.
- 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 11/7/15: [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)
- Picard Center for Child Development and Lifelong Learning (2007-08). *LA 4 & starting points: Prekindergarten program evaluation*. Lafayette, LA: Author.
- Picard Center for Child Development and Lifelong Learning (2013). *Eighth-grade outcomes for LA 4 cohort 1 students* (Technical Brief). Lafayette, LA: Author.
- Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, D., Mashburn, A. & Downer, J. (2012). *Third grade follow-up to the Head Start Impact Study final report*, OPRE Report # 2012-45, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Quint, J., Zhu, P., Balu, R., Rappaport, S., & DeLaurentis, M. (September 2015). *Final report from the Investing in Innovation (i3) evaluation*. New York, NY, MDRC. Retrieved from [http://www.mdrc.org/sites/default/files/SFA\\_2015\\_FR.pdf](http://www.mdrc.org/sites/default/files/SFA_2015_FR.pdf)
- Quinton, S. (September 4, 2015). *States agree on need for 'preschool,' differ on definition*. The Pew Charitable Trusts/Research & Analysis/Stateline. <http://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/> Retrieved September 5, 2015.

- 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*. New York: Russell Sage Foundation.
- ReadyNation (n.d.). *Business case for early childhood investments*. Washington, DC: ReadyNation, [www.ReadyNation.org](http://www.ReadyNation.org). Retrieved September 12, 2015.
- Rosinsky, K. (2014). *The relationship between publicly funded preschool and 4<sup>th</sup> grade math test scores: A state-level analysis*. (Master's thesis, Georgetown University, Washington, DC). 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), August 31, 2015
- Strategic Research Group (2011). *Assessing the impact of Tennessee's pre-kindergarten program: Final report*. Columbus, OH: Author.
- U.S. Department of Education and Department of Health and Human Services (August 8, 2014). *Applications for New Awards; Preschool Development Grants – Expansion Grants*. Washington, DC: *Federal Register*, 79, No. 159.
- What Works Clearinghouse (2015). *Find what works/early childhood interventions*. Washington, DC: Institute of Education Sciences, US Department of Education. Retrieved 11/7/15 from: [ies.ed.gov/ncee/wwc/findwhatworks.aspx#Social-emotional development](http://ies.ed.gov/ncee/wwc/findwhatworks.aspx#Social-emotional%20development)
- 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
- Xiang, Z., & Schweinhart, L. J. (2002). Effects five years later: The Michigan School Readiness Program evaluation through age 10 (Research Report). Ypsilanti, MI: High/Scope Educational Research Foundation. <http://www.highscope.org>

Table 1  
*Effects Reported at the End of Pre-k or Beginning of K*  
*(Standardized mean difference effect sizes)*

	Literacy	Language	Math	Other
<b>Randomized Studies</b>				
<b><i>Tennessee</i></b> (Lipsey et al., 2015)				
WJ Composite				
WJ Letter-Word Identification	.41*			
WJ Spelling	.29*			
WJ Oral Comprehension		.09		
WJ Picture Vocabulary		.20*		
WJ Applied Problems			.17*	
WJ Quantitative Concepts			.27*	
TR Preparedness for K				.22*
TR Work-related Skills				.20*
TR Social Behavior				.19*
TR Peer Relations				.04
TR Behavior Problems				-.04
TR Feelings About School				-.03
<b>Regression-Discontinuity Studies</b>				
<b><i>Tennessee</i></b> (Lipsey et al., 2011)				
WJ Composite				
WJ Letter-Word Identification	.82*			
WJ Spelling	.99*			
WJ Oral Comprehension		.26*		
WJ Picture Vocabulary		.48*		
WJ Applied Problems			.48*	
WJ Quantitative Concepts			.50*	
<b><i>Arkansas</i></b> (Jung et al., 2013)				
CTOPPP Print Awareness	.82*			
PPVT		.28*		
WJ Applied Problems			.33*	
<b><i>California</i></b> (Barnett et al., 2009)				
CTOPPP Print Awareness	1.17*			
PPVT		.30*		
WJ Applied Problems			.38*	
<b><i>Georgia</i></b> (Peisner-Feinberg et al., 2014)				
Letter Knowledge	.89*			
WJ Letter-Word Identification	1.05*			
WJ Sound Awareness	.59*			
WJ Word Attack	1.20*			
WJ Picture Vocabulary		.01		
WJ Applied Problems			.51*	

	Literacy	Language	Math	Other
Counting			.86*	
Social Awareness				.43*
SSiS Social Skills				.23
SSiS Problem Behavior				.10
<b>Michigan</b> (Wong et al., 2008)				
CTOPPP Print Awareness	1.09*			
PPVT		-.13		
WJ Applied Problems			.53*	
<b>New Jersey</b> (Wong et al., 2008)				
CTOPPP Print Awareness	.32*			
PPVT		.36*		
WJ Applied Problems			.23*	
<b>New Mexico</b> (Hustedt et al., 2010)				
ELSA Early Literacy	1.30*			
PPVT		.24*		
WJ Applied Problems			.37*	
<b>North Carolina</b> (Peisner-Feinberg et al., 2011)				
Phonological Awareness (TOPEL)	.56*			
Print Knowledge (TOPEL)	1.16*			
WJ Letter-Word Identification	1.14*			
PPVT		.06		
WJ Applied Problems			.34*	
Counting			.81*	
<b>Oklahoma</b> (Wong et al., 2008)				
CTOPPP Print Awareness	.42			
PPVT		.28*		
WJ Applied Problems			.34	
<b>South Carolina</b> (Wong et al., 2008)				
CTOPPP Print Awareness	.78*			
PPVT		.04		
<b>West Virginia</b> (Wong et al., 2008)				
CTOPPP Print Awareness	.92*			
PPVT		.16		
WJ Applied Problems			.06	
<b>Post Hoc Matched Studies</b>				
<b>Arkansas</b> (Jung et al., 2013)				
CTOPPP Print Awareness	.15*			
PPVT		.12*		
WJ Applied Problems			.05*	
<b>California</b> (Barnett et al., 2009)				
CTOPPP Print Awareness	.30*			

	Literacy	Language	Math	Other
PPVT		.24*		
WJ Applied Problems			.15*	
<b>Michigan</b> (Xiang & Schweinhart, 2002)				
COR Language & Literacy	.37*	.37*		
COR Logic & Mathematics			.15	
COR Initiative				.34*
COR Social Relations				.37*
TR Readiness to Learn				.27*
TR Social Relations				-.13
<b>New Jersey</b> (Frede et al., 2007)				
CTOPPP Print Awareness	.29*			
PPVT		.21*		
WJ Applied Problems			.20*	
<b>Virginia</b> (Huang et al., 2012)				
PALS	.74*			

Notes: COR=High/Scope Child Observation Record. CTOPPP= Comprehensive Test of Phonological and Print Processing. ELSA=Early Literacy Skills Assessment. PALS=Phonological Awareness Literacy Screening. PPVT=Peabody Picture Vocabulary Test. SSiS=Social Skills Improvement System. TOPEL=Test of Preschool Early Literacy. TR=teacher ratings. WJ=Woodcock Johnson.

\*  $p < .05$