

# We need more evidence in order to create effective pre-K programs

Dale C. Farran

## Executive Summary

The proposition that expanding pre-K will improve later achievement for children from low-income families is premature. Premature as well is the presumption that solid research exists to guide the content and structure of pre-K programs. Despite more than 50 years of preliminary work on pre-K as an early intervention for young children from poor backgrounds, the field of early childhood education has a relatively small database to use as a guide to effective practice. Lack of evidence about which skills and dispositions are most important to effect in pre-K and what instructional practices would affect them has led us to the current situation of poorly defined, enormously varied programs, all called pre-K, as well as a reliance on a set of quality measures with no empirical validity. Despite being included in national and state policies and used to hold pre-K providers accountable, none of the widely used measures of classroom and center quality relates strongly, if at all, to child growth on the school readiness outcomes on which most pre-K programs are focused. The outlook for poor children is too dire to allow this situation to continue.

State investment in pre-K programs is predicated on the belief that pre-K is effective at promoting school readiness skills, and, most importantly, that those skills will then be linked to long-term school success, i.e., closing the achievement gap.<sup>1</sup> Oddly, despite huge support for pre-K from politicians and advocates alike, three critical questions related to this belief have not been addressed. The first is which skills and dispositions among young children are the most important to influence early for long-term success in school. The second is which classroom and center practices will work to promote those skills and dispositions. The final question is which assessment systems will validly capture the quality of those practices when they are scaled up for delivery in thousands of pre-K settings. In this paper, I argue that essential research to answer those questions has not yet been done.

Currently, participating in pre-K is voluntary and programs are varied in their content, staffing, and aegis. With the prospect of continued expansion, it is important to make sure public dollars are being invested wisely so that we do not squander this chance to positively affect the lives of young, vulnerable children.

## ***Important skills for long term school success***

---

After more than 50 years of offering pre-K as an intervention for children whose families live in poverty, early childhood researchers cannot yet give confident answers to the fundamental question of which early childhood competencies are most important for long-term school success. Existing longitudinal studies provide primarily correlational and cross-sectional evidence about the trajectory of school achievement over some span of time,<sup>ii</sup> but they do not provide definitive information about which skills to focus on in order to alter the trajectory.

At least part of the reason for the lack of precision in what to target can be found in the history of early intervention for children from poor families. According to Zigler,<sup>iii</sup> the roots of early intervention lie in an experimental program from the 1930s that successfully altered the intellectual development of children of mentally retarded adolescents who were institutionalized in a state orphanage.<sup>iv</sup>

During the 1960s, a number of experimental early childhood intervention programs were launched, including Head Start. And inspired by the orphanage

intervention of 30 years earlier, the common assessment used to measure the success of these programs was children's IQ scores.<sup>v</sup> Rick Heber and the Milwaukee Project made *Time* magazine with claims to have improved some children's IQs by as much as 70 points (compared to the IQ scores of their mothers), with an average gain of 30 points.<sup>vi</sup> For the next 20 years, the focus remained on producing higher scores on the global assessment of IQ, and early, general, enriched experiences were the presumed mechanism for transforming the way children's brains developed.

The major experimental early childhood programs cited frequently today to support the push for an expansion of pre-K programs reflected the dominant perspective of that same time period. In summarizing the history of the Perry Preschool Project (one of the most cited early intervention projects), Schweinhart<sup>vii</sup> describes how Weikart, a special education teacher, was concerned with school failure and retention of poor children. At that time, the special education category "mildly retarded" was used to categorize such children. It was presumed that if IQ could be raised through early intervention, children would avoid placement in special education classes and would perform better in school. Therefore, the outcome initially measured for the Perry Preschool Project was Stanford Binet IQ test scores.<sup>viii</sup>

Similarly, much of the early writing about the Abecedarian Project appeared in journals devoted to mental retardation (the term at the time for what is now called intellectual disability) and intelligence, and even those appearing in the broadly focused leading academic journal in the field, *Child Development*, had such titles as "The plasticity of intellectual development: Insights from preventive intervention"<sup>ix</sup> and "Biological nonoptimality and quality of postnatal environment as codeterminants of intellectual development."<sup>x</sup>

IQ test scores tend to correlate with all other more differentiated areas of development, such as language and memory, and the tests include samples of many types of skills to create one general and global measure. The primary difficulty with this approach as a basis for designing interventions is that there is no way to identify what specifically changed about children's abilities that enabled them to perform better in school or to link those changes to any particular set of active ingredients in the treatment. Neither Perry nor Abecedarian explicitly describes beyond the broadest level the "treatment" that brought about their positive effects.

In these earlier programs, general enrichment was associated with some initial improvement in a general measure of development. When original programs have been expanded to broader implementation (for example, Head Start<sup>xi</sup>), the same levels of effects have not been found. Indeed, [troubling signs](#) of unforeseen possible negative effects for statewide pre-K implementation have been found in the Tennessee statewide program.<sup>xii</sup> The reasons for the lack of generalizability from the 1960s programs to policy implementation illustrate how difficult it is to replicate effects or adjust a program to enhance effects when both the outcome and the treatment are presented as global and general.<sup>xiii</sup>

Abandoning IQ as the outcome leaves the field with the challenge of determining which specific competencies at the end of preschool are most important for children's long-term academic success. In other words, what particular skills should be targeted in pre-K classrooms and thus, what should be measured to determine the success of the program? The answers to these questions are not as straightforward as many curriculum publishers and policy advocates would have the field believe.

Addressing these important questions is hampered by several aspects of research in early childhood development. First, almost all of the studies are correlational. They link children's test scores at one point in time to test scores at a later point in time. Moreover, with few exceptions,<sup>xiv</sup> the correlations are based on the level of children's performance and not on the gains in particular areas as a result of intervention. For example, children's math and literacy scores as measured during the pre-K period predict later school and life success, but those scores correlate with many other contemporaneous measures of development, including IQ test scores, and they may simply be proxies for other influences over which pre-K programs have little or no control, including the quality of parent-child interactions in the home.

Those areas of development that are both amenable to intervention and for which intervention-induced gains are correlated with later performance should obviously be the ones targeted in pre-K programs, but there is relatively little in the existing literature to determine which those are.

Second, preconceptions of which skills are important for school success as well as the existence of established measures have strongly influenced the things that are measured and thus taught. What are

most easily measured are "school readiness" skills, taken most often to mean concrete measures of letter knowledge and early numeracy. But what are easily measured may not at all be the most central developmental processes to affect.

Recent neurocognitive research findings might provide the sort of guidance needed for a starting point. Noble and her colleagues reported results from 1,099 individuals aged 3 to 20.<sup>xv</sup> The development of brain surface area was strongly linked to family income, with the effect driven largely by the lowest income levels. Those areas most affected by low income were ones involving language, reading, executive functions (attending, working memory) and spatial skills. (It is important to note that the analyses controlled for age, genetic ancestry, and gender.)

Smaller neurocognitive studies of young children have found consequences for these same areas of development.<sup>xvi</sup> Moreover, emerging research involving assessments of children's early competencies confirm the neurocognitive evidence that these domains are already affected by poverty by age 3<sup>xvii</sup> or possibly even earlier.<sup>xviii</sup>

Neither the neurocognitive nor the assessment findings, of course, immediately suggest what interventions would be successful to interfere with the deleterious consequences of living in a high poverty environment. Careful, systematic research is needed to figure out what experiences might be effective as these same areas of development are the ones most resistant to change in general pre-K programs.

## ***Classroom practices promoting important skills***

---

Without clear, verified competencies identified as the desired child outcomes of pre-K programs, it should be no surprise that the measures defining high quality classrooms are general, global, and for the most part not empirically validated.

Quality measures have been dominated by three primary approaches—the benchmark ratings from the National Institute of Early Education Research (NIEER) and two classroom rating measures—the Early Childhood Environmental Rating System-Revised (ECERS-R) and the Classroom Assessment Scoring System (CLASS). Each of these measures has some notable psychometric problems, yet each of them has been woven into quite consequential policies. None of them was developed on the basis of empirical

knowledge of which skills are most important to affect in pre-K.

In order to create a standard for evaluating state-funded pre-K programs, NIEER developed a set of 10 benchmarks on which states are graded each year. None of the benchmarks requires actual observations of the classrooms. Instead, they deal with regulatory issues such as the adoption of early learning standards in the state, the requirement that lead teachers in each classroom have a bachelor's degree, the condition that assistants have a child development associate's degree (or equivalent), as well as issues of staff-child ratios and group size. Each year NIEER gathers the data related to the benchmarks from state officials and issues a yearbook giving state-by-state evaluations.<sup>xix</sup> States have looked to the NIEER benchmarks for guidance on creating high quality pre-K programs, assuming that these standards have been validated through research.

The ECERS-R is an essential component of many states' Quality Rating and Improvement Systems (QRIS). ECERS-R total scores help define different categorical levels (often called Stars as in the Tennessee system<sup>xx</sup>). The star rating a center earns is public and can affect parental interest in enrolling their children; in many states, the rating is also linked to reimbursement rates for childcare vouchers. Thus, the ECERS-R score can have serious financial implications for a center.

Similarly, the CLASS also now has important programmatic effects. It is being included in many states' QRIS plans. Moreover, in the reauthorization of the Head Start program, Congress specified that a reliable and valid observational measure had to be conducted as part of the renewal process for individual programs. The Federal Office of Head Start chose the CLASS to be that instrument. Use of the CLASS went into effect in December 2011. Head Start programs have to re-compete for their program award if any meet the following condition:

(2) After December 9, 2011, to have an average score across all classrooms observed that is in the lowest 10 percent on any of the three CLASS: Pre-K domains from the most recent CLASS: Pre-K observation among those currently being reviewed unless the average score across all classrooms observed for that CLASS: Pre-K domain is equal to or above the standard of excellence that demonstrates that the classroom interactions are above an exceptional level of quality. For all three

domains, the "standard of excellence" is a 6.<sup>xxi</sup>

Each year, the Office of Head Start publishes a list of programs by state that will be required to re-compete.<sup>xxii</sup> This is an extremely distressing event for a program.

The research into the relationship between these classroom "quality" measures and child outcomes has for the most part focused on literacy and math skills; language has received some attention, almost exclusively vocabulary, and executive function skills have actually received relatively little attention though that may be increasing. For these outcomes, the evidence that these quality measures matter is remarkably small to non-existent.

The most ambitious and comprehensive study of the relationship between the NIEER benchmarks and child outcomes can be found in the National Center for Early Development and Learning's Multi-State Study of Pre-Kindergarten (Multi-State) and the State-Wide Early Education Programs Study (SWEET).<sup>xxiii</sup> The researchers created individual measures of the 10 benchmarks for the 671 classrooms included in the study.

Associations between gains in language, literacy and math measures and the benchmarks, collectively and individually, were examined. The total NIEER benchmark score was related to none of the outcomes. Similarly, none of the individual benchmarks predicted gains on the outcomes.

A great deal of effort has gone into investigating the relationship between both ECERS-R and CLASS scores and child outcomes over the past many years. So far, the obtained relationships between these measures and child outcomes at the end of the pre-K year are modest, at best. In a major analysis involving hundreds of pre-K classrooms across four large samples,<sup>xxiv</sup> researchers found that an increase of 1 standard deviation in classroom quality (or about 1 point on the 7-point ECERS and CLASS scales) was associated with only about 1/20th of a standard deviation of language growth for children in those classrooms, the equivalent of moving up less than 1 point by the end of the year on a scale on which the typical gap between poor children and their more affluent peers at that age is about 10 points. The effects on math outcomes were even weaker. And there were no effects at all from any quality measure on social emotional skills or behavior problems, the so-called "soft skills."

In a study of the Boston pre-K classrooms (often praised for their high quality),<sup>xxv</sup> where there were stronger child outcomes, no relationships, not even weak ones, were found between the measures of quality (ECERS-R, CLASS) and child language outcomes (the only academic measure used).

Once again, the reasons for these disappointing results may lie in the history of these instruments. None of them was developed empirically. Each of them represents a set of developers' ideas about what was likely to be important for young children's development. Each has a great deal of face validity and thus each has convinced policy makers to incorporate them into policy in consequential ways. But while face validity and ideology can provide the starting point of an investigation, they must be supplemented with rigorous research actually validating their use before they are adopted into policy.

NIEER guidelines focus on the "structural" aspects of classrooms, elements easily incorporated into policy. The ECERS-R reflects a perspective that the materials in the classroom, the ways they are organized, and the amount of time children are allowed to explore them freely are critical quality features. The empirical work to determine which aspects of organization, which and how many types of materials, and how to facilitate children's focus during free play has not occurred.<sup>xxvi</sup> The CLASS proceeds from a perspective that the emotional atmosphere of a classroom and the teachers' interactions with children are the critical quality features. The point is not to call these ideological perspectives into question, but to argue that these are beliefs and not empirically validated measures of quality demonstrated to link to any

outcomes that might be of interest to the programs.

## Conclusion

The proposition that expanding pre-K will improve later achievement for children from low-income families is premature. Premature as well is the presumption that solid research exists to guide the content and structure of pre-K programs. Despite more than 50 years of preliminary work on preschools as an early intervention for young children from poor backgrounds, the field of early childhood education has a relatively small database to use as a guide to effective practice. One reason for this is that the first 20 years were focused on IQ gains as the measure of programmatic success. Those early studies continue to be used to justify policies for massive expansion of pre-K despite the fact that the goal of present day programs is school readiness rather than IQ gains, and ignoring the fact that little about the content of those early demonstration projects is being replicated in today's pre-K programs.

The needs of young children are just as great—maybe greater—as they were in 1965 when Head Start began. The approach the field should take in response is to begin a rigorous research effort to determine which malleable competencies in early childhood are most related to the developmental trajectories of poor children, which experiences within pre-K settings actually facilitate the development of those skills, and how the success of pre-K programs in transmitting those skills can be validly measured for purposes of accountability and improvement. None of these critical areas is well understood today.

<sup>i</sup> 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).

<sup>ii</sup> Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology*, 43, 1428-1446; 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.

<sup>iii</sup> Zigler, E. & Anderson, K. (1979). An idea whose time had come: The intellectual and political climate. In E. Zigler & J. Valentine (Eds), *Project Head Start: A legacy of the war on poverty*. (pgs. 3-20). New York, NY: The Free Press.

<sup>iv</sup> Skeels, H.M. & Dye, H.R. (1939). A study of the effects of differential stimulation on mentally retarded children. *Proceedings of the American Association of Mental Deficiency*, 44, 114-136.

<sup>v</sup> Consortium for Longitudinal Studies (1983). *As the twig is bent: Lasting effects of preschool programs*. Hillsdale, NJ: Lawrence Erlbaum.

<sup>vi</sup> "Nurturing intelligence," *Time Magazine*, Monday, January 03, 1972.



- vii Schweinhart, L. J. (2002). How the High/Scope Perry Preschool study grew: A researcher's tale. *Phi Delta Kappa Center for Evaluation, Development, and Research* (No. 32). Available online: <http://www.highscope.org/Content.asp?ContentId=232>.
- viii Lazar, I. & Darlington, R. (1982). Lasting effects of early education: A report from the Consortium for Longitudinal Studies. *Monographs of the Society for Research in Child Development*, Serial No. 195, Vol. 47, Nos. 2-3.
- ix Ramey, C., Yeates, K. & Short, E. (1984). The plasticity of intellectual development: Insights from preventative intervention. *Child Development*, 55, 1913-1925.
- x Breitmayer, B. & Ramey, C. (1986). Biological nonoptimality and quality of postnatal environment as codeterminants of intellectual development. *Child Development*, 57, 1151-1165.
- xi 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.
- xii Farran, D. & Lipsey, M. (October 8, 2015). Expectations of sustained effects from scaled up pre-K: Challenges from the Tennessee study. *Evidence Speaks*, 1(4). Washington, DC: Brookings. <http://www.brookings.edu/research/papers/2015/10/08-expectations-of-sustained-effects-from-scaled-up-prek-challenges-tennessee-study-farran-lipsey>.
- xiii Farran, D.C. (2005). Developing and implementing preventive intervention programs for children at risk: Poverty as a case in point. In M. Guralnik (Ed.), *A developmental systems approach to early intervention: National and international perspectives* (pp. 267-304). Baltimore: Paul Brookes, Publisher.
- xiv Watts, T., Duncan, G., Siegler, R., & Davis-Kean, P. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, 43, 352 -360. DOI: 10.3102/0013189X14553660.
- xv Noble, K., Houston, S., Brito, N., Bartsch, H., Kan, E., Kuperman, J., Sowell, E. (2015). Family income, parental education and brain structure in children and adolescents. *Nature Neuroscience*, 18, 773-778. doi:10.1038/nn.3983.
- xvi Farah, M. & Hackman, D. (2012). SES, childhood experience, and the neural basis of cognition. In V. Maholmes & R. King (Eds.), *The Oxford handbook of poverty and child development*. (pp. 307-318). New York, NY: Oxford University Press.
- xvii Burchinal, M, Raver, M., Vernon-Feagon, L., Cox, M., & Blair, C (2015). Poverty and early child development in rural low-income regions: Results from the Family Life Project. In E. Votruba-Drzal. *Place, Poverty, and child development*. Symposium presented at the Biennial Meeting of the Society for Research in Child Development, Philadelphia, PA.
- xviii Fernald, A., Marchman, V., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, 16, 234-248. DOI: 10.1111/desc.12019.
- xix Barnett, S., Carolan, M., Squires, J., Brown, K., & Horowitz, M. (2014). The state of preschool 2014: State preschool yearbook. National Institute for Early Education Research (NIEER), Graduate School of Education, Rutgers, the State University of New Jersey.
- xx <http://tnstarquality.org/html/star-quality.htm>.
- xxi <https://eclkc.ohs.acf.hhs.gov/hslc/standards/hsp/1307;YXN0cm9uZw==#1307.3>.
- xxii <http://eclkc.ohs.acf.hhs.gov/hslc/grants/dr/cohort-4-listsby-state.html>.
- xxiii Mashburn, A., Pianta, R., Hamre, B., Downer, J., Barbarin, O., Bryant, D., Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79, 732-749.
- xxiv Keys, T., Farkas, G., Burchinal, M., Duncan, G., Vandell, D., Li, W., Howes, C. (2013). Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics. *Child Development*, 84, 1171-1190. DOI: 10.1111/cdev.12048.
- xxv Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, 84, 2112-2130. doi.org/10.1016/j.ecresq.2012.12.002.
- xxvi Gordon, R., Hofer, K., Fujimoto, K., Risk, N., Kaestner, R. & Korenman, S. (2015). Identifying high-quality preschool programs: New evidence on the validity of the Early Childhood Environment Rating Scale-Revised (ECERS-R) in relation to school readiness goals, early education and development. *Early Education and Development*, 26, 1086-1110. DOI:10.1080/10409289.2015.1036348.