

Early Childhood Education at Scale: Lessons From Research for Policy and Practice

Daphna Bassok

University of Virginia

Mimi Engel

University of Colorado Boulder

This AERA Open special topic takes stock of the current knowledge about scaled-up early childhood education (ECE) programs. Rapidly rising ECE participation rates in recent decades have coincided with wide variability in the types of programming available for young children. This expansion makes understanding heterogeneity in effects of ECE crucial. Three articles in this special topic explore heterogeneity in ECE program effects with regard to timing, context, domains, programmatic features, counterfactual conditions, and child characteristics. In addition to exploring heterogeneity, understanding possible mechanisms for improving ECE experiences and outcomes for children and their families is crucial. Articles in this special topic explore three potential drivers for improving ECE services and outcomes, including professional development, curriculum and instruction, and parental supports for ECE educators.

Keywords: *achievement, curriculum, early childhood, educational policy, mixed methods, parents and families, professional development*

When education researchers and policymakers discuss “scale-up,” they often want to know whether a program that worked well in one context can be made widely available while retaining its benefits. Discussing effective scale-up in early childhood education (ECE) is different. In the United States, ECE is *already* operating at scale. Fueled by increased knowledge regarding young children’s development (National Research Council, 2000), interest in ameliorating socioeconomic and racial achievement gaps early (Bassok, Finch, Lee, Reardon, & Waldfogel, 2016), and the needs of working families, scaled-up ECE has become the norm.

In 1985, just over 150,000 children were enrolled in pre-kindergarten. By 2016, that figure had increased nine-fold to almost 1.4 million (Snyder, de Brey, & Dillow, 2018). Multiple states and cities now offer large or universal public preschool programs. Across the United States, over 65% of 4-year-old children were enrolled in some type of center-based care, including child care, Head Start, or prekindergarten, in 2016. And while participation rates are lower among younger children, roughly a third of all children ages 0 to 6 (not yet enrolled in kindergarten) are enrolled in some type of center-based ECE setting (Snyder et al., 2018).

The scale at which ECE operates today implies is time to move away from the question, “Do ECE programs have

positive impacts?” Just as we do not debate whether fourth grade “works,” asking whether ECE programs, broadly defined, work is counterproductive. This is not because the evidence on the impacts of ECE is so unequivocal and so compelling that no further research is needed. And it is not because we should simply assume that increased public investment in ECE will yield benefits for children and society. Rather, unidimensional questions about whether ECE programs work are no longer practical at a time when out-of-home care is a matter of course for most young children in the United States.

About one-third of mothers with children under 3 were employed in 1975. In contrast, over 60% were in the labor force in 2015 (U.S. Department of Labor, 2016). Further, in 2017, a quarter of children were living with one parent, compared with 13% in 1968 (Livingston, 2018). ECE research needs to acknowledge this new status quo. The questions worth asking now are not about *whether* ECE works. Rather, we must explore the conditions under which ECE programs are most effective, for which children, and through what mechanisms.

The goal for this *AERA Open* special topic is to inform this process by taking stock of what we have learned thus far from the rapid scale-up of ECE programs in the United



States and elsewhere. We explore the implications of these findings for researchers, policymakers, and practitioners while considering areas where further research is most needed.

When developing this special topic, we sought two types of papers. First, we looked for contributions that synthesized the evidence from scaled-up ECE programs such as the federal Head Start program, state prekindergarten initiatives, and other large-scale implementations of ECE programming. The articles in this category highlight the promise of ECE programs while also emphasizing the vast heterogeneity in both the types of experiences ECE programs provide and the impacts these programs have on children's development. They suggest a need for more nuanced and specific research addressing under what conditions ECE works and for whom.

In addition to exploring the effects of varied ECE programming, we sought a second set of papers that dove deeply into promising areas of focus for policymakers as they consider large-scale ECE expansion and improvement efforts. While measuring the effects of programs on children is essential, so is rigorous work examining strategies to foster improvement. There is strong consensus on the importance of ECE "quality" broadly defined. However, as a field, we often lack precision and specificity regarding which malleable factors are most likely to move the needle to improve young children's ECE experiences and outcomes. This special topic includes careful examinations of three potential levers for improvement: professional development to support the ECE workforce, curriculum and instruction, and parental engagement.

Taken together, the articles in this special topic report and synthesize evidence on the effects of recent, large-scale preschool initiatives and explore promising strategies for improving the quality of scaled-up preschool programming. In this introduction, we summarize current evidence on scaled-up ECE programs in the United States and introduce the six papers that comprise this special topic.

What We Know About the Effects of ECE at Scale

The evidence most often cited in support of increased public investment for ECE comes from rigorous studies examining the short- and long-term impacts of small-scale interventions that took place decades ago. In his 2013 State of the Union address, President Obama noted that, "Every dollar we invest in high-quality early education can save more than seven dollars later on—by boosting graduation rates, reducing teen pregnancy, even reducing violent crime" (*The Atlantic*, 2013). The experimental studies alluded to in that speech (e.g., Belfield, Nores, Barnett, & Schweinhart, 2006; Campbell et al., 2012) indicate that children randomly assigned to intensive preschool programs experienced positive long-term outcomes, including

increased rates of educational attainment and employment as well as higher earnings.

While compelling, the "model" programs described in those studies were implemented half a century ago. Participants attended preschool during the 1960s and 1970s. Much has changed in the interim with respect to young children's lives in general and their ECE experiences in particular. A growing body of research has evaluated more recent and larger scale programs, which are typically both less intensive and less expensive than the model programs often cited to tout the benefits of ECE participation (Camilli, Vargas, Ryan, & Barnett, 2010; Yoshikawa et al., 2013).

Large-scale state prekindergarten programming is a relatively new phenomenon. Between 2002 and 2017, state spending on preschool rose from \$2.4 billion to \$7.6 billion. Adjusting for inflation, the change is nearly \$4 billion (Friedman-Krauss et al., 2018). Reviewing the relatively large evidence base on the short-term impacts of state prekindergarten programs, a recent consensus statement from a group of ECE experts concludes that on average, children who attend prekindergarten enter kindergarten with stronger readiness skills, particularly with respect to academics, than their peers who do not attend (Phillips et al., 2017). Similar patterns of short-term benefits are commonly observed for Head Start and other ECE settings (Bassok, Gibbs, & Latham, 2018; Puma et al., 2010).

A more limited set of studies has examined the medium and longer term impacts of scaled-up ECE programming, including both Head Start and prekindergarten. Here, results are mixed. Rigorous quasi-experimental studies indicate a host of benefits of Head Start participation for children and their families, including improvements in parenting practices (Bauer & Schanzenbach, 2016; Gelber & Isen, 2013), reduced incidence of behavioral and health problems (Carneiro & Ginja, 2014), and increases in participants' educational attainment and economic well-being (Bauer & Schanzenbach, 2016; Deming, 2009). Because state prekindergarten programs are a relatively recent addition to the ECE landscape, longer term impact analyses of these initiatives are limited. However, a number of studies suggest persistent benefits from state prekindergarten participation on children's achievement outcomes in elementary and middle school (Cascio & Schanzenbach, 2013; Fitzpatrick, 2008; Ladd, Muschkin, & Dodge, 2014; Phillips, Gormley, & Anderson, 2016).

Importantly, however, results from the only two experimental studies tracking the impacts of scaled-up ECE programs in the United States through elementary school are discouraging. Findings from the national Head Start Impact Study (HSIS) showed initial benefits for children who were randomly assigned to Head Start (Puma et al., 2010). However, by third grade, on average, there were no differences between the treatment and control groups (Puma et al., 2012). Similarly, results from a lottery-based experiment to

evaluate Tennessee's prekindergarten program indicate that by third grade, preschool participants were performing at the same level as, or below, their peers who had not been randomly assigned to the program (Lipsey, Farran, & Durkin, 2018).

Given their rigor, scope, and lack of positive effects, these two large studies have garnered substantial attention, yielding commentaries, policy briefs, and media articles with titles such as, "Are the Effects of State Pre-K Overrated?" (Samuels, 2018); "Preschool Is Not Magic" (Slavin, 2018); and "Trouble in the Land of Early Childhood Education?" (Haskins & Brooks-Gunn). These commentaries sensibly caution against viewing ECE programs as a silver bullet. The results from these experiments, coupled with a large set of studies documenting rapid fade-out of ECE effects, suggest that it is unrealistic to expect participation in a single year of "typical" ECE programming to meaningfully alter young children's learning trajectories (Bassok et al., 2018; Leak et al., 2012).

It would be unwise, however, to use these findings to justify reductions in public investment for ECE. Recently, a number of thoughtful commentaries have wrestled with this complicated evidence base to consider the best path forward (e.g., Jones & Lesaux, 2018; Weiland, 2018). Researchers need to focus on understanding the conditions under which ECE programs yield positive and persistent results. Doing so will support policymakers in their efforts to better align future investments with programs and practices that have the most potential to benefit young children.

New Lessons From Scaled-Up ECE Initiatives

The first two articles in this special topic are about the impact of large-scale public preschools programming with a focus on exploring heterogeneity in effects.

In "New Findings on Impact Variation From the Head Start Impact Study: Informing the Scale-Up of Early Childhood Programs," Morris et al. (this special topic) synthesize results from a set of studies produced by the Secondary Analysis of Variation in Impacts (SAVI) Center, which was created to support researchers in reanalyzing data from the HSIS to explore heterogeneity in program impacts. The authors note that much of the research on Head Start and other ECE programs focuses on *average* treatment effects. When variation in program impacts is examined, the focus is often on child or family characteristics as drivers of variability (e.g., exploring whether effects are larger for children from families with the lowest income) (Bitler, Hoynes, & Domina, 2014; Puma et al., 2010).

They then highlight the more limited body of research exploring heterogeneity in ECE effects by program characteristics, neighborhood context, and the counterfactual experience (Kline & Walters, 2016; Walters, 2015; Zhai, Brooks-Gunn, & Waldfogel, 2014). Morris and colleagues

make several key points with regard to heterogeneity in ECE effects. First, there is substantial variation across Head Start programs with respect to effects on three of four cognitive outcomes considered. Thus, even within a federal program with many explicit regulations regarding program operations and program quality, impacts vary substantially. Second, program impacts differ based on child characteristics. Estimated effects are larger for children who are dual language learners/Spanish speakers and children who enter Head Start with lower baseline skills. Estimated effects also vary across urban and rural Head Start settings.

Finally, Head Start impacts vary substantially depending on the counterfactual conditions used for comparison. Using a principal stratification framework, Morris and colleagues find no evidence that Head Start yields benefits for children who would have otherwise attended another center-based ECE option. However, among children who would have attended a home-based setting in the absence of an offer to attend Head Start, the benefits of the program on receptive vocabulary were meaningful. These findings underscore the importance of focusing on the "counterfactual" ECE options. The authors highlight the need to target scarce resources to communities and age groups for whom there are few center-based ECE alternatives. Expanding access to publicly-funded ECE programs for the youngest children, *ages 0 through 3*, may be particularly useful given that in many contexts, very few formal options are available for this age group.

In "State Prekindergarten Effects on Early Learning at Kindergarten Entry: An Analysis of Eight State Programs," Barnett et al. (this special topic) also explore heterogeneity in ECE program impacts, focusing on state-funded prekindergarten programs. Currently, 43 states are implementing some type of state-funded programming for children ages 3 to 4. Thus, understanding both average effects of state-funded preschool as well as the extent to which effects vary by context and across learning domains is crucial. This evidence is needed for guiding legislation and administrative decision making toward programmatic choices that are likely to produce the largest and most lasting benefits. Building on Wong, Cook, Barnett, and Jung (2008), Barnett and colleagues estimate the effects of 1 year of state-funded prekindergarten programming for children at age 4 across eight states with widely varying participation rates and programmatic components. Using an age-cutoff regression discontinuity approach, they assessed the short-term impacts of programs on children's language, early mathematics skills, and early literacy skills at school entry.

The average effect sizes for language, math, and literacy across state contexts were 0.24, 0.45, and 1.1, respectively. While the positive impacts are encouraging, the differences in effect sizes across domains are striking. The study also highlighted wide variation in impacts across state contexts. For example, the authors report an effect size of 1.1 for

math in Michigan, compared to a statistically insignificant effect size of 0.24 for math in New Mexico. Similarly, the effect of state prekindergarten on literacy at kindergarten entry in West Virginia was 1.72, compared with 0.5 in New Jersey.

Barnett and colleagues note that state-funded prekindergarten programs vary tremendously in terms of per-pupil expenditures, programmatic components, and population served, among other areas. Thus, the substantial variation they find across contexts is not surprising. Their conclusions echo the recent consensus statement on the state of knowledge with regard to state prekindergarten, which notes that programming is so widely varied that it should not be considered as a singular entity (Phillips et al., 2017).

The wide variability in program impacts indicates that effects of state prekindergarten in one context or a single domain are unlikely to generalize. Continued efforts to implement state prekindergarten should be cognizant of this wide variability. Additional research should unpack the mechanisms that explain heterogeneity in effect size estimates across contexts and outcome areas.

A third paper exploring the impacts of a scaled-up ECE program uses data from Norway to provide new evidence on the effects of scaled-up ECE programs serving infants and toddlers. In the United States, most publicly funded ECE programs serve children who are between 3 and 5 years old. These programs often target children from families with low income. In contrast, Norway has gradually scaled up a universal ECE program for children beginning at age 1. In “Estimating the Consequences of Norway’s National Scale-Up of Early Childhood Education and Care (Beginning in Infancy) for Early Language Skills,” Dearing, Zachrisson, Mykletun, and Toppelberg (this special topic) provide new evidence on the effects of a large-scale ECE program on the development of very young children. The authors use population data to estimate the effect of the program on the language skills of infants and toddlers, disaggregating their findings for children from households with low, middle, or high income. Results indicate that providing universal access to ECE for younger children in Norway was positively associated with their early language skills and that effects were over twice as large for children from low-income families relative to children from middle- or high-income families.

Generalizing from these results to the United States is potentially problematic for a variety of reasons including how far the United States lags behind Norway with respect to the provision of public programming and supports for families with infants and toddlers. Still, the findings from this study highlight heterogeneity in treatment effects based on families’ socioeconomic status, results that are consistent with evidence on ECE in the United States (Bitler et al., 2014; Puma et al., 2010). Further, Dearing and colleagues’ work suggests that providing high-quality ECE for children under age 3 may hold potential for ameliorating socioeconomic achievement

gaps. This is an important result in light of evidence showing socioeconomic gaps are observed as early as 9 months (Halle et al., 2009) and that these gaps have widened substantially in recent decades in the United States (Reardon, 2011).

It is worth noting that the evidence base on the effects of ECE for infants and toddlers is mixed. Using experimental data from a high-quality infant and toddler program, Duncan and Sojourner (2013) project that high-quality care for children ages 1 to 3 has the potential to dramatically reduce income-based gaps in achievement. In contrast, quasi-experimental evidence on universal early care in Quebec suggests negative impacts that persist as children progress through school, with the most recent evidence suggesting negative effects on health outcomes and crime rates (Baker, Gruber, & Milligan, 2017; Japel, Tremblay, & Cote, 2005). It is important to note that the care provided in Quebec was considered to be of low quality.

Taken together, these three studies highlight that the benefits of ECE vary substantially based on child characteristics (e.g., primary home language, socioeconomic status), contextual factors (e.g., urbanicity, availability of ECE alternatives), and outcome domains (e.g., mathematics vs. language). Fully unpacking and understanding this heterogeneity is essential for designing ECE policies that maximize benefits. At the same time, it is crucial to consider which malleable factors of ECE programs should be altered through policy to help improve program effectiveness. The remaining three papers in the special topic address this issue.

Levers for Improving ECE at Scale

There is widespread agreement that ensuring ECE programs are of high quality is essential. Further, there is consensus that many of the scaled-up ECE programs that currently serve young children may not be providing experiences that are likely to foster sustained benefits (Yoshikawa et al., 2013). However, there is surprisingly little consensus on the specific characteristics or combinations of programmatic features that are most essential for ensuring the effectiveness of ECE programs. This lack of clarity is a major barrier for policymakers seeking to mandate or incentivize means for systematically improving ECE experiences and outcomes.

Quality in ECE has typically been operationalized via structural features of programs and classrooms (e.g., student-teacher ratios, teacher education levels). Recently, recognition of the importance of more process-oriented ECE features such as responsive teacher-child interactions or high-fidelity use of engaging and well-designed curricula has increased. However, the ways in which important programmatic components are operationalized or measured in policy and practice are often poorly aligned with evidence on best practices. Thus, the second set of articles in this special topic address three aspects of ECE programming that

have potential to improve quality: professional development to support the ECE workforce, curricula and instruction, and parental engagement. The three articles summarize current knowledge on these topics and suggest avenues for increasing the use of evidence-based policies and practices at scale. These articles also provide suggestions for areas for further research regarding programmatic components to target for improving quality in ECE.

In “Enhancing the Impact of Professional Development in the Context of Preschool Expansion,” Hamre, Partee, and Mulcahy (this special topic) synthesize the evidence on professional development for ECE teachers, making recommendations about how best to support this workforce in scaled-up ECE programs. Young children’s experiences in their classrooms and thus the educators who plan and facilitate their learning opportunities are essential to ensuring positive effects of ECE. However, the ECE workforce is characterized by low levels of education and pay and high levels of turnover. In turn, many ECE classrooms fail to provide the types of engaging and supportive interactions young children need to thrive.

Hamre et al. note that although there are rigorous studies showing that professional development initiatives *can* impact both teaching practices and child outcomes, the types of professional development that most ECE teachers experience are unlikely to foster improvement, particularly when implemented at scale. The authors note that most ECE educators receive professional development that is not focused on evidence-based teaching practices, is insufficient in terms of duration and intensity, and is not presented in a format that is likely to support sustained changes in teaching practices. These findings are similar to evidence on the lack of effectiveness of the vast majority of professional development experiences in K–12 settings (e.g., Hill, Beisiegel, & Jacob, 2013; Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Hamre and colleagues suggest concrete policy changes to help ameliorate these problems. Their recommendations include changing professional development regulations to focus on the type and quality of professional development offered (rather than just the quantity), providing more supports to ECE program leaders to help them facilitate staff development, and creating systems to certify the quality of ECE professional development providers.

In a related article titled “Preschool Curricula and Professional Development Features for Getting to High-Quality Implementation at Scale: A Comparative Review Across Five Trials,” Weiland, McCormick, Mattered, Maier, and Morris (this special topic) articulate an approach to improving instructional quality in ECE programs operating *at scale*. While Hamre et al., which focuses broadly on policy changes needed to support ECE teachers’ professional development, Weiland et al. explore microlevel features of curricula and instruction that have the potential to facilitate effective teaching practices in ECE at scale. They

make the case for domain-specific, play-based curricula coupled with access to and regular contact with coaches who observe teachers and support their efforts to implement the curricula with high levels of fidelity. The authors, who were directly involved in five large-scale ECE studies that used domain-specific curricula combined with training and coaching, conduct a cross-study comparative review to identify common supports that may foster high-fidelity implementation.

They identify several promising programmatic features, including the availability of detailed scripts for classroom teachers to use during instruction; opportunities for teachers to voice their perspectives and concerns to coaches rather than a more traditional, top-down, coaching model; the use of real-time data for monitoring instruction and providing feedback; and the provision of shared planning time for teachers. The authors do not claim that these features are causally linked to improvements in ECE teaching practices or child outcomes. Rather, they highlight them as implementation features worthy of more attention from both researchers and practitioners eager to push forward on taking evidence-based practices to scale.

Finally, in “A New Approach to Defining and Measuring Family Engagement in Early Childhood Education Programs,” Sabol, Sommer, Sanchez, and Busby (this special topic) address the disconnect between the evidence on effects of parental involvement in ECE settings and the way family engagement is oftentimes operationalized and incentivized through ECE policy and practice. Based on their review of the literature, the authors note that if a goal of ECE programming is to boost child outcomes through family engagement, the best way to do so is through the provision of services that meet parents’ needs and help build human and social capital. Sabol and colleagues provide examples of studies that document how direct provision of services to parents can benefit children and families.

In contrast to the examples they provide, however, the authors note that family engagement in ECE is often operationalized to focus on activities such as volunteer opportunities in classrooms, attendance at parent-teacher conferences, and the availability of a parent handbook. These features have not been shown to measurably benefit either parental or child outcomes. Nearly all state Quality Rating and Improvement Systems (QRIS), accountability systems aimed at incentivizing improvements in ECE settings, include these types of measures of parental engagement. While these aspects of parental involvement are relatively easy to measure, and potentially beneficial, they are not types of parental supports that are most likely to improve child and family outcomes.

Sabol and colleagues make concrete recommendations about how to move QRIS away from simplistic measurement of parental involvement. Their suggestions focus on assessing parental needs through questionnaires and focus groups to

facilitate the provision of targeted, evidence-based services for parents aimed at meeting their most pressing needs. A starting point for maximizing the ways that QRIS can be used to better support parents of ECE participants is to develop QRIS measures around parental participation that are likely to tap whether and how ECE programs are assessing and making efforts to help meet those needs through either direct service provision or referrals stemming from the development of robust networks of local social service providers.

Conclusion

The evidence presented in this special topic highlights that the effects of ECE are highly variable even within program types. The temptation to draw a brief, clear, and resounding conclusion about ECE is misguided. Early childhood education is neither a silver bullet nor a poor use of public dollars. Rather, it is a large and growing sector of education.

With rising participation rates and increased investment, the vast majority of children in the United States now participate in some type of early childhood programming prior to entering kindergarten. Thus, exploring the programmatic, contextual, and individual factors that are likely to result in the largest benefits to children, their families, and society is crucial for making evidence-based decisions about the future of ECE.

The wide heterogeneity in effects of ECE as well as the limitations of our current knowledge regarding mechanisms and processes for improving ECE quality and outcomes present a vast array of questions for researchers to tackle. In the meantime, careful and systematic review of existing evidence should guide policy and decision making around ECE to potentiate positive results for young children and their families. ECE programming should be aligned as closely as possible with the best evidence from research.

References

- The Atlantic*. (2013). Obama's 2013 State of the Union speech: full text. Retrieved from <https://www.theatlantic.com/politics/archive/2013/02/obamas-2013-state-of-the-union-speech-full-text/273089/>
- Baker, M., Gruber, J., & Milligan, K. (2017). *Long run impacts of a universal child care program*. Retrieved from <http://faculty.arts.ubc.ca/kmilligan/research/papers/BGM-Final.pdf>.
- Barnett, W. S., Jung, K., Friedman-Krauss, A., Frede, E. C., Nores, M., Hustedt, J. T., . . . Daniel-Echols, M. (2018). State prekindergarten effects on early learning at kindergarten entry: An analysis of eight state programs. *AERA Open*, 4(2). doi:10.1177/2332858418766291
- Bassok, D., Finch, J. E., Lee, R., Reardon, S. F., & Waldfogel, J. (2016). Socioeconomic gaps in early childhood experiences: 1998 to 2010. *AERA Open*, 2(3). doi:10.1177/2332858416653924.
- Bassok, D., Gibbs, C. R., & Latham, S. (2018). Preschool and children's outcomes in elementary school: Have patterns changed nationwide between 1998 and 2010? [Advance online publication]. *Child Development*. doi: 10.1111/cdev.13067
- Bauer, L., & Schanzenbach, D. W. (2016). *The long-term impact of the Head Start program*. Washington, DC: The Hamilton Project.
- Belfield, C. R., Nores, M., Barnett, S., & Schweinhart, L. (2006). The High/Scope Perry preschool program cost-benefit analysis using data from the age-40 followup. *Journal of Human Resources*, 41(1), 162–190.
- Bitler, M. P., Hoynes, H. W., & Domina, T. (2014). *Experimental evidence on distributional effects of Head Start* (No. w20434). Cambridge, MA: National Bureau of Economic 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(3), 579–620.
- Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., . . . Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: An Abecedarian Project follow-up. *Developmental Psychology*, 48(4), 1033–1043.
- Carneiro, P., & Ginja, R. (2014). Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start. *American Economic Journal: Economic Policy*, 6(4), 135–173.
- Cascio, E. U., & Schanzenbach, D. W. (2013). *The impacts of expanding access to high-quality preschool education* (No. w19735). Cambridge, MA: National Bureau of Economic Research.
- Dearing, E., Zachrisson, H. D., Mykletun, A., & Toppelberg, C. O. (2018). Estimating the consequences of Norway's national scale-up of early childhood education and care (beginning in infancy) for early language skills. *AERA Open*, 4(1). doi:10.1177/2332858418756598.
- Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3), 111–134.
- Duncan, G. J., & Sojourner, A. J. (2013). Can intensive early childhood intervention programs eliminate income-based cognitive and achievement gaps? *Journal of Human Resources*, 48(4), 945–968.
- Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement. *The BE Journal of Economic Analysis & Policy*, 8(1).
- Friedman-Krauss, A. H., Barnett, W. S., Weisenfeld, G. G., Kasmin, R., DiCrecchio, N., & Horowitz, M. (2018). *The state of preschool 2017: State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research.
- Gelber, A., & Isen, A. (2013). Children's schooling and parents' behavior: Evidence from the Head Start Impact Study. *Journal of Public Economics*, 101, 25–38.
- Halle, T., Forry, N., Hair, E., Perper, K., Wandner, L., Wessel, J., & Vick, J. (2009). *Disparities in early learning and development: Lessons from the Early Childhood Longitudinal Study—Birth Cohort (ECLS-B)*. Bethesda, MD: Child Trends.
- Hamre, B. K., Partee, A., & Mulcahy, C. (2017). Enhancing the impact of professional development in the context of preschool expansion. *AERA Open*, 3(4). doi:10.1177/2332858417733686
- Haskins, R., & Brooks-Gunn, J. (2016). *Trouble in the land of early childhood education*. Retrieved from https://futureofchildren.princeton.edu/sites/futureofchildren/files/media/starting_early_26_2_policy_brief.pdf

- Hill, H. C., Beisiegel, M., & Jacob, R. (2013). Professional development research: Consensus, crossroads, and challenges. *Educational Researcher*, 42(9), 476–487.
- Japel, C., Tremblay, R. E., & Cote, S. (2005). Quality counts: Assessing the quality of daycare services based on the Quebec Longitudinal Study of Child Development. *Choices*, 11, 1–42.
- Jones, S., & Lesaux, N. (2018). Asking if early-childhood education “works” is the wrong question. Here are the right ones. *The Washington Post*. Retrieved from https://www.washingtonpost.com/education/2018/12/27/asking-if-early-childhood-education-works-is-wrong-question-here-are-right-ones/?noredirect=on&utm_term=.52b9549606dc
- Kline, P., & Walters, C. R. (2016). Evaluating public programs with close substitutes: The case of Head Start. *The Quarterly Journal of Economics*, 131, 1795–1848. doi:10.1093/qje/qjw027
- 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(1), 162–187.
- Leak, J., Duncan, G., Li, W., Magnuson, K., Schindler, H., & Yoshikawa, H. (2012). *Is timing everything? How early childhood education program cognitive and achievement impacts vary by starting age, program duration and time since the end of the program*. Unpublished manuscript.
- Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children’s achievement and behavior through third grade. *Early Childhood Research Quarterly*, 45, 155–176.
- Livingston, G. (2018). *About one-third of U.S. children are living with a single parent*. Retrieved from <http://www.pewresearch.org/fact-tank/2018/04/27/about-one-third-of-u-s-children-are-living-with-an-unmarried-parent/>
- Morris, P. A., Connors, M., Friedman-Krauss, A., McCoy, D. C., Weiland, C., Feller, A., . . . Yoshikawa, H. (2018). New findings on impact variation from the Head Start Impact Study: Informing the scale-up of early childhood programs. *AERA Open*, 4(2). doi:10.1177/2332858418769287
- National Research Council. (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: Author.
- Phillips, D., Gormley, W., & Anderson, S. (2016). The effects of Tulsa’s CAP Head Start program on middle-school academic outcomes and progress. *Developmental Psychology*, 52(8), 1247–1281.
- Phillips, D., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., & Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge on pre-kindergarten effects—A consensus statement*. Retrieved from <https://www.brookings.edu/research/puzzling-it-out-the-current-state-of-scientific-knowledge-on-pre-kindergarten-effects/>
- Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., & . . . Downer, J. (2012). *Third grade follow-up to the Head Start Impact Study: Final report*. OPRE Report 2012-45. Washington, DC: Administration for Children & Families.
- Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., . . . Spier, E. (2010). *Head Start Impact Study: Final report*. Washington, DC: Administration for Children and Families.
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity* (pp. 91–116). New York, NY: Russell Sage.
- Sabol, T. J., Sommer, T. E., Sanchez, A., & Busby, A. K. (2018). A new approach to defining and measuring family engagement in early childhood education programs. *AERA Open*, 4(3). doi:10.1177/2332858418785904
- Samuels, C. A. (2018). Are the effects of state pre-k overrated? *Education Week*. Retrieved from <https://www.edweek.org/ew/articles/2018/07/18/are-the-effects-of-state-pre-k-overrated.html>
- Slavin, R. (2018). *Preschool is not magic. Here’s what is*. Retrieved from <https://robertslavinsblog.wordpress.com/2018/08/02/preschool-is-not-magic-heres-what-is/>
- Snyder, T. D., de Brey, C., & Dillow, S. A. (2018). *Digest of education statistics 2016* (NCES 2017-094). Retrieved from <https://nces.ed.gov/pubs2017/2017094.pdf>
- U.S. Department of Labor. (2016). *Working mothers issue brief*. Retrieved from https://www.dol.gov/wb/resources/WB_WorkingMothers_508_FinalJune13.pdf
- Walters, C. R. (2015). Inputs in the production of early childhood human capital: Evidence from Head Start. *American Economic Journal: Applied Economics*, 7(4), 76–102.
- Weiland, C. (2018). Pivoting to the “how”: Moving preschool policy, practice, and research forward. *Early Childhood Research Quarterly*, 45, 188–192.
- Weiland, C., McCormick, M., Matterna, S., Maier, M., & Morris, P. (2018). Preschool curricula and professional development features for getting to high-quality implementation at scale: A comparative review across five trials. *AERA Open*, 4(1). doi:10.1177/2332858418757735
- 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(1), 122–154.
- Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (REL 2007-No. 033). Washington, DC: Institute of Education Sciences.
- Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., . . . Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. Ann Arbor, MI: Society for Research in Child Development.
- Zhai, F., Brooks-Gunn, J., & Waldfogel, J. (2014). Head Start’s impact is contingent on alternative type of care in comparison group. *Developmental Psychology*, 50, 2572–2586.

Authors

DAPHNA BASSOK is an associate professor of education and public policy at the University of Virginia and is also the associate director of EdPolicyWorks, a joint collaboration between the Curry School of Education and the Frank Batten School of Leadership and Public Policy. Her research addresses early childhood education policy, with a particular focus on the impacts of policy interventions on the well-being of low-income children.

MIMI ENGEL is an associate professor of research and evaluation methodology in the School of Education at the University of Colorado Boulder. She studies educational policies and practices that have the potential to improve student outcomes, particularly among youth from traditionally underserved populations.