



MONEYBALL *for* HEAD START:

Using Data, Evidence, and Evaluation to Improve Outcomes for Children and Families

Sara Mead and Ashley LiBetti Mitchel

JANUARY 2016



RESULTS FOR AMERICA

Results for America (RFA) is improving outcomes for young people, their families, and communities by shifting public resources toward evidence-based, results-driven solutions. RFA is working to achieve this goal through a three-part strategy: building a strong, bipartisan Invest in What Works coalition; developing and advancing the next generation of evidence-based, results-driven practices, policies, and programs; and supporting leaders at all levels of government to invest in what works.



Bellwether Education Partners is a nonprofit dedicated to helping education organizations in the public, private, and nonprofit sectors become more effective in their work and achieve dramatic results, especially for high-need students. To do so, we provide a unique combination of exceptional thinking, talent, and hands-on strategic support.



The National Head Start Association (NHSA) is a nonprofit organization committed to the belief that every child, regardless of circumstances at birth, has the ability to succeed in life. NHSA is the voice for more than 1 million children, 200,000 staff and 1,600 Head Start grantees in the United States. Since 1974, NHSA has worked diligently for policy changes that ensure all at-risk children have access to the Head Start model of support for the whole child, the family and the community. NHSA's vision is to lead - to be the untiring voice that will not be quiet until every vulnerable child is served with the Head Start model of support for the whole child, the family and the community - and to advocate - to work diligently for policy and institutional changes that ensure all vulnerable children and families have what they need to succeed. NHSA's mission is to coalesce, inspire, and support the Head Start field as a leader in early childhood development and education.



The Volcker Alliance was launched in 2013 by former Federal Reserve Board Chairman Paul A. Volcker to address the challenge of effective execution of public policies and to help rebuild public trust in government. The nonpartisan Alliance works toward that broad objective by partnering with other organizations—academic, business, governmental, and public interest—to strengthen professional education for public service, conduct needed research on government performance, and improve the efficiency and accountability of governmental organization at the federal, state, and local levels.

Cover Photo: Keith Dunlop

MONEYBALL *for* HEAD START:

Using Data, Evidence, and Evaluation to Improve Outcomes for Children and Families

EXECUTIVE SUMMARY

Head Start is a valuable federal program that improves the lives of our nation's most vulnerable children and their families. Research shows that Head Start programs improve children's learning at school entry and have a positive impact on long-term life outcomes. Yet research also suggests that Head Start could have a stronger impact on children's early learning, school, and life outcomes. The key question is, how can policymakers and practitioners maximize outcomes for Head Start children and their families?

This paper — the product of a combined effort of Results for America, Bellwether Education Partners, the National Head Start Association, and the Volcker Alliance — outlines a vision for a continuous improvement approach that uses data, evidence, and evaluation to improve outcomes at all levels of the Head Start program.

At the *local grantee level*, all Head Start grantees need systems of data collection and analysis that support data-informed, evidence-based continuous improvement, leading to better results for children and families. Head Start grantees collect and report data on a variety of outcomes, but effectively using this information to improve quality and outcomes requires a high level of intentionality, planning, and expertise in analyzing, interpreting, and acting on data. The best Head Start grantees already do this, but making data-informed continuous improvement the rule, rather than the exception, will require cultural shifts and increased capacity across all grantees. Federal policymakers can support these efforts by reviewing and revising reporting requirements, changing requirements that limit grantees' staffing flexibility, and allocating existing training and technical assistance funds to help build grantees' capacity to analyze and act on data. Groups of local grantees and researchers should also work together in Networked Learning Communities to share and analyze data to improve grantee practices and child outcomes.

At the *federal oversight level*, the Office of Head Start (OHS), within the Administration for Children and Families of the U.S. Department of Health and Human Services, needs a valid accountability and performance measurement system that allows federal officials to detect trends and patterns in performance; identify, learn from, and disseminate lessons from trends and patterns in performance; differentiate oversight, support, and interventions on the basis of grantee performance; and support continuous improvement across Head Start as a whole. Head Start grantees are already accountable through the Aligned Monitoring System and the Designation Renewal System. But supporting continuous improvement in Head Start will require measuring grantee quality with more results-based, robust, and fair systems that use valid, reliable, and accurate ways of measuring child outcomes; family outcomes; and indicators of program quality, safety, health, and compliance. Establishing such systems will require an iterative process that engages a variety of federal agencies, researchers, practitioners, and the

philanthropic sector to identify the outcomes or components of program performance that are most important, develop ways to measure those outcomes or performance indicators, and pilot and refine new measurement tools. Once new tools or measures are ready to be implemented at scale, OHS should give grantees time to become familiar with and buy into the new systems before attaching consequences. OHS should publicly report data on common measures of program performance using transparent, interactive data reports that enable stakeholders to access data on common program performance indicators and compare and analyze data across subsets of grantees and the program as a whole. OHS should also use these data to differentiate grantees by levels of performance — high-performing, adequately performing, and low-performing — across multiple domains.

At the *research and evaluation* level, federal policymakers and the philanthropic sector need to support research that builds the knowledge base of what works in Head Start and informs changes in program design and policies. Funding for Head Start research, demonstration, and evaluation should be increased from less than 0.25 percent of total appropriations to 1 percent, and should focus on activities that are most likely to result in building knowledge that local grantees can use to improve their quality and outcomes, specifically:

- 1. Identifying and understanding effective practices of high-performing grantees;**
- 2. supporting measured trials of new, scalable practices and approaches** to improve child and family outcomes; and
- 3. developing high-quality assessment tools and measures** that provide valid and reliable information on child and family outcomes, program quality, and other key domains of performance.

Local grantees bear the primary responsibility for integrating data-informed practice into their work. Other actors, such as the philanthropic sector, universities and other research institutions, and the private sector, can also play a crucial role in building grantee capacity or supporting the development, evaluation, and dissemination of promising practices. But federal policies play a major role in facilitating, incentivizing, or deterring local grantees' ability to shift culture and build capacity to support continuous learning and improvement. And as a result, changes in federal policies are needed.

Federal Policy Recommendations

The necessary changes — both in grantee practice and in federal oversight — cannot happen overnight. Fully realizing the vision outlined in this report will require a multiyear commitment to research and cycles of experimentation to address outstanding technical and measurement challenges. But there are steps that Congress and the administration can take now to both advance this vision in the near term and support the research needed to fully realize it in the future. These steps include:

- 1. Congress and the Secretary of Health and Human Services should make data-informed, continuous improvement a key priority in any legislative or regulatory policy action on Head Start.**
- 2. The Office of Head Start and the philanthropic sector should invest in building grantee capacity to use data to improve performance.**
- 3. Federal policymakers should initiate an iterative process to develop robust, common performance indicators for Head Start and should engage researchers, the philanthropic sector, and Head Start grantees as partners in this process.**

- 4. Federal research agencies should work with researchers and the philanthropic sector to support the development of solid, trusted metrics of Head Start child outcomes, family outcomes, and program capacity.**
- 5. The Office of Head Start should provide transparent, interactive, public reporting on grantee performance.**
- 6. Once the Office of Head Start has developed a sufficiently robust system to measure grantee performance, it should use this system to differentiate grantee performance** in order to identify high-performing grantees and learn from and scale their effective practices; support improvement in adequately performing grantees; intervene in low-performing programs; and, when necessary, defund or require them to compete for grants.
- 7. The Office of Head Start should continue, learn from, and build on efforts to make program monitoring more performance focused and less compliance oriented.**
- 8. The Secretary of Health and Human Services should implement a robust research agenda for Head Start, and Congress should increase the cap on Head Start research, demonstration, and evaluation spending from \$20 million to 1 percent of total appropriations.**
- 9. Congress should authorize the Secretary of Health and Human Services to grant additional flexibility to allow cohorts of programs working with researchers to pilot new approaches to serving children and families.**

Through these steps, federal policymakers can support Head Start grantees in using data, evidence, and evaluation to strengthen their practice and improve results for children and families.

INTRODUCTION: WHY DATA, EVIDENCE, AND EVALUATION MATTER FOR HEAD START CHILDREN AND FAMILIES

Head Start is a valuable federal program that improves the lives of our nation’s most vulnerable children and their families. Abundant evidence shows that children growing up in poverty are more likely to have adverse early childhood experiences that place them at risk for poor education and life outcomes and are less likely than their higher-income peers to have high-quality early learning experiences.¹ Head Start, which delivers early childhood education and comprehensive family services to more than 1 million preschoolers, infants, and toddlers living in poverty each year,² plays a crucial role in helping address the real challenges these children and their families face and putting them on track to succeed in school and in life.

For decades, policymakers and researchers have debated whether Head Start “works.” But that’s the wrong question. Research, including the Head Start Impact Study³ and the Head Start Family and Child Experiences Survey,⁴ shows that Head Start improves children’s learning at school entry. Although the Head Start Impact Study found that Head Start’s positive effects on test scores fade as children move through the early elementary grades, other quasi-experimental studies, which followed children into adulthood, show that Head Start alumni are more likely to graduate high school and have better adult life outcomes than similar children who did not attend Head Start.⁵ Moreover, a large and growing body of evidence shows that a range of high-quality early childhood programs can produce lasting gains in children’s school and life outcomes.⁶ Given this evidence and the pressing need to close education achievement gaps for low-income children, the key question is not whether Head Start works, but how policymakers and practitioners can maximize outcomes for these children and their families.

Federal investment in Head Start is significant — \$8.6 billion in fiscal year 2015.⁷ Given both the resources invested and the stakes for our nation’s most vulnerable children, it is crucial to invest these funds in ways that produce the best results they can for both the children and families Head Start serves and the taxpayers who fund those services.

In early 2015, four organizations — Results for America, Bellwether Education Partners, the National Head Start Association, and the Volcker Alliance — came together to explore ways to better use data, evidence, and evaluation to improve outcomes for Head Start children and families. Together, we reviewed current Head Start policies, practices, and research on effective early childhood education programs. We also brought together a wide variety of experts and stakeholders — researchers, current and former federal policymakers, and grantees themselves — to discuss the current use of data, evidence, and evaluation in Head Start and to identify opportunities for improvement. Based on this work, we examined how federal policies and program practices could apply the following three Moneyball principles, developed by Results for America, to maximize results of Head Start programs:

- » **Build** evidence about the practices, policies, and programs that will achieve the most effective and efficient results so that policymakers can make better decisions;
- » **Invest** limited taxpayer dollars in practices, policies, and programs that use data, evidence, and evaluation to demonstrate they work; and

- » **Direct** funds away from practices, policies, and programs that consistently fail to achieve measurable outcomes.⁸

While our four organizations reflect a variety of perspectives, expertise, and experience, we are united in three common beliefs that inform our joint work:

- » **Head Start plays a crucial role in the lives of our nation's most vulnerable children and families — but it could do even better.** Because of this, our work focused on how to improve Head Start, not whether the program should exist.
- » **Both investment and quality are important to improving Head Start outcomes.** The recommendations in this report focus on how to maximize the impact of existing Head Start funds, not on what the level of funding should be. Implementing some of the recommendations in this report will require increased capacity and resources at the grantee level, but this report is intentionally silent on whether this capacity should be supported by increased federal funding, philanthropic resources, or reallocation of existing federal funds. When public policymakers allocate funds based on evidence of what works, however, improving program effectiveness and demonstrating results can strengthen the case for increases in funding.
- » **Improving outcomes, rather than mere compliance, should be the focus of Head Start regulations and monitoring.** Historically, Head Start requirements and program monitoring have focused on compliance — whether grantees follow rules and spend money appropriately — rather than on the outcomes they achieve. While we may not agree on all the outcomes for which Head Start grantees should be accountable, we do agree that child and family outcomes — not compliance — must be the North Star of Head Start.

Based on these shared beliefs, and the principles outlined above, this paper presents a vision for a continuous improvement approach that uses data, evidence, and evaluation to improve outcomes at all levels of the Head Start program.

- » At the **local grantee** level, all Head Start grantees need systems of data collection and analysis that support data-informed, evidence-based continuous improvement and lead to better results for children and families. Groups of local grantees and researchers should also work together in Networked Learning Communities to share and analyze data to improve grantee practices and child outcomes.
- » At the **federal oversight** level, the Office of Head Start (OHS), within the Administration for Children and Families of the U.S. Department of Health and Human Services, needs a valid accountability and performance measurement system that allows federal officials to detect trends and patterns in performance; identify, learn from, and disseminate lessons from trends and patterns in performance; differentiate oversight, support, and interventions on the basis of grantee performance; and support continuous improvement across Head Start as a whole.
- » At the **research and evaluation** level, federal policymakers and the philanthropic sector need to support research that builds the knowledge base of what works in Head Start and informs changes in program design and policies.

In June 2015, the U.S. Department of Health and Human Services proposed revisions to the Head Start Program Performance Standards — the rules and regulations that govern Head Start programs — including significant steps to promote the use of data-informed continuous improvement. The current rule requires grantees to complete planning processes and create

plans to address key issues but does not address the outcome of those plans. The proposed rule, in contrast, emphasizes grantees' use of data to set goals for improving practice or outcomes in key areas, track progress toward goals, and make changes based on what they learn. Beyond a new section of the Performance Standards specifically focused on continuous improvement, the proposed rule integrates the use of data throughout, including in standards related to program governance, education and child development, and staff professional development. The draft Performance Standards are deeply informed by research on both how young children learn and the practices of effective early childhood programs.

The Performance Standards, however, cannot do everything required to foster a culture of continuous improvement at all levels of the Head Start program. Successfully implementing the new Performance Standards will require investments in provider and system capacity to support use of data and continuous improvement at the grantee level, as well as new strategies for monitoring programs' continuous improvement systems. Moreover, supporting data-informed continuous improvement in Head Start programs requires shifting the emphasis in grantee oversight from *compliance* to *performance*. This, in turn, necessitates changes in program accountability and performance measurement, issues the proposed changes in the Performance Standards do not address. For that reason, the analysis in this paper addresses both the proposed changes in the Performance Standards and other policy and practice changes that extend beyond the scope of the Standards themselves.

SECTION 1: USING DATA AND EVIDENCE TO SUPPORT IMPROVEMENT AT THE LOCAL GRANTEE LEVEL

At the **local grantee level**, all Head Start grantees need systems of data collection and analysis that support data-informed, evidence-based continuous improvement and lead to better results for children and families. Groups of local grantees and researchers should also work together in Networked Learning Communities to share and analyze data to improve grantee practices and child outcomes.

Use of data to inform instruction and continuous program improvement is an essential feature of early childhood education programs — including many high-performing Head Start grantees (see box 1) — that produce lasting learning gains for children.⁹ In these programs:

- » teachers use child assessment and observation data to inform and individualize instruction and identify needs for early intervention;
- » teachers and family support staff collect and analyze data on family needs, experiences, and engagement and use data to communicate with parents about their children’s progress and goals, as well as how the parents can support children’s learning;
- » center directors and coaches use attendance, teacher observation, and child assessment data to provide feedback and support to teachers and family support staff; identify children and families in need of customized support; and identify opportunities for improvement; and
- » program-level leadership uses data to identify trends and needs across programs and make adjustments to budget, staffing, and programming to improve child and family outcomes.

Effective continuous improvement systems use data from multiple sources, including child and family demographic data; child outcomes; results of health and developmental screenings; attendance records; family participation and outcomes metrics; safety and compliance information; classroom observations; professional development records; staff qualifications, retention, and job satisfaction information; and financial information. Head Start grantees already collect data on all of these measures and use this information to develop community needs assessments, comply with federal mandates, and inform their program practices, among other purposes. But simply collecting required information is not sufficient. Using this information to improve quality and outcomes requires a high level of intentionality, planning, and expertise in analyzing, interpreting, and acting on data.

Box 1: Using Data for Continuous Improvement at a Local Program Level: Community Action Project of Tulsa

CAP Tulsa is an anti-poverty organization that operates early childhood programs in Tulsa, Oklahoma. To improve program practice and performance, CAP Tulsa collects and analyzes a variety of data in three categories: child achievement, classroom quality, and program operation.

CAP Tulsa developed the Child Assessment Study to measure child achievement. Through this system, the program selects a sample of children who are assessed on a battery of assessments to measure progress in six key areas of development: applied math, literacy beyond letter recognition, expressive language, initiative, self-control, and behavior. These assessments supplement the Teaching Strategies GOLD and other assessment data that CAP teachers use to assess progress for all children. Collecting additional assessments on a sample of children provides program administrators with a more sophisticated picture of trends in child achievement and growth that they can use to inform ongoing improvement.

CAP Tulsa measures classroom quality using the Classroom Assessment Scoring System (CLASS) and the Teacher Effectiveness system. CLASS is a research-based observation instrument that assesses the quality of teacher-child interactions. The Teacher Effectiveness system is based on a rubric that CAP Tulsa developed to define, evaluate, and provide feedback on teacher performance. Using these two measures, the program tracks classroom quality data and supports individual educators in acting on the data. For example, CAP Tulsa tracks the performance of educators over time, ties goals to specific areas of the performance rubric in which the educator needs to improve, targets training and support based on those goals, and evaluates performance benchmarks as educators make progress. Program leadership can also roll up individual educator performance to identify staff development needs by site, position, or years of experience, or system-wide. Recently, the CAP Tulsa Teacher and Leader Effectiveness steering committee used high-level performance data to refine position-specific performance rubrics.

Finally, CAP Tulsa analyzes program operation data to monitor the inputs and processes behind classroom quality and teacher effectiveness. For example, the program tracks staff turnover, child participation and dosage rates, child average daily attendance, chronic absenteeism, child and family demographics, referrals to disability or mental health services, and family satisfaction with early childhood programs. These measures give context to the child achievement and classroom quality data, allowing program leadership to identify patterns and respond accordingly. For example, CAP Tulsa instituted several initiatives to improve child attendance after monitoring and benchmarking attendance data. The program now continues to monitor attendance data to see whether these new efforts are improving attendance overall.

Improved child achievement outcomes and program quality are CAP Tulsa's ultimate goals, but the program is careful to maintain a positive working environment for teachers and staff as they work toward those goals. Program leadership decided, for example, to use effectiveness data exclusively in conversations about educator growth, rather than for compensation or other personnel decisions. And although the program requires teachers to collect information on a number of different data points, leadership instituted processes that allow teachers to seamlessly collect the information without focusing less on actual teaching. Program leadership regularly conducts staff satisfaction surveys to gauge how effective they've been in their efforts. The result of these decisions is a culture that uses data to support educators, rather than induce fear.

Some Head Start grantees have created robust continuous improvement systems that they use to inform program practice and improve quality and results over time (see box 1). These systems can be models for both other grantees and the early childhood field as a whole. But not all Head Start programs have the capacity, systems, or culture to support data-informed continuous improvement. How Head Start grantees collect, analyze, and use data varies widely and depends on many factors, including the capacity and expertise of the local grantee; membership in state, regional, and national Head Start associations that facilitate relationships and resources; the efforts of Head Start state collaboration offices; the quality and relevance of training and technical assistance provided by the Office of Head Start's 10 regional offices; and access to federally funded National Training and Technical Assistance centers. State policies, programs, and systems, including quality rating and improvement systems, state longitudinal data systems, and state-funded pre-k may also affect both the resources available to grantees and the demands placed on them for the collection, reporting, analysis, and use of data.

Practitioners, policymakers, and researchers identify several barriers to using and analyzing data to improve quality and outcomes:¹⁰

- » **Lack of strong measurement tools:** There are few valid and reliable tools for measuring some outcomes that Head Start grantees and researchers agree are important, including children's social and emotional development, family well-being and engagement, and outcomes for dual-language learners and children with disabilities. This makes it difficult for grantees to know what goals to set for these outcomes or track data on their progress toward achieving their goals.
- » **Reporting requirements that do not support continuous improvement:** Current Head Start reporting requirements are oriented toward compliance — making sure that programs provide the information the federal government requires or carry out required activities — than facilitating the collection of information in a format that teachers, administrators, and program managers can use to support continuous improvement. Existing systems report data upstream but don't share that data with grantees in ways that add value to their work, nor do the systems aggregate data to identify trends among grantees. Head Start grantees that also receive state pre-k or child care funds, or other state, local, or philanthropic funds, are often subject to multiple, duplicative, and uncoordinated reporting requirements that impose additional compliance burdens on program staff.
- » **Data systems that do not help grantees understand and use their data:** Multiple commercial providers offer technology systems that Head Start grantees use to track and analyze data on attendance, enrollment, quality, child assessments, and other key program activities and outcomes. But these tools were primarily designed to support reporting and compliance rather than to help grantees understand and use their data. Although grantees can, in theory, use data from these tools to inform continuous improvement, doing so requires effort, sophistication, and significant staff time. Cost is also a barrier for some grantees.
- » **Staff capacity that is inadequate for analyzing and using data:** Most Head Start programs have staff assigned to collect and report data, but often in low-level data-entry roles and without the necessary skills to support teachers, center directors, and program leadership in analyzing and using data to inform practice. Head Start's prescribed staffing structure makes it difficult to reallocate resources or redefine existing roles in order to enhance capacity to analyze and use data. And Head Start administrators and center directors — who should ultimately be responsible for using data to inform continuous improvement — often lack training in the use of data and have numerous demands on their time, making it difficult to focus on improving quality and outcomes.

- » **Lack of support for building capacity:** Head Start invests \$200 million annually in training and technical assistance (see box 2), some of which is intended to support grantees' capacity for data-informed continuous improvement. For example, the National Center on Program Management and Fiscal Operations,¹¹ one of seven national centers that receive Head Start training and technical assistance funds, has created useful tools and resources to help grantees use and learn from their data. But changing practice and culture requires more direct, hands-on support to build grantee capacity. Half of training and technical assistance funds go directly to local grantees, who may use these funds to support staff time or training in order to analyze and use data. Relatively few do so, however, because current policies create strong incentives for grantees to use funds in other ways. Twenty-five percent of training and technical assistance funds go to regional offices, the multistate structure that administers Head Start, to help build grantees' capacity. But regional-office staff members often lack expertise in continuous improvement.

- » **A culture that does not support continuous improvement:** Continuous improvement is not just a set of specific practices, but a culture and mind-set that programs must cultivate at all levels. As long as grantees are primarily accountable for demonstrating compliance with numerous Head Start rules and regulations, however, it will be difficult for grantees to shift from a culture of compliance to one of data-informed continuous improvement. Other state and federal early childhood programs often reflect a similar compliance-based approach. Since many Head Start grantees also receive funds from these programs, changing the culture of compliance in Head Start may also require changes to, and improved coordination with, these other programs. Moreover, by micromanaging so many aspects of Head Start practice, current federal rules, regulations, and monitoring processes can limit grantees' ability to reallocate resources or make changes in response to what they learn from data.

Box 2: Head Start Research and Training and Technical Assistance

Under the Improving Head Start for School Readiness Act of 2007 (Head Start Act), the Secretary of Health and Human Services allocates annual funding to support continuous improvement in Head Start programs through two distinct types of activities: research and technical assistance.

Research

The Head Start Act allows the Secretary to provide up to \$20 million to fund research, demonstrations, and evaluation activities. This funding is designed to support ongoing assessment of program quality, contribute to the knowledge base of program practices that help and hinder child development, disseminate use of findings from research, and fill gaps in the existing knowledge base.¹² The Secretary can directly carry out these activities or contract with public or private entities to do so.¹³ In FY2014, Head Start spent the full \$20 million on research, evaluation, and demonstration activities.¹⁴

Training and Technical Assistance

The Head Start Act also provides funding to support training and technical assistance (T/TA) for grantees. The Secretary may reserve between 2.5 and 3 percent of all Head Start dollars for T/TA;¹⁵ in FY2014, the Secretary committed \$203 million,¹⁶ or about 2.1 percent of Head Start funding.

T/TA funding is split among actors at three levels. At least 50 percent of funding goes to local grantees. Grantees can use the funding for a range of activities, such as correcting deficiencies identified during monitoring, but most grantees spend T/TA funds to help staff pay the costs of earning required degrees. The other 50 percent of funding is split between regional offices and national centers. Regional offices use these funds to engage state-based T/TA specialists, who provide on-site support to grantees. National Training and Technical Assistance centers provide grantees and regional offices with resources and training that often serve as general information for the field rather than customized support. There are eight national centers that support grantees on topics such as health, professional development, and program management and fiscal operations.¹⁷

The best Head Start programs find ways around these obstacles — but making data-informed changes to achieve better outcomes should be the norm, not the exception. Meeting the proposed Head Start Program Performance Standards, which would require programs to “implement a process for using data to identify program strengths and needs, develop and implement plans that address program needs, and continually evaluate progress towards achieving program performance goals,”¹⁸ will require addressing these barriers. While individual Head Start grantees bear the primary responsibility for establishing data collection and continuous improvement systems, federal policymakers must support these efforts by reviewing and revising reporting requirements, changing requirements that limit grantees’ staffing flexibility, and allocating existing training and technical assistance funds to capacity building that helps programs meet new requirements.

The Performance Standards set expectations for practice at the individual grantee level. But federal policies should also seek to foster the development of Networked Learning Communities that bring together groups of Head Start grantees and researchers to share and analyze data to inform continuous improvement. By pooling the work and cost of data analysis across multiple grantees, this approach can be more cost-effective than each grantee trying to build data-analysis capacity alone. Collaborating with researchers enables grantees to access a higher level of expertise and more sophisticated data analysis and evaluation than they could obtain on their own. Moreover, looking at data from across multiple grantees can help them learn much more than merely looking at their own data. Using a set of common data points and indicators, Networked Learning Communities allow participants to see how their population and performance compare with those of other network participants, to identify areas of relative strength and weakness, to identify and learn from other programs that are producing particularly strong results in specific domains, and to work together with their peers and research partners to develop solutions to challenges identified in the data. By connecting grantees to researchers and grounding researchers’ work and questions in the needs and experiences of practitioners, this approach can help accelerate the pace of learning and improvement for both participating grantees and the field as a whole.

Some grantees, such as the members of the Minnesota Head Start Association (see box 3), have already created Networked Learning Communities on their own accord. Federal policies could encourage the development of these networks by allocating a share of training and technical assistance funds to support the creation and operation of Networked Learning Communities. Philanthropic funds can also support the creation of Networked Learning Communities in particular communities or regions, or on specific topics, and provide ongoing support for their work.

Box 3: Minnesota Head Start Association's Networked Learning Community

The nonprofit Minnesota Head Start Association (MHSA) works on behalf of Minnesota's 33 federal- and state-funded Head Start and Early Head Start grantees. The MHSA offers a compelling example of how Networked Learning Communities can foster continuous improvement in Head Start programs. In 2002, MHSA members formed four learning networks, called Learning Communities for Change, that give grantees the opportunity to drive continuous improvement initiatives in a collaborative environment. There is one learning community for each of these four topics: oral health, early childhood mental health, child and family outcomes, and school readiness goals. Grantees in each learning community establish shared goals; engage experts on the topic; define, measure, and communicate around common indicators of progress; and modify their work to adapt to new data or information from the field.

Participants in the Child and Family Outcomes Learning Community work together to improve the quality and comparability of the child and family assessment data they collect and to use that data to improve programming. The MHSA purchased group licenses for two child assessment tools, which allowed participating grantees to generate comparable data across programs. This work ultimately spurred the development of a voluntary statewide child-outcomes database, which the MHSA operates.

In 2010, MHSA created the School Readiness Goals Learning Community for grantees interested in aggregating and sharing data from the statewide database. Since then, nearly two-thirds of the Minnesota Head Start grantees have submitted data for sharing and analysis. The MHSA engaged an independent researcher to help analyze the data across participating programs, enabling the Learning Community to learn from trends, patterns, and variations in outcomes among programs.

SECTION 2: IMPROVING PROGRAM OVERSIGHT

At the **federal oversight level**, the Office of Head Start (OHS) needs a valid accountability and performance measurement system that allows federal officials to detect trends and patterns in performance; identify, learn from, and disseminate lessons from trends and patterns in performance; differentiate oversight, support, and interventions on the basis of grantee performance; and support continuous improvement across Head Start as a whole.

Head Start grantees are subject to multiple forms of accountability. Through the Aligned Monitoring System, OHS monitors grantees' compliance with the 1,400 requirements in the Head Start Program Performance Standards. The Designation Renewal System (DRS), mandated by the Improving Head Start for School Readiness Act of 2007 (Head Start Act), also creates a form of accountability by requiring grantees that hit one of seven "triggers" to compete to renew their grants. The criteria that trigger competition are:

- » deficiencies in compliance with Head Start Program Performance Standards, as identified through the Aligned Monitoring System;
- » license revocations by state or local licensing agencies;
- » suspensions by HHS;
- » disbarment by any state or federal agencies;
- » significant financial risk;
- » failure to establish or track and analyze data on goals for improving children's school readiness; and
- » scores on the Classroom Assessment Scoring System (CLASS) — a tool that OHS adopted to observe the quality of teaching in Head Start classrooms — that fall below a certain threshold or within the lowest 10 percent of grantees on any of the three CLASS domains.

The problem isn't a lack of accountability in Head Start, but rather what grantees are held accountable for. Head Start monitoring is largely compliance-based, using detailed checklists and requiring grantees to provide extensive paper documentation of meetings, plans, and other compliance requirements. Most of the seven factors that trigger designation renewal are related to grantees' compliance with federal, state, local, or financial requirements. Only the last requirement on the trigger list, CLASS scores, emphasizes program quality beyond mere compliance, and it focuses on a very narrow domain of Head Start performance. Further, the DRS does not differentiate between deficiencies that indicate serious shortcomings in programs' health and safety practices, finances, or quality and deficiencies that reflect minor compliance problems. As a result, designation renewal has actually intensified the culture of compliance in Head Start.

Designation renewal is grounded in sound principles — identify grantees that consistently fail to achieve desired results; take action to correct underperformance; and, where necessary, create opportunities for higher performers to replace the underperforming grantees. Data and reports from the field also suggest that designation renewal has led to the replacement of historically underperforming grantees — roughly 5 percent of all Head Start grantees have lost their grants in the past three years — and has stimulated improvement in underperformers. But the system’s reliance on imperfect measures has also resulted in identifying some high-performing grantees for competition, while failing to identify others that are low-performing.

Further, the DRS — and compliance-based monitoring more generally — identifies only those programs that fall short of standards, not exemplary performers. The Centers of Excellence program, authorized in the Head Start Act, was designed to identify, recognize, and reward high-performing Head Start grantees. It identified only 10 grantees a year, however — out of more than 1,400 — and has not been funded in recent years.

To really support continuous improvement, Head Start needs a fairer and more results-based and robust way of measuring grantee quality. This is a necessary precondition for both identifying and learning from programs that are producing exemplary results, and for achieving the intended goals of the DRS. While OHS should be responsible for putting in place and administering such a performance measurement system, it cannot do this work on its own. Identifying, developing, and implementing a robust set of performance measures for Head Start will require the engagement of a broader range of actors, including researchers, the philanthropic sector, state and local policymakers, and — crucially — Head Start grantees. Ultimately, the Head Start statute itself should incorporate references to the existence, purposes, and expected use of performance measures. But Congress should allow OHS to pilot and refine these measures with grantees before requiring them in a performance management system, in order to ensure that measures are valid, reliable, and useful to the field. Otherwise congressional action will be premature and potentially damaging. This paper seeks to outline an iterative process for developing and adopting more robust performance measures for Head Start programs, making information about grantee performance more transparent, and ultimately using this information to differentiate grantee performance in ways that support continuous improvement at both the individual grantee level and across Head Start nationally.

CURRENT CHALLENGES IN MEASURING HEAD START PERFORMANCE

Head Start stakeholders agree that federal oversight of Head Start grantees needs to shift from monitoring compliance to focusing more on results.¹⁹ The challenge is knowing what to measure and how to measure it. Head Start is a complex program that is delivered in a variety of contexts across the country and seeks to address multiple child and family outcomes. Therefore, it needs to track multiple measures of program performance, including measures of child and family outcomes and other indicators of program quality. Using multiple measures provides a more comprehensive picture of program performance and avoids placing too much weight on any single measure. A meaningful system of performance measures for Head Start programs should include:

- » **Child outcomes:** Head Start’s primary purpose is to improve child outcomes. Current monitoring systems ask grantees to demonstrate a variety of things — but not how they’re improving children’s outcomes. A meaningful system for differentiating program performance should incorporate multiple measures of children’s outcomes across the domains of child development included in the Head Start early learning standards, using tools that are valid and reliable for this purpose and for the population of students that Head Start programs serve — including dual-language learners and children with developmental delays and disabilities.

- » **Family outcomes:** Head Start is a two-generation program that seeks to drive and sustain improvements in child outcomes by engaging parents to support their children’s learning and development. A meaningful system of differentiating program performance should include measures of family outcomes or changes in parent behavior that are associated with improved outcomes for children.
- » **Leading indicators:** Leading indicators are intermediate outcomes that have a strong predictive relationship with ultimate outcomes of interest or reflect short-term outputs that lead to longer-term outcomes. Child attendance, for example, is both a strong predictor of children’s learning outcomes and an outcome of family engagement efforts.
- » **Safety, health, and compliance data:** Even in a more outcomes-oriented system, there will still be some minimum input and compliance requirements — such as those related to health and safety or compliance with civil rights laws — that all programs must meet. These minimum standards should continue to be included in any performance measurement system but should indicate the floor, not the ceiling, of program performance. Further, system designers should also consider where outcomes — such as injuries or reported safety incidents — could potentially replace or supplement more compliance-oriented measures.
- » **Program quality data:** Moneyball principles argue that, where possible, performance measurement and accountability systems should focus on results rather than inputs. Public policies, however, must also acknowledge that many factors outside of programs themselves can influence child and family outcomes in a given time period. Therefore, outcome measures should be complemented by measures of program quality and practices.
- » **Grantee health and capacity data:** Producing strong outcomes requires effective organizations that use resources well, support and develop their staff, and use data to inform continuous improvement. While programs should have broad flexibility in how they choose to carry out these activities, a well-designed system of performance management should include measures of the organizational health and financial well-being needed to support program delivery.

Knowing how to measure program performance in some of these domains is challenging, however.

For example, Head Start grantees are required to establish goals for children’s school readiness and to collect child-level assessment data on their progress toward those goals.²⁰ But the child assessments that most grantees use were primarily designed to inform teachers’ instruction and individualize learning for each child — not to provide information about the performance of Head Start programs.

Researchers have developed a number of promising, evidence-based tools to measure children’s learning in a variety of domains — including language, literacy, and science²¹ — and have validated or are currently validating these assessments. But these tools have not yet been widely implemented in Head Start or early childhood settings, or validated for large samples of children reflecting the Head Start population. Moreover, many of these tools focus on children’s development in a single domain, meaning that a comprehensive picture of children’s development outcomes would require using multiple assessments.

Similarly, researchers have developed a variety of promising new tools to measure quality in early childhood programs. Traditionally, early childhood programs have measured quality in terms of inputs, such as teacher credentials, square footage, and adult-child ratios. In contrast, the newer tools provide a systematic way of observing children's actual experiences in early childhood programs and are more predictive of children's learning than are traditional input measures.²² The Classroom Assessment Scoring System, used to measure the quality of adult-child interactions in Head Start classrooms, is one example. But, since the 2007 Head Start reauthorization, researchers have developed and validated, or are currently validating, a variety of other tools that look at children's experiences, instructional quality, and other dimensions of program practice correlated with children's learning.²³

The emergence of new, evidence-based measures of both child learning and program quality creates the potential to develop much more informative and robust measures of Head Start performance. But many of these measures are very new, and a great deal of work remains to determine whether they should be implemented at scale for performance measurement purposes. Even after new assessment tools are vetted for validity and reliability, they must still be piloted across a variety of Head Start and other early childhood programs before policymakers can understand the practical issues involved in implementing the tools at scale and the usefulness of the information they provide, and establish baseline data on current norms and variations on these measures across the Head Start population.

Additional research is also needed in order to develop valid and reliable measurement tools in domains where they are lacking. For example, while we know that family engagement and stability are crucial to children's early development and future school success, there is little evidence about which measurable parent and family outcomes are most important for Head Start programs to measure (see box 4). Similarly, there are few valid and reliable tools measuring children's social-emotional development.

Even where essential indicators seem easy to measure, it is not always clear how to best collect or present data to provide useful information about program performance. Child attendance offers a case in point. Measuring attendance seems like a no-brainer — either a child is there or she isn't — and Head Start programs already track data on attendance. But the way attendance data are collected and shared directly affects how they can be used to address key questions: Which children are absent the most? Is there a pattern to their absenteeism? How do interventions or changes in practice affect children's attendance? Similar questions arise about many potential indicators of safety, health, and financial quality. If a performance measurement system is to include information on these indicators, OHS will need to establish common standards for how programs collect and report data, in ways that maximize the data's value to programs and other users.

These challenges are real — but they cannot become an excuse for maintaining the status quo. Rather, what's needed is a strong commitment to moving toward more results-based measures of Head Start performance, coupled with an iterative and thoughtful process for developing and refining those performance measures and evaluation methods over time.

Box 4: Developing Measures of Parent, Family, and Community Engagement Study

Parent engagement and leadership is a core principle of Head Start, but efforts to measure program results have historically focused more on measuring child development and learning and less on family growth and engagement. Programs lack appropriate tools for gathering quantitative data about family engagement and must rely on measures of outputs, such as parent attendance at events, rather than capture meaningful changes in family outcomes over time.

To address that issue, in spring 2015, the National Head Start Association partnered with NORC — a research institute at the University of Chicago — and the Region V Head Start Association to launch a new study of Head Start parent, family, and community engagement. *Assessing Parent, Family, and Community Engagement through Parent Report* will:

- » inform the national understanding of parent engagement; and
- » create a tool to help local programs collect better information about parents' experiences in Head Start.

During the first phase of the project, in fall 2015, the study gathered stories from parents of children who participated in a random sample of Head Start, Early Head Start, Migrant and Seasonal Head Start, and American Indian and Alaska Native Head Start programs. The team at NORC will use computational text analysis to identify the themes parents expressed as being the most valuable Head Start contributions to their children and families. The analysis will also explore how these themes align with the Head Start Parent, Family, and Community Engagement Framework.²⁴ Ultimately, the study will produce a measurement tool that programs across the country can easily use to measure parent, family, and community engagement based on parent stories, focusing on the very dimensions Head Start families have identified as most important. The tool is intended to enable programs to measure individual families' changes in engagement over time, support measurement of program changes from year to year, and enable comparisons among Head Start grantees. It will provide programs with useful information to both enhance engagement with families and support ongoing program evaluation activities.

— Emmalie Dropkin, National Head Start Association

MOVING TOWARD STRONGER MEASURES OF HEAD START PERFORMANCE

The recent history of education is full of examples in which policymakers decided to hold programs or individuals accountable for outcomes, mandated accountability, and then create or adopted measurement tools to comply with accountability mandates. The National Reporting System adopted by the George W. Bush administration offers one such example, as do the seven designation renewal criteria adopted following the 2007 Head Start reauthorization. In K-12 education, many states developed systems of teacher evaluation in response to similar mandates. As these examples indicate, this approach can lead to measurement tools that don't actually measure the most important outcomes, causing unintended consequences and public backlash. A more sensible approach would be to identify the outcomes or components of program performance that are most important, develop tools to measure those outcomes or domains of program performance, pilot those tools in a small number of sites, refine tools based on pilot experience before rolling them out at scale, and give the field time to become familiar with the tools before attaching them to consequences.

How would this look in Head Start?

First, federal policymakers should define the key domains of child and family outcomes, safety, health, compliance, program quality, and grantee health and capacity that are crucial to measure in order to understand and improve the performance of Head Start grantees. This is a crucial first step toward adopting meaningful performance measures. Because decisions about which key domains to measure are not simply technical but also reflect judgments about the priorities and purposes that Head Start programs should emphasize, the process for defining these domains should engage a variety of federal policymakers, researchers, practitioners, and other stakeholders.

Once the key domains of program performance have been defined, OHS can identify the data for each of those domains that it currently collects — and stop collecting unrelated data. It is likely, however, that there will be key domains of performance in which consistent, valid, and reliable information is not currently available or collected across grantees. Where this is the case, federal officials must initiate a process, working with researchers, practitioners, and the broader field, to iteratively develop, refine, and adopt new measures of grantee performance.

Box 5: The Office of Head Start's Efforts to Create a More Performance-Oriented Monitoring Process

The Office of Head Start is already taking steps to make its monitoring process more performance-oriented and use it to better differentiate program performance. The Head Start Key Indicators, adopted in 2014, allow grantees that meet certain criteria to be reviewed against a streamlined set of 27 compliance measures, rather than complete a full monitoring review. Grantees undergoing the streamlined review also receive Classroom Assessment Scoring System observations and a review of environmental health and safety.²⁵ Grantees that fail the streamlined review must complete a comprehensive monitoring review, and those that pass a streamlined review in one review cycle must complete a comprehensive review in their next cycle. The Key Indicators are a positive — but only partial — step toward reducing the monitoring and compliance burden on programs with a track record of meeting standards, and it is too early to tell what effect they will have in practice.

Monitoring protocols developed for the FY2016 monitoring cycle include both compliance measures, which assess compliance with federal Head Start Program Performance Standards in key performance areas, and pilot quality measures, which are a preliminary step toward outlining a continuum of program quality to support ongoing quality improvement. The FY2016 Head Start Monitoring Protocol also calls on reviewers to meet with program directors and early childhood development coordinators to review the data that grantees collect on children's progress toward school readiness goals, as well as grantees' analysis of this data and any changes made in response to data. This approach stops short of holding programs accountable for the child learning outcomes they produce. But it does incorporate into the monitoring process greater attention to child outcomes, while also emphasizing programs' capacity to collect data on children's progress toward school readiness goals and use this data to inform continuous improvement. OHS should continue to build on this approach as an interim step until more robust program-wide measures of child and family outcomes can be adopted.

In some cases, relevant data on key domains of program performance may exist but not be collected consistently across grantees. Adopting a broader range of common performance measures across Head Start grantees will require developing standardized ways of collecting, reporting, and analyzing data on these performance domains. Rather than a top-down approach, however, OHS should work with the field to pilot common ways to collect and analyze data, building on work that the National Head Start Association and leading Head Start grantees have already done, in order to select feasible approaches that generate the most useful information for both grantees and OHS.

Where valid and reliable measurement tools do not exist, federal research agencies must work with researchers, the philanthropic sector, and the private sector to develop new tools to measure key outcomes or domains of program quality. A variety of federal agencies — including the Office of Planning, Research and Evaluation, within the U.S. Department of Health and Human Services; the Institute of Education Sciences, within the U.S. Department of Education; the National Institute of Child Health and Human Development within the National Institutes of Health; and the National Science Foundation — fund research on young children’s learning and development. Many existing and emerging tools used to measure quality and learning in early childhood programs grew out of this federally funded research. Other federal entities, such as the White House Social and Behavioral Sciences Team,²⁶ may also have valuable insights to contribute to this work. As a result, the federal government is well-positioned to fund both the validation and the refinement of existing tools for use in Head Start programs, as well as fund the development of new tools in domains where they are lacking.

This work will require collaboration, however. The federal agencies that fund relevant research will need to work together to develop a common research agenda for funding work that addresses gaps in the existing set of tools and knowledge. This would allow these agencies to share information about currently funded projects, pool resources across agencies where appropriate, avoid duplication of effort, and set common expectations for new measurement tools. For example, new measurement tools developed with federal funds should be designed to be valid and reliable for the diverse population of children that Head Start programs serve — including dual-language learners and children with disabilities — regardless of the agency that funds the research. Wherever possible, federal funds should also support the development of assessment tools that measure children’s progress, not just offer a snapshot of skills at one point in time.

Federal research efforts must also be coordinated with philanthropic and industry efforts. Philanthropic funders can typically make spending decisions more quickly and nimbly than federal agencies can, and can take greater risks with their funding. As a result, some stages of the research and development process, such as basic research to inform the creation of tools, or large-scale validation studies of promising models, may be well-suited to federal investment, while others, such as rapid iteration and refinement of new tools, may be better supported by philanthropic or private industry funds. To catalyze private efforts, federal research or philanthropic funds could also support prize-awarding “challenges,” along with traditional contracts or grants, as a strategy for accelerating the identification of outcome-measuring tools. Developing, piloting, and refining of new measurement tools also requires close collaboration among researchers, funders, and practitioners in the field. Finally, making new tools available at scale will likely require partnerships between researchers and commercial publishers or technology companies.

The vast majority of this work must be driven by actors outside the federal Office of Head Start, but OHS can play a valuable role in helping broker collaboration among various stakeholders and groups. For example, OHS could work with the National Head Start Association and other networks of grantees to facilitate collaboration between researchers and Head Start grantees in piloting new measurement tools.

This collaborative, iterative process would allow OHS to adopt or refine the performance measures it collects over time as research and the field generate new tools or new lessons about existing ones. New tools or indicators should initially be piloted in a subset of grantees to help OHS and researchers understand how these tools actually work in practice, collect baseline data on the distribution of grantee performance, and identify potential implementation challenges or considerations before adopting tools for program-wide use. As new tools and indicators are rolled out system-wide, they should initially be on a “no stakes” basis. Programs should be required to report information on these measures to OHS, and OHS should report information on grantees’ performance, both individually and in the aggregate, back to grantees and to the broader public. But there should be no consequence attached to new measures in their first two years of system-wide use so that grantees can become accustomed to new tools or indicators before they are used to inform decisions. As better tools for measuring child and family outcomes are adopted, OHS may be able to reduce its use of compliance and input-based indicators to measure program performance, while maintaining a key set of safety, health, and compliance measures to ensure that all Head Start grantees meet minimum standards.

MAKING GRANTEE PERFORMANCE INFORMATION MORE TRANSPARENT

As OHS adopts measures of program performance in key domains, it should publicly report the information on these measures in a transparent and accessible format that supports comparison across grantees on key indicators of program performance. OHS collects a great deal of data from grantees but does not share that data back with grantees, researchers, or the public in a way that informs continuous improvement. For example, OHS posts individual grantees’ CLASS scores and monitoring reports to its website, but this information is presented in a static way that does not allow users to easily access and compare information across multiple grantees, or to connect information from grantees’ CLASS and monitoring reports with other data included in their Program Information Report. Transparent interactive data reports would enable grantees, researchers, policymakers, and other stakeholders to access data on common program performance indicators, track trends in grantee or program-wide performance over time, and compare and analyze data across subsets of grantees serving similar populations. Interactive reports should enable users to sort data by child, family, and community demographic characteristics; grantee size; and other relevant grantee characteristics in order to facilitate constructive comparisons across grantees serving similar populations. Reports should also be designed to allow users to track correlations between different performance measures and trends in program-performance data over time.

OHS should publicly report on the grantee performance data it already collects, including information from Head Start Monitoring reports; CLASS scores; aggregate rates of child attendance; staff qualifications; and percentage of children with health insurance. While these data provide only a partial picture of program performance, they can still provide useful information to grantees, researchers, and other stakeholders. Data should be made available at the grantee, delegate agency, and program-wide level, with appropriate safeguards to prevent disclosure of individually identifiable child-level data. As new measures are developed and adopted for program-wide use, they should be added to transparent, interactive, and public data reports.

As noted above, all Head Start grantees are required to set goals for improving children’s school readiness and to collect and analyze child-level assessment data,²⁷ but programs are free to select the assessments they use to meet these requirements. In the absence of common, valid, and reliable measures of child learning outcomes, transparent, interactive reports should include information from the assessments that programs use to measure children’s learning and

development, with appropriate information from assessment publishers that places program-reported data in context.

To the extent that grantees collect data on additional measures that reflect their unique context, philosophy, community needs, or population served, they should also be permitted to submit this information for inclusion in transparent, interactive reports. Programs serving dual-language learners, for example, might submit data on children’s development in their home language.

DIFFERENTIATING GRANTEE PERFORMANCE

An effective system of performance measurement should not only collect information on grantee performance but also differentiate multiple levels of performance across various domains. One of the weaknesses of compliance-based systems is that they measure performance in strictly binary terms — either a grantee is meeting a standard or it isn’t. To be sure, there are some minimum standards, such as compliance with civil rights laws, or certain health and safety requirements, that all programs must meet, and where further differentiation is unnecessary or inappropriate. But, across many areas of grantee quality or outcomes, such as parent engagement or teacher-child interactions, the distribution of grantee performance is far more continuous. Understanding where grantees fall along a continuum of performance is often more informative than simply knowing that they met a minimum standard.

Differentiating performance across multiple indicators allows grantees to understand where they are doing well and where they need to improve, helping them focus their continuous improvement efforts. Ideally, the performance measurement system should allow OHS and grantees not only to see where their performance falls within the continuum of all Head Start grantees but also to compare their current and trend performance with that of “peer” grantees serving demographically similar children and families.

OHS should use performance data to identify those grantees that appear to produce meaningfully better results than their peers on specific outcomes, for specific populations, over a sustained period of time — as well as those demonstrating notable improvements in performance — in order to identify the practices that contributed to improved outcomes and disseminate information about those practices to other programs (see section 3). OHS should also share data with researchers to investigate relationships between grantees’ child outcomes and other indicators of program performance, grantee health, and compliance and determine which of these factors are most closely associated with better or improving child outcomes.

The primary goal of any performance measurement system should be to encourage and support grantees to improve their own performance. But another key function of performance measurement is to enable oversight agencies to identify and take action to improve grantees that fail to meet acceptable standards of performance — whether in a specific domain or overall. These grantees should receive increased support and interventions, targeted to areas where they need to improve. Those that do not improve should be required to compete to renew their grants or, in the most severe cases, defunded. A robust performance measurement system would improve on the current DRS by providing a more performance-focused and accurate way to identify low-performing grantees than the current seven designation-renewal criteria do. Grantees could be identified as low-performing based on either unacceptably low levels of performance on specific performance indicators or a pattern of low performance across multiple indicators over time.

But OHS shouldn’t just use performance measurement to identify low-performing grantees. It should also develop criteria to identify and publicly recognize high-performing grantees, work

with researchers and these grantees to identify the practices that contribute to their success, and disseminate those practices to the broader field. Grantees with a history of high performance should also be subject to less-frequent monitoring and encouraged to pursue opportunities to further enhance their impact on children and families. For example, high-performing grantees that want to adopt innovative approaches to improving child and family outcomes should be given flexibility to do so, provided they partner with independent researchers to evaluate the impact of changes in practice. Others could be encouraged to expand by applying for Head Start grants that come up for competition in nearby or similar communities. Through its system of Key Indicators, OHS has already taken steps to differentiate monitoring, allowing programs with a history of clean reviews and audits to be reviewed against a streamlined set of compliance measures (see box 5).²⁸ A more robust system of performance measures would allow OHS to further differentiate monitoring and streamline requirements for programs that produce strong results.

In a well-designed performance management system, the vast majority of Head Start grantees will be neither low-performing nor high-performing, but will likely be identified as adequately performing. Performance measurement can support continuous improvement for these grantees by creating incentives to improve, identifying areas of strength and weakness, and helping them learn from and adopt practices from higher performers in the areas where they need to improve. OHS could also use performance measurement data to identify trends and common needs across grantees in order to inform training, technical assistance, and other grantee supports. For example, if data suggest that many grantees perform poorly on measures of children's math development, the agency may need to increase resources for training in early math and monitor subsequent data on children's math learning outcomes to assess the effectiveness of these investments.

SECTION 3: USING RESEARCH AND EVALUATION TO IMPROVE PROGRAM OUTCOMES

At the **research and evaluation** level, federal policymakers need to support research that builds the knowledge base of what works and use this research to inform changes in program design and policies.

Over the past 50 years, research has dramatically expanded our understanding of young children’s learning and development, and of “what works” in educating them. But there is much that we still do not know. We know a lot about practices that improve learning in Head Start and other early childhood programs, but much less about how to best support teachers and programs in consistently implementing those practices at scale. To support programs in continuously improving outcomes, the federal government also needs to fund research to identify effective practices, address gaps in the existing knowledge base, and develop new tools to support implementation of effective practices at scale.

Evaluations and other forms of research do what performance measurement alone cannot. Performance measurement tracks progress toward intended program outcomes, but does not compare outcomes with outcomes from alternative programs or the status quo. Evaluations determine whether specific programs, interventions, or policies produce outcomes superior to alternative policy choices, or to no policy at all. Other forms of research increase understanding of problems or opportunities that policies seek to address. These forms of research can inform priority setting or decisions about policy trade-offs, as well as the development of new models or approaches that seek to improve outcomes or lower costs.

The federal government currently spends \$20 million annually on research, demonstration, and evaluation activities related to Head Start. While this sounds like a substantial amount of money, it represented less than 0.25 percent of total Head Start spending in FY2014.²⁹ To build the knowledge base for the field, federal policymakers should increase funding for Head Start research, demonstration, and evaluation to 1 percent of total appropriations. But how funds are used is as important as the amount. These additional funds should support research that is carried out in partnership with grantees to build knowledge that is responsive to grantee needs and is immediately applicable to improving quality and outcomes in Head Start settings. Federal research funds should support three primary activities:

1. **Identification of effective practices of high-performing grantees and the testing of their replicability:** As noted above, a robust performance measurement system should allow OHS to identify grantees that are producing meaningfully better results on specific indicators, for specific populations of children, over a sustained period of time. Federal research funds should support partnerships between these grantees and researchers to identify and understand the practices and program characteristics that contributed to the improved child and family outcomes and develop effective mechanisms to replicate these practices and characteristics in other Head Start grantee and early childhood settings.

- 2. Measured trials of new practices and approaches:** Dramatically improving results for Head Start children and families will require developing better, scalable approaches to teacher professional development, curriculum, family engagement, and a range of other key program practices and services. Innovative approaches are also needed to help Head Start grantees prevent and reduce chronic absenteeism, mitigate the impact of early trauma and toxic stress, and respond effectively to children’s challenging behaviors. A more robust system of performance measurement would allow grantees to demonstrate performance based on the results they produce, rather than on compliance with regulatory requirements, which often limit grantees’ ability to innovate. Once such a system is in place, Congress could give OHS the authority to allow groups of grantees with strong performance track records to experiment with new approaches and to work with researchers to evaluate those new approaches based on their results. Crucially, these innovative approaches should be chosen and designed by or with practitioners, in response to their own experience and data, and should be sustainable using existing Head Start funds. Using rapid iteration and a series of incremental improvements over time, researchers and practitioners can work together to improve the effectiveness of existing program practices and tools; develop new, more effective practices and tools; use real-time data to evaluate their impact; and refine approaches in response to data.
- 3. Development of valid and reliable tools to measure key outcomes:** As noted above, the lack of valid and reliable measures in some key domains of program quality and child and family outcomes is a challenge not just for Head Start but also for the early childhood field as a whole. Federal research funds should prioritize the development of valid and reliable tools to measure:
- » children’s social-emotional development;
 - » adaptations or additional tools for measuring learning outcomes for dual-language learners;
 - » adaptations or alternative tools for measuring the learning progress of children with disabilities;
 - » key indicators of family outcomes and appropriate tools for measuring those outcomes; and
 - » more efficient and cost-effective tools for assessing children’s learning in each domain of learning.

Beyond funding research, federal policymakers must use the results of research to improve Head Start policies and program design. The recently proposed changes to the Head Start Program Performance Standards offer an example of policy and program design choices based in research. Key proposed changes — such as a shift from half-day to full-day preschool programs, an increased focus on job-embedded professional development for Head Start teachers, and a requirement that programs use evidence-based curricula to build parents’ confidence and skills in supporting their children’s learning — are deeply informed by research findings on what works in early childhood programs.

But the evidence on what works in early childhood programs will continue to grow and evolve after the revisions to the Performance Standards are finalized. The last extensive revision of the Head Start Program Performance Standards happened 17 years ago — and the knowledge base on Head Start has evolved significantly since then. Given the new ways in which cognitive and neuroscience research are expanding our knowledge of children’s learning and development, we cannot afford to wait another 17 years to use new information to improve Head Start programs. Going forward, the Secretary of Health and Human Services should regularly update the Performance Standards as the knowledge base continues to grow and evolve.

Finally, to support Congress and the administration in a future reauthorization of Head Start, the Advisory Committee on Head Start Research and Evaluation should conduct a periodic review of the relevant evidence on the full range of education and comprehensive services that Head Start programs provide.

SECTION 4: BROADENING SUPPORT

Local grantees bear the primary responsibility for implementing systems of data-informed continuous improvement, and federal policies create conditions and incentives that support or hinder these efforts. But other actors, including the philanthropic sector, researchers, states, and the private sector, also have a valuable role to play in supporting and accelerating this work.

The philanthropic sector can play a crucial role both in developing grantee-level capacity and in accelerating the development of new performance measurement tools. As noted above, many grantees will need to build additional capacity to collect, analyze, and act on data in the ways envisioned by both this paper and the draft Head Start Program Performance Standards. Philanthropic funders can make one-time investments in the professional development, technology, and data infrastructure needed to enhance program capacity. Philanthropy can also fund third-party organizations, such as nonprofit organizations or university-based researchers, to work with Head Start grantees to analyze their data. And philanthropic funds can help grantees make changes in their practices in response to data. Because Head Start staffing and other programmatic requirements limit grantees' flexibility to reallocate their funds, marginal philanthropic dollars can have a disproportionate impact in enabling programs to fund improvements or activities that they identify as crucial to improving outcomes but that are not mandated by Head Start rules. Philanthropic funders can also support the creation of Networked Learning Communities by awarding grants to consortia of grantees working with researchers to share and analyze data and inform program improvement. Using philanthropic funds to demonstrate the potential of Networked Learning Communities to improve child and family outcomes in Head Start could strengthen the case for Congress to authorize dedicated training and technical assistance funds for these networks in future Head Start reauthorizations.

Philanthropic funds can also help accelerate the development of new measurement tools for use in Head Start programs. Because philanthropic funders are typically able to act more quickly and take more risks than government funders can, they are better suited to funding rapid-cycle testing and iteration of new assessment tools in order to develop and pilot the tools to the point where they are ready for large-scale validation with public or industry funds.

Both philanthropic and public investments in research and data analysis should seek to cultivate deep, long-term partnerships between researchers and local Head Start grantees or groups of grantees. In exchange for conducting analyses and exploring questions of interest to grantees, researchers in these partnerships gain access to extensive program data sets. These types of partnerships can ensure that research is grounded in the experience and needs of Head Start grantees, rather than those of researchers, and results in tools and resources that are useful to practitioners and can be feasibly implemented in real program contexts.

Developing scalable solutions will also require engagement from the private sector, including both commercial publishers and technology developers. Making new assessment tools or observational quality measures widely available to Head Start grantees requires delivery capacity and platforms that most researchers and early childhood providers lack. Similarly, emerging technologies make it possible to design systems that collect and share data in real time in ways that are useful to the people submitting the data, as well as to the people monitoring programs and the people doing research about them. Tablets and online game-playing testing capacities can also help reduce the challenges involved in collecting data on children's learning or classroom quality. But realizing the potential of these technologies will require collaboration among Head Start grantees, researchers, and technology developers.

Data systems offer yet another potential avenue for public-private collaboration to enhance grantee capacity. Most grantees use one of several commercially available data systems, such as ChildPlus; the Program Resources and Outcomes Management Information System; and the Child Outcome Planning & Administration system, but these systems were primarily designed to support mandated reporting and compliance, not to help grantees make sense of their data. Groups of grantees, however, could exercise their collective market power to demand changes or additional functionalities in these systems to better support grantees in analyzing and using data for program improvement. The National Head Start Association, OHS, or the philanthropic sector could help facilitate this by convening data-system vendors, grantees, experts, and federal officials to identify necessary changes or new functionalities. Alternatively, philanthropic funders could invest in the development of new, open-source data systems that would be available to grantees at a lower cost than commercial options are and be designed to support continuous improvement as well as compliance and reporting.

OHS could also work with the philanthropic sector, technology innovators, and independent researchers to develop and implement the transparent and interactive reports on grantee performance described in section 2. Because of federal contracting regulations, developing these interactive online reporting capacities under the auspices of OHS would likely be a time-consuming and inflexible process. Rather than building these tools itself, OHS could investigate opportunities to enter into data-sharing agreements with an independent research organization, which could then access philanthropic funds and in-kind technology support to develop interactive, transparent reporting tools in a much more rapid, iterative, and nimble way than OHS could do through standard technology contracts. The resulting interactive reports could also be more easily and quickly revised as OHS incorporates new data sources into its performance measurement system.

Efforts to improve the use of data, evidence, and evaluation in Head Start must also engage state-level actors. Over the past decade, at least 37 states adopted quality rating and improvement systems (QRIS) that seek to differentiate quality in early childhood programs and support quality improvement, and more than 25 states have developed kindergarten entry assessments (KEA) to collect data on children's school readiness at the beginning of kindergarten.³⁰ The design of QRIS and KEA varies widely across states, as does the extent to which Head Start grantees are included in QRIS. But some states — particularly those that received Race to the Top-Early Learning Challenge grants or Enhanced Assessment Grants for KEA from the U.S. Department of Education — are generating useful information about and new models for quality measurement and child assessment that can also inform performance measurement in Head Start. Many Head Start grantees also participate in state-funded pre-k or child care subsidy programs that require them to report data to the state agencies that administer these programs. As more states build out early childhood longitudinal data systems that collect data on children's participation in publicly funded early childhood programs, Head Start grantees are increasingly participating in these systems. OHS should work with state-level officials to identify opportunities to use common measurement tools, data-reporting standards, and time lines that reduce the burden of duplicative data-reporting requirements on Head Start grantees. Ideally, OHS could develop data-sharing agreements with state agencies that allow Head Start grantees to submit a single report on data indicators of interest to both state agencies and OHS, which these agencies could share with one another. OHS should also pursue similar data-sharing agreements with other federal agencies and programs, such as the Child and Adult Care Food Program.

While the philanthropic sector, researchers, states, and the private sector all have valuable roles to play, federal policies and actions can influence other actors' desire and ability to engage with Head Start programs. For example, the lack of transparent, publicly available data on Head Start grantees' quality may dissuade some funders from engaging with Head Start programs. Making

information about grantee performance more accessible might help to catalyze philanthropic investment. More directly, the Secretary of Health and Human Services should work with an intermediary organization or collaborative of early childhood funders to encourage philanthropic investment in building programs' capacity for continuous improvement.

SECTION 5: FEDERAL POLICY RECOMMENDATIONS

This paper outlines a broad vision for how Head Start can and should use data, evidence, and evaluation to create a system of continuous improvement that leads to better results for children and families. Realizing this vision will require significant changes at the local program or grantee level. Other organizations, including the Urban Institute and the Head Start National Center on Program Management and Fiscal Operations, have already provided extensive recommendations and guidance about the changes that must occur at the local program level.³¹

But federal policies play an important role in facilitating, incentivizing, or deterring local grantees in their ability to make these changes, and they also impact the context in which other key actors — such as states, the philanthropic sector, the private sector, and researchers — engage with Head Start grantees. Ultimately, federal policies can best support local grantees and other actors in continuously improving quality by measuring Head Start grantees' performance against a clear set of common indicators that accurately reflect program quality and correlate with improved child and family outcomes; by transparently reporting on grantees' performance on these indicators; and by using this information to identify and learn from high-performing programs, identify and intervene in low-performing programs, and help other programs improve.

Fully realizing this vision will require additional research and cycles of experimentation to address outstanding technical and measurement challenges. This work must be done in collaboration with researchers, the philanthropic sector, the private sector, states, and — most importantly — Head Start grantees themselves. There are steps, however, that Congress and the administration can take now, both to advance this vision in the near term and to support the research needed to fully realize this vision in the future:

- 1. Congress and the Secretary of Health and Human Services should make data-informed, continuous improvement a key priority in any legislative or regulatory policy action on Head Start.** While implementing systems of data collection, use, and continuous improvement is primarily a grantee responsibility, federal policies must support grantees in implementing systems of ongoing continuous improvement:
 - » The U.S. Department of Health and Human Services should ensure that the final rule on the revised Head Start Program Performance Standards maintains requirements from the proposed rule related to continuous improvement,³² and OHS should develop effective ways to monitor these standards.
 - » Congress should ensure that any future reauthorization of Head Start emphasizes the importance of ongoing continuous improvement at the program level.
- 2. The Office of Head Start and the philanthropic sector should invest in building grantee capacity to use data to improve performance.** Many grantees need additional support and tools to enhance their capacity to use data to improve child and family outcomes. OHS can help build this capacity through training and technical assistance funding and should also partner with the philanthropic sector to encourage philanthropic investment in building program capacity.
 - » Congress should authorize the Secretary of Health and Human Services to dedicate a percentage of training and technical assistance funds to competitive grants for

networks of local grantees, working in partnership with researchers, to create or operate regional or affinity-based Networked Learning Communities, like the Minnesota Head Start Association's Learning Communities for Change.

- » OHS should require grantees to target their own technical assistance funds to build capacity for data analysis and continuous improvement, as well as to address areas of weaker performance identified through their continuous improvement processes.
- » OHS should require future regional-office contracts to stipulate that contractors must employ staff with expertise in data-informed continuous improvement.
- » OHS should use performance measurement data to evaluate the effectiveness of training and technical assistance offerings and inform allocation of federal and regional training and technical assistance funds.
- » OHS or philanthropic leaders should work with providers and vendors to encourage development or improvement in existing data systems to better support daily operations, continuous improvement, and research, as well as group-purchasing options that would lower the cost to grantees. For example, OHS could support focus groups of vendors, providers, and experts to identify changes in data-system functionalities that would support grantees in analyzing and using data for program improvement, or facilitate the creation of purchasing pools to lower the cost of data systems for grantees.
- » OHS should work with an intermediary organization or collaborative of early childhood funders to encourage philanthropic investment in building grantees' capacity for continuous improvement.

3. Federal policymakers should initiate an iterative process to develop robust, common performance indicators for Head Start and should engage researchers, the philanthropic sector, and Head Start grantees as partners in this process. Supporting continuous improvement across all Head Start grantees requires better ways of measuring grantee performance. While OHS must adopt and administer a performance measurement system for Head Start, identifying, developing, and implementing a robust set of performance measures for Head Start will require the engagement of a broader range of actors, including researchers, the philanthropic sector, the private sector, state and local policymakers, and Head Start grantees.

- » Federal policymakers, working with researchers and practitioners, should define the key domains of child and family outcomes, safety, health, compliance, program quality, and grantee health and capacity that are crucial to measure.
- » OHS should identify the data it collects on each of those domains — and stop collecting less-useful data.
- » OHS should partner with groups of local grantees to pilot common ways to collect and analyze data on domains of performance where valid and reliable measures exist but are not consistently reported.
- » OHS should develop formal processes for vetting, piloting, adopting, and revising new valid and reliable measures of program performance in key domains as researchers develop them. These processes should ensure that any new measurement tools are piloted in a subset of grantees before being mandated for all Head Start grantees, and that no consequences are attached to new measurement tools during their first two years in program-wide implementation.
- » The Secretary of Education should work with the Secretary of Health and Human Services and OHS to ensure that lessons from the Enhanced Assessment Grants awarded to Texas, North Carolina, and Maryland to enhance kindergarten entry

assessments, as well as the results of quality rating and improvement system validation studies conducted with Early Learning Challenge Grant funds, inform the development and adoption of performance measures for Head Start programs.

4. Federal research agencies should work with researchers and the philanthropic sector to support the development of solid, trusted metrics of Head Start child outcomes, family outcomes, and program quality.

The lack of solid, trusted assessments or metrics of child and family outcomes is the major barrier to making Head Start less compliance oriented and more performance oriented. To address this problem:

- » The Secretary of Health and Human Services and the Secretary of Education, and the directors of the National Institutes of Health and the National Science Foundation, working with the Advisory Committee on Head Start Research and Evaluation, should develop a common research agenda for the development of valid and reliable assessments of young children's learning, family outcomes, and early childhood program quality to address gaps in the field.
- » The Secretary of Health and Human Services and the Secretary of Education should direct a share of federal Head Start and education research funds to support a multiagency effort to develop, identify, and validate measures of child and family outcomes and program quality in key domains over the next five years.
- » OHS should work with local grantees and researchers to facilitate piloting and validation of new measurement tools in Head Start programs.
- » The Secretary of Education should work with state pre-k programs, school districts, and charter schools to facilitate piloting and validation of new measurement tools in other early childhood settings.

5. The Office of Head Start should provide transparent, interactive public reporting on grantee performance.

OHS must not only collect information from grantees but also share data and analysis back with them in a way that helps them to improve. Grantee- and program-level data should also be made available to researchers and the broader public, with appropriate safeguards to prevent disclosure of individually identifiable child-level data:

- » OHS should develop transparent, interactive public reports on grantee performance indicators that allows grantees, researchers, and the public to access and analyze information on grantee performance across multiple domains and indicators, as well as overall.

6. Once the Office of Head Start has developed a sufficiently robust system to measure grantee performance, it should use this system to differentiate grantee performance.

Differentiating grantee performance is essential to supporting continuous improvement in Head Start programs but requires accurate and fair ways of measuring grantee performance. The criteria used to identify grantees for designation renewal have significant shortcomings, but the preceding recommendations can help to establish fairer and more comprehensive ways of assessing Head Start grantees' performance.

- » Once such systems are in place, the Secretary of Health and Human Services should use information about differentiated program performance to:
 - identify high-performing grantees and support research to learn from and scale their effective practices;
 - support improvement in adequately performing grantees by providing transparent information about their performance, enabling them to identify

strengths and weaknesses in their performance, and helping them identify and learn from high performers; and

- intervene in low-performing programs and, when necessary, defund or require them to compete for grants.
- » For the time being, continue to implement the Designation Renewal System but revise the system in the future to incorporate more comprehensive and accurate measures of program performance as they are adopted.

7. The Office of Head Start should continue, learn from, and build on efforts to make program monitoring more performance focused and less compliance oriented. The Head Start Key Indicators represent a positive step toward reducing the monitoring and compliance burden on grantees and shifting Head Start monitoring from emphasizing compliance with numerous regulations to focusing on key domains of quality and performance. But this is only a partial step. To build on this work:

- » OHS should continue the Key Indicators model.
- » OHS should study and learn from its experience, and that of its grantees, with the Key Indicators. Based on this experience, it should seek to identify ways to further reduce monitoring burdens on programs.
- » OHS should continue to seek opportunities to further reduce compliance monitoring requirements and increase its use and analysis of regularly reported data to inform program monitoring.

8. The Secretary of Health and Human Services should implement a robust research agenda for Head Start. Improving Head Start outcomes requires investments in research, development, and evaluation to build the knowledge base on what works in Head Start and directly address the needs of Head Start grantees:

- » Congress should increase the cap on Head Start research, demonstration, and evaluation spending from \$20 million to 1 percent of total appropriations.
- » The Secretary of Health and Human Services should shift the focus of Head Start research funding to emphasize research that directly addresses the needs of practitioners in the field and is carried out in collaboration with them, including research to:
 - identify, understand, and replicate the effective practices of high-performing grantees;
 - conduct measured trials of new practices and approaches to delivering Head Start services; and
 - develop, pilot, and validate new tools to measure key domains of child and family outcomes and other indicators of Head Start grantee performance.

9. Congress should authorize the Secretary of Health and Human Services to grant additional flexibility to allow cohorts of programs working with researchers to pilot new approaches to serving children and families. Innovation and experimentation are crucial to dramatically improving child outcomes at scale in Head Start and other early childhood programs but require increased flexibility and ongoing, rigorous evaluation. The proposed revisions to the Head Start Program Performance Standards allow flexibility for grantees to implement “local variation” approaches to program options, curriculum, and professional development, provided the grantees can show that these approaches meet the needs of their communities (for program option variations) and that they will work with researchers (for curriculum and professional development variations) to evaluate the effectiveness of the local approach. Congress

should build on this model by authorizing the Secretary of Health and Human Services to create enhanced local variations or “experimental sites” that would allow grantees or networks of grantees with a proven track record of results to radically experiment with one or more aspects of program design in return for rigorous evaluation of program outcomes. Under this model:

- Programs would have increased flexibility to pilot new models of service delivery, professional development, curriculum, assessment, family services, or other Head Start services, and in exchange they would commit to tracking data on agreed-upon outcomes and performance indicators.
- Pilots would be accompanied by rigorous, independent evaluation.
- Local variations would be granted for a period that is limited but sufficient to support rigorous evaluation of program outcomes and could be extended beyond the initial time frame only if the program demonstrates strong results.

CONCLUSION

Assessment and continuous learning are central tenets of quality early childhood practice. Effective early childhood teachers are constantly observing children’s progress and adjusting learning environments and customizing learning for children based on what they observe. Achieving excellent outcomes for Head Start children requires programs to implement similar norms and cultures of ongoing assessment, analysis, and learning at the program and federal oversight level. Inculcating continuous learning within Head Start at the federal policy level can support ongoing learning at the program level and ultimately lead to better results for children and families.

END NOTES

1. See, for example, David Burkum and Valerie E. Lee, *Inequality at the Starting Gate* (Washington, DC: Economic Policy Institute, 2002), http://www.epi.org/publication/books_starting_gate/; Tamara Halle et al., *Disparities in Early Learning and Development: Lessons from the Early Childhood Longitudinal Study - Birth Cohort (ECLS-B)*, Child Trends, Executive Summary, June 2009, <http://www.childtrends.org/wp-content/uploads/2013/05/2009-52DisparitiesELExecSumm.pdf>.
2. “Head Start Program Facts: Fiscal Year 2014,” Office of Head Start, Department of Health and Human Services, accessed August 19, 2015, <http://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/docs/hs-program-fact-sheet-2014.pdf>.
3. *Head Start Impact Study*, Final Report (Washington, DC: US Department of Health and Human Services, Administration for Children and Families, 2010).
4. Nikki Aikens et al., *Getting Ready for Kindergarten: Children’s Progress During Head Start*, FACES 2009 Report, OPRE Report 2013-21a (Washington, DC: Office of Planning, Research and Evaluation; Administration for Children and Families; and US Department of Health and Human Services, 2013), http://www.acf.hhs.gov/sites/default/files/opre/faces_2009_child_outcomes_brief_final.pdf.
5. David Deming, “Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start,” *American Economic Journal: Applied Economics* 1, no. 3 (2009): 111–34, http://www.people.fas.harvard.edu/~deming/papers/Deming_HeadStart.pdf; Jens Ludwig and Douglas L. Miller, “Does Head Start Improve Children’s Life Chances?,” *The Quarterly Journal of Economics* 122, no. 1 (2007).
6. Hirokazu Yoshikawa et al., *Investing in Our Future: The Evidence Base on Preschool Education* (New York: Foundation for Child Development, 2013), <http://fcd-us.org/sites/default/files/Evidence%20Base%20on%20Preschool%20Education%20FINAL.pdf>.
7. “Head Start Program Facts,” Office of Head Start.
8. “Moneyball Principles,” Moneyball for Government, accessed August 19, 2015, <http://moneyballforgov.com/moneyball-principles/>.
9. Jim Minervino, *Lessons from the Classroom: Implementing High-Quality Pre-K that Makes a Difference for Young Children* (Washington, DC: Bill & Melinda Gates Foundation, 2014), https://docs.gatesfoundation.org/documents/Lessons%20from%20Research%20and%20the%20Classroom_September%202014.pdf.
10. Teresa Derrick-Mills, *Understanding Data Use for Continuous Quality Improvement in Head Start* (Washington, DC: Urban Institute and Office of Planning, Research and Evaluation, 2015), <http://www.urban.org/research/publication/understanding-data-use-continuous-quality-improvement-head-start>.

11. "Program Management and Fiscal Operations," Office of Head Start Early Childhood Learning and Knowledge Center, accessed August 19, 2015, <http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/operations>.
12. Improving Head Start for School Readiness Act of 2007, Pub. L. No. 110-0134, Sec. 640 (2007).
13. Improving Head Start for School Readiness Act of 2007, Pub. L. No. 110-0134, Sec. 649 (2007).
14. "Head Start Program Fact Sheet FY 2014," Office of Head Start, Administration for Children and Families, accessed November 23, 2015, <http://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/docs/hs-program-fact-sheet-2014.pdf>.
15. Improving Head Start for School Readiness Act of 2007, Pub. L. No. 110-0134, Sec. 640 (2007).
16. "Head Start Program Facts," Office of Head Start.
17. "T/TA Resources," Office of Head Start, accessed August 19, 2015, <https://eclkc.ohs.acf.hhs.gov/hslc/tta-system>.
18. Notice of Proposed Rulemaking on Head Start Performance Standards, 1302.102(c).
19. See, for example, testimony provided by Head Start agency director Tim Nolan, former Office of Head Start director Yvette Sanchez Fuentes, and co-author Sara Mead at an October 2015 U.S. House of Representatives Education and the Workforce Committee hearing on Head Start reauthorization: <http://edworkforce.house.gov/calendar/eventsingle.aspx?EventID=399395>.
20. 45 C.F.R. § 1307.3 (b)(2) (2007).
21. See, for example the early language screener developed by researchers at Temple University's Temple Infant and Child Laboratory, the University of Delaware, and Smith College (<http://www.cla.temple.edu/infantlab/>); the Lens on Science computer-adaptive science assessment for preschool children (<http://ies.ed.gov/ncer/projects/grant.asp?ProgID=7&grantid=805>); and the Minnesota Executive Function Scale (<http://www.cehd.umn.edu/icd/research/carlson/projects.html>).
22. Andrew Mashburn et al., "Measures of Classroom Quality in Prekindergarten and Children's Development of Academic, Language, and Social Skills," *Child Development* 79, no. 3 (2008): 732-49.
23. See, for example, the FirstSchool Snapshot, developed by researchers at the University of North Carolina's Frank Porter Graham Child Development Institute (<http://fpg.unc.edu/node/5691>) and the Five Essentials for Early Education, developed by the University of Chicago Consortium on Chicago School Research and the Ounce of Prevention Fund (see <https://ccsr.uchicago.edu/sites/default/files/uploads/Five%20Essential%20Supports-Early%20Ed%20Two-Page%20FINAL%20PDF.pdf>).

24. "Welcome to the Parent, Family, and Community Engagement Interactive Framework!," Office of Head Start, accessed August 19, 2015, <http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/family/framework>.
25. "Overview: Head Start Key Indicator-C Tool," Office of Head Start, RIKInstitute, accessed October 16, 2015, <http://rikinstitute.wikispaces.com/file/view/HSKI-C2.pdf>.
26. For more information, see "Executive Order: Using Behavioral Science Insights to Better Serve the American People," White House Office of the Press Secretary, *Politico*, last modified September 15, 2015, <http://static.politico.com/56/62/0efb293b4d8c899d8c554cc63af3/executive-order-on-behavioral-science.eo.rel.pdf>.
27. 45 CFR § 1307.3 (b)(2) (2007).
28. "Overview: Head Start Key Indicator-C Tool," Office of Head Start.
29. "Head Start Program Facts," Office of Head Start.
30. "Quality Rating and Improvement Systems," BUILD Initiative, accessed November 25, 2015, <http://www.buildinitiative.org/TheIssues/EarlyLearning/QualityQRIS.aspx>; "Kindergarten Entry Assessments," BUILD Initiative, accessed November 25, 2015, <http://www.buildinitiative.org/TheIssues/EarlyLearning/StandardsAssessment/KEA.aspx>.
31. See Teresa Derrick-Mills, *Understanding Data Use for Continuous Quality Improvement in Head Start*, Urban Institute and Office of Planning, Research and Evaluation, last modified February 2015, http://www.urban.org/research/publication/understanding-data-use-continuous-quality-improvement-head-start/view/full_report; Teresa Derrick-Mills et al., *A Resource Guide for Head Start Programs: Moving Beyond a Culture of Compliance to a Culture of Continuous Improvement*, OPRE Report 2015-02 (Washington, DC: Office of Planning, Research and Evaluation; Administration for Children and Families; and US Department of Health and Human Services, 2015), http://www.urban.org/research/publication/resource-guide-head-start-programs/view/full_report; "Data in Head Start and Early Head Start," National Center for Program Management and Fiscal Operations, Administration for Children and Families, and US Department of Health and Human Services, accessed November 25, 2015, <http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/operations/data>.
32. Section 1302, Subpart J; see Bellwether Education Partners, *Analysis of Notice of Proposed Rulemaking in Head Start Performance Standards*, pp. 34-36, last modified August 2015, http://bellwethereducation.org/sites/default/files/NPRM_Analysis.pdf.

ABOUT THE INVEST IN WHAT WORKS POLICY SERIES

This report is part of Results for America's Invest in What Works Policy Series, which provides ideas and supporting research to policymakers to drive public funds toward evidence-based, results-driven solutions. Results for America is committed to improving outcomes for young people, their families, and communities by shifting public resources toward programs and practices that use evidence and data to improve quality and get better results.



Invest in
What Works
Policy
Series