Data Utilization Practices in Head Start Programs

Ashley LiBetti

Table of Contents

Click on each title below to jump directly to the corresponding section.

Background

This brief is part of Bellwether Education Partners' *Leading by Exemplar* project, a multi-year study researching the practices of five exemplary Head Start programs. This brief provides in-depth information about exemplary programs' data utilization practices.

The *Leading by Exemplar* project has three goals: to identify Head Start programs that are producing powerful results for children, elevate them as proof points of what is possible for the field, and learn from their practices to inform policy and efforts to improve early learning outcomes. To identify potential exemplars, Bellwether Education Partners drew on publicly available quantitative data and recommendations from experts and stakeholders in the field. A program was eligible for this study if it had demonstrable evidence, via an external evaluation or internal analysis of longitudinal data, of positive impacts on children's learning that were either substantially larger than those of typical Head Start or other early childhood programs or sustained beyond kindergarten entry. We believe there are many more Head Start programs that meet this criterion, but our analysis focuses on these five programs.

Among our findings was that data utilization is a key driver in a program's performance. In this brief, we synthesize common data utilization practices across exemplars and provide lessons for other early childhood programs — Head Start and otherwise — as well as implications for the field.

Additional information about the *Leading by Exemplar* project, including methodology, lessons for the field, and other analysis, is available here. Case studies of each exemplary program as well as a brief synthesizing the programs' instructional models are also available.

Head Start Examplars

Acelero Learning Camden/Philadelphia

CAP Tulsa

Educare Miami-Dade

Fairfax County Public Schools

Utah Community Action

Introduction

Knowing what effective early childhood programs do in practice is different from knowing how to ensure consistent, high-quality implementation of those practices.

There is no "one best formula" that ensures early childhood programs will effectively serve children and families. Research has identified program practices and characteristics that are associated with improved quality in early childhood programs, as well as some that are less effective. But knowing what effective early childhood programs do in practice is different from knowing how to ensure consistent, high-quality implementation of those practices. Moreover, there are many areas of early childhood program practice where research provides less clear guidance for providers.

Further, because Head Start and other early childhood programs vary widely in their contexts, resources, workforce, and populations of children they serve, each grantee must customize both the services it offers and its approach to delivery, based on local resources and needs and often refined over time based on experience.

All this means that, in order for Head Start and other early childhood programs to realize their full potential, they must be able to learn from their experiences and identify strategies that will be effective within their unique context. A key vehicle for doing is the use of data to inform ongoing continuous improvement.

Data-informed continuous improvement is a cyclical process through which programs use data to understand their performance and make adjustments to improve. In this process a program uses data to identify areas for improvement, designs interventions to address those issues, implements the intervention, uses data to assess the results of interventions, and begins the cycle again.

Specifically, programs execute this cycle through three buckets of activities:

- **Collecting data**, often multiple types (e.g., input, process, and outcome data) and levels (e.g., child-, classroom-, and program-level data)

- **Analyzing data** to identify potential correlations and relationships between data points, as well as potential opportunities for improvement and areas of strength

- **Acting on data**, which includes both communicating the data to stakeholders and identifying appropriate interventions, if any, based on what information the program wants to act on

In recent years, federal policies and leaders in the field have encouraged and provided support for Head Start programs to adopt these continuous improvement cycles. The Office of Planning, Research, and Evaluation within the Administration for Children and Families has funded research and resources to build the field's understanding of data-informed continuous improvement.[1] Resources and tools to help programs understand and implement data-informed continuous improvement practices have also been disseminated through Head Start's training and technical assistance network. The 2016 Head Start Program Performance Standards include a new section dedicated to developing goals and the systems that programs should use to monitor progress towards those goals (see sidebar). And the National Head Start Association (NHSA) regularly provides trainings and workshops to programs to support their continuous improvement efforts.

Head Start Program Performance Standards Requirements on Achieving Program Goals

Subpart J, Section 1302.102 of the 2016 Head Start Performance Standards outlines requirements to support programs in achieving their program goals. This section has four components, but the one most relevant here is *1302.102(c): Using data for continuous improvement*. Its provisions require programs to develop and implement a process for using data to identify program strengths and needs, develop plans to address those needs, and monitor progress towards goals and compliance with standards.

As part of that process, programs must aggregate and analyze data, including analyzing child-level assessment data three times a year, in the aggregate and disaggregated by subgroups (e.g., dual language learners and children with special needs). Using these data (as well as data on teaching practice, staffing and professional development, family needs assessments, and comprehensive services), programs must identify specific needs and come up with a plan to address them. If necessary, programs should use program improvement plans to improve program practice around professional development, change program scope and services, refine school readiness and other program goals, and adapt strategies to better address the needs of sub-groups.[2]

However, despite these efforts and a strong desire among many Head Start grantees to build their capacity to use data for ongoing improvement, many Head Start and other early childhood programs face challenges in implementing or maximizing the benefits of data-informed continuous improvement.

In 2016 Bellwether Education Partners, working in partnership with the National Head Start Association, Results for America, and the Volcker Alliance, engaged a wide range of Head Start stakeholders and experts in the field to assess the current state of continuous improvement in Head Start and outline a vision for using data, evidence, and evaluation to improve results for Head Start children and families. This work included several high-level recommendations for programs, such as increasing staff capacity to conduct internal analyses and fostering a culture of continuous improvement rather than compliance. It did not, however, offer detailed recommendations for how programs should implement these approaches.

Exemplar programs are highly intentional in their use of data to inform continuous improvement, and the processes they have put in place to do so are key parts of how these programs produce exemplary results.

Over the past three years, as part of the *Leading by Exemplar* project, Bellwether has conducted extensive research to identify and learn from the practices of exemplary Head Start programs. Exemplary Head Start programs, for the purposes of this project, are programs that are producing child learning outcomes that exceed those of typical early childhood programs and/or are sustained as children progress through the public education system.

We found that all these exemplar programs are highly intentional in their use of data to inform continuous improvement, and that the processes they have put in place to do so are key parts of how these programs produce exemplary results. This paper analyzes common practices of these programs and provides specific implementation examples to help other programs identify how they can adopt — and adjust to their context — some of these practices themselves. For more information on how the programs implement these approaches, please see each program's case study.

Common Practices of Exemplary Programs

These programs' experiences feature strategies and important conditions for effectively implementing data-informed continuous improvement that haven't been highlighted in previous research.

The exemplars profiled here meet the continuous improvement requirements of the Head Start Program Performance Standards and implement nearly all of the best practices recommended by early childhood analysts and experts in data-informed continuous improvement. At the same time, these programs' experiences feature strategies and important conditions for effectively implementing data-informed continuous improvement that haven't been highlighted in previous research. Through their internal trial and error, they have developed rigorous processes that other Head Start programs, analysts, and policymakers can learn from. Specifically, these programs' continuous improvement processes share several core components:

First, data **inform teacher practice**. Teachers use data and observations on an ongoing basis to monitor children's progress, set goals for individual children, and differentiate learning opportunities based on children's progress. These data are also used to identify and craft targeted interventions for children who need more intensive support. Programs also use data to **support staff development**. Coaches also rely on data to inform one-on-one sessions with teachers: Teachers and coaches set goals together; use observation, child performance, and other data to assess progress towards those goals; identify strategies to build on strengths and develop challenge areas; and then start the process over again with new goals. Programs roll up those individual-level data to identify shared needs among teachers, design professional development content to address them, and then track the progress closely to determine if additional supports are necessary. All supervisor-supervisee relationships in these programs follow similar coaching and development cycles. Integrating data into daily

practice and staff development builds staff capacity and confidence in using data and also encourages staff to view data as a means for improvement and development, creating a culture in which data are seen as a tool rather than a threat.

Programs also use data to **monitor their overall effectiveness**. They conduct regular "pulse checks," which pull together data from a variety of sources to tell the story of the overall health of the organization and pinpoint areas of concern. As discussed later in this paper, pulse checks do not automatically trigger action, but they do keep programs aware of where things are working well and where improvements may be needed. Similarly, programs use data to track their impact on the children and families they serve, creating a form of self-accountability. Program leadership shares the findings with all staff and other stakeholders, creating transparency around data and performance that contributes to a culture that embraces data as a tool for improvement.

Finally, programs use data-informed continuous improvement cycles to **assess the quality and impact of specific program practices** in both instruction and operations and to identify, pilot, and test innovations to improve those practices. To that end, programs may use child outcome data and teacher feedback to identify strengths and weaknesses in their existing curricula, pilot curricular modifications or new curricula in a sample of classrooms, and use this cycle to determine if the modifications are worth rolling out to all classrooms or should be dropped entirely. Similarly, the teams that manage a program's staffing regularly analyze staff performance and hiring process data to determine whether their hiring practices are bringing the right staff into teaching and other roles, and to revise and test new selection criteria to bring more effective teachers into the classroom.

These practices form a backbone of data-informed continuous improvement processes that are common across all the exemplars studied. The specific ways that exemplars implement these practices vary, highlighting multiple strategies that programs can use to meet similar goals. At the same time, common themes emerge across these programs' varied approaches.

These practices form the backbone of data-informed continuous improvement processes that are common across all the exemplars studied, but the specific ways that exemplars implement these practices vary.

Implications for the Field

This section explores the practices that exemplars use to implement rigorous and effective data-informed continuous improvement processes. Seven practices and key themes emerged as particularly important and common across all exemplar providers:

1 Intentionally select tools to measure child, classroom, and program outcomes.

2 Analyze outcome data in the context of input and process data.

3 Conduct frequent pulse checks…

4 …and use pulse check data to determine if improvement efforts are necessary.

5 Develop a program culture that views data as a tool for learning, not punishment.

6 Commit resources to building staff capacity in collecting, analyzing, and using data.

7 Assess program effectiveness via unbiased research partnerships.

① Intentionally select tools to measure child, classroom, and program outcomes.

Early childhood programs often cite the limitations of existing measurement tools as a barrier to or challenge in data-informed continuous improvement efforts. Although child development —and these programs' work — is complex and multi-faceted, many existing measurement tools focus on only one or a few developmental domains (e.g. language development) or areas of program practice (e.g. adult-child interactions). As a result, no existing measure gives programs all the information they need to monitor program effectiveness, making it difficult to identify areas of need, set goals, or develop strategies for improvement. Further, programs also struggle with a lack of valid and reliable measurement tools for several important child and family outcomes and program practices.

Exemplary programs have mitigated this risk by strategically selecting measurement tools that reflect their purposes and goals and employing multiple assessments from a variety of different sources. To assess classroom quality, for example, all programs in this sample use the Classroom Assessment Scoring System (CLASS), but supplement it with classroom environment checklists, developed either internally or by an outside company, that look at dimensions of teacher practice and classroom environment not reflected in the CLASS.

Exemplary programs strategically select measurement tools that reflect their purposes and goals, and employ multiple assessments from a variety of different sources.

Similarly, programs track child outcomes using a combination of criterion-referenced assessments (which assess mastery of or ability to meet predefined objectives) and norm-referenced assessments (which compare children's progress to the performance of the hypothetical "average" child). Programs may use a curriculum-companion assessment, such as Teaching Strategies' GOLD® (TS GOLD®), to inform teacher practice, and a separate tool — either from an external assessment company or developed internally — to track organization results over time. **Utah Community Action (UCA)**, for example, developed a new assessment tool called the Pre-K Skills Assessment after they learned that children's performance on curriculum-companion assessments did not predict or correlate to the measures of kindergarten readiness used by the public school districts where UCA Head Start children entered kindergarten. The Pre-K Skills Assessment measures foundational children's literacy and numeracy knowledge, allowing UCA to better understand and track children's progress in these areas. UCA administers the Pre-K Skills Assessment three times a year, taking a baseline at the beginning of the year and assessing students again in mid-winter and spring to monitor progress throughout the year.

The **Fairfax County Public Schools (FCPS)** early childhood program also developed their own assessment, the Early Childhood Rubric, but they rely on theirs to both inform classroom instruction and monitor program health. Teachers use the EC Rubric as a formative assessment to measure children's progress throughout the year and differentiate instruction. At the same time, data from the EC Rubric informs professional development, changes to the curriculum, and other program improvement efforts at a system-wide

level. FCPS developed and uses their own assessment because it's aligned to the Virginia state standards, Head Start standards, and their own curriculum, but they also want to ensure that children are making progress compared to national norms. FCPS regularly compares performance on their internally developed assessment with the Devereux Early Childhood Assessment and the Phonological Awareness Literacy Screening for Pre-K. These assessments are norm-referenced, meaning they can show how a child's performance compares to that of a hypothetical "average" child. This is crucial for gauging child performance and program effectiveness, not just against their own past performance but against other programs nationally, and also helps FCPS ensure that the Early Childhood Rubric is valid — that is, that the assessment data are providing the type of information that they want it to.

Programs may also use multiple assessments to track multiple types of outcomes, such as the Minnesota Executive Function Scale to track social emotional development and executive function, the Peabody Picture Vocabulary Test to track verbal ability, or the Woodcock Johnson Applied Problems and Quantitative Concepts assessment to track early math skills. Using multiple assessments allows a program to effectively conduct an internal validation process.

❷ Analyze outcome data in the context of input and process data.

Over the past 25 years, there has been an increased focus on the importance of measuring outcomes in education and early childhood programs. All the exemplar programs here focus intently on outcomes for the children and families they serve, but they also realize that outcomes are, at best, a lagging indicator. Looking at outcome data alone doesn't provide a clear guide for action to improve those outcomes. Rather, these programs demonstrate that consistently linking outcome data with data on inputs and processes is necessary to identify opportunities for improvement and deliver high-quality programming.

Exemplar providers look at a mix of data on a continuum, from inputs to processes to outcomes, and regularly analyze relationships between the different data points to inform their work.

Exemplar providers look at a mix of data on a continuum, from inputs to processes to outcomes, and regularly analyze relationships between the different data points to inform their work. **Utah Community Action**, for example, looked at the relationship between child attendance and performance on assessments and found that lower attendance was correlated with lower performance. In response, they developed visual tools to show parents the relationship between attendance and performance, specific to each classroom, which teachers posted outside their doors. Analysis conducted soon thereafter showed that attendance increased after teachers shared the visuals. Similarly, **Acelero Learning's** internal analysis showed a tight relationship between child outcomes and family characteristics. This analysis led Acelero Learning to invest heavily in families, including developing a family engagement curriculum that gives parents activities and tools to participate in their children's learning at home.

Programs have also used interconnected input, process, and outcome data to inform changes to their teacher recruitment and selection criteria. For example, FCPS's internal data analysis indicated that many new teachers were not prepared to work with children from high-need backgrounds. In response, program leaders revised their interview process so as to better assess a candidate's ability to work with children living in poverty or those who have experienced trauma, as well as a candidate's level of empathy for families experiencing those circumstances, and revised the performance task to assess how candidates would respond to children who exhibit disruptive behaviors. Similarly, **CAP Tulsa** analyses found that prospective teachers who did not have teaching experience were, on average, as effective as teachers who did. CAP Tulsa revised their hiring criteria to consider candidates with and without teaching experience equally. CAP Tulsa also recently found that a multiple-choice behavioral assessment, which was previously required during the teacher hiring process, was negatively correlated with success in teaching roles. The program has since dropped that assessment.

❸ Conduct frequent pulse checks…

As discussed previously, exemplary programs conduct frequent "pulse checks." In a pulse check, program leaders pull a specific set of data points that reflect program progress towards goals. Pulse checks serve two purposes: They allow leadership to monitor overall organizational health and progress, and they identify areas that may require attention or improvement.

Programs' pulse checks serve two purposes: They allow leadership to monitor overall organizational health and progress, and they identify areas that may require attention or improvement.

Each program conducts its pulse checks differently. All programs conduct both comprehensive pulse checks, in which they track crucial data points across all priority areas (financials, human capital, academic, family engagement, etc.) to get a full picture of program operations, and targeted pulse checks that drill down into more detailed data on specific priority areas. But when and why programs conduct these two pulse checks, and if and how they share out those data, varies.

CAP Tulsa, for example, conducts regular comprehensive pulse checks, which include data on internal process metrics (such as volunteer hours, recruitment inquiries, and website traffic), financial metrics (such as cost per child analyses and fund development), and employee learning and growth metrics (such as turnover rates and staff satisfaction). CAP Tulsa's leadership team reviews these data every other month, then reviews the program's performance over time at the end of every year. **Utah Community Action** follows a similar structure: The executive team conducts comprehensive pulse checks every month and does a formal analysis annually. Every spring, the program conducts self-assessment to determine the degree to which each program area has achieved its comprehensive services and school readiness goals. As part of this self-assessment, the leadership team collects quantitative and qualitative data from 18 different sources, including teacher and parent

interviews, TS GOLD® child performance data, and classroom observations using several different tools. The results of the self-assessment are shared with the entire leadership team as a way to summarize the program's progress and challenges over the previous year.

CAP Tulsa, Utah Community Action, and **Acelero Learning** also regularly perform targeted pulse checks on child and family data to review with academic staff. They pull data at the program level on indicators such as chronic absenteeism, child achievement in literacy and math, and family engagement in events, then disaggregate those data by site. Each program follows a different timeline, but eventually these data are shared with all staff.

④ … and use pulse check data to determine if improvement efforts are prudent.

As is to be expected, pulse checks often reveal aspects of program performance that have room for improvement. But pulse checks do not automatically trigger immediate action in these areas. Exemplary programs consider pulse check information in the context of what else is happening in the program before deciding whether or not to act on it. **CAP Tulsa**, for example, explicitly asks site leaders if the pulse check data align with what they are seeing day to day.

Because these programs are deliberate in their continuous improvement processes, they also do not automatically respond to every challenge with an improvement intervention. Instead, programs balance new data with their other values and priorities, and may decide that the best course of action is no action at all. Programs decide not to act on a challenge area for different reasons. It may be because they are in the process of piloting an intervention and want to keep as many variables constant as possible. In other cases, program leadership may want to limit the burden of a new intervention on staff — both to avoid negative effects on morale and to avoid jeopardizing the quality of implementation of new and existing programming. **Utah Community Action**, for example, learned from a survey that the pace of change and improvement was too fast for teachers and staff; in response, the program took steps to titrate their continuous improvement efforts.

In this way, exemplar programs employ intentional, strategic decision-making to determine if they should act on the pulse check data before they even begin to discuss *how*.

But this deliberate approach also shapes how programs act in response to identified challenges. Exemplar programs rarely enact sweeping program-wide changes in response to data. Rather, they more typically identify a range of potential responses or solutions and pilot or test them on a small scale before deciding whether to apply them more broadly. **Acelero Learning's** "pencil-pen-Sharpie" approach (discussed in their case study) offers a case in point. Acelero Learning deliberately fosters a culture in which any staff member can propose a new initiative or intervention based on program data and staff members'

Exemplary programs consider pulse check information in the context of what else is happening in the program before deciding whether or not to act on it.

own observations and experiences. These new initiatives are initially implemented or tested on a very small scale — what Acelero Learning refers to as the "pencil" stage — or held in the proposal stage until leadership deems it prudent to act. Initiatives that prove successful in the pencil stage may then advance to wider implementation in the "pen" stage and eventually be embedded into the core of the program's practices (the "Sharpie" stage). But many initiatives never move past the idea or pencil stage. This approach allows Acelero Learning to create space for innovation and staff agency for improvement, while carefully vetting new innovations and avoiding churn or swings in program practice that could be destabilizing for staff.

5 Develop a program culture that views data as a tool for learning, not punishment.

These exemplar programs intentionally cultivate a program culture that views data as a tool for understanding and supporting progress, rather than a source of punishment.

These exemplar programs intentionally cultivate a program culture that views data as a tool for understanding and supporting progress, rather than a source of punishment. This culture, and the degree to which it enables staff and leadership to learn from their performance, is a key driver of these programs' exemplary results.

As discussed above, these programs foster a positive culture around data by deliberately and carefully deciding if and how they will respond to new data. In doing so, they communicate to staff that data are not a driver to action but rather an invitation to curiosity and inquiry around what the data show and why this might be occurring.

Additionally, leadership teams build a positive data culture by modeling data as an improvement tool and by publicly holding the program — and themselves — accountable. As mentioned, **CAP Tulsa**, **Utah Community Action**, and **Acelero Learning** all facilitate regular meetings with all staff to share out updates and findings on the organization's health and performance based on pulse check data. Crucially, these meetings focus on program operation goals as well as instructional quality. In other words, the meetings include data and goals that reflect the performance of program leadership as well as line staff. Leadership does not attempt to hide areas where their performance falls short of goals; instead, within this full-staff setting they flag challenge areas, hypothesize about why those challenge areas exist, ask for additional context and feedback from staff, and propose next steps and solutions (which, as discussed, may be to take no action at all). In this way, program leadership signals their comfort with data as a reflection of performance and their willingness to be vulnerable enough to draw lessons from, rather than make excuses for, challenge areas.

Further, staff are constantly exposed to data as a tool to improve their own performance and better individualize instruction. All of these programs ground coaching and professional development conversations in individual and aggregated observation and child- and

classroom-level performance data. Over the course of the year, teachers see that coaches are not using the data as a "gotcha," but rather as information that provides a backdrop to the conversation. Teachers receive training to support their use of data in improving their instruction, which further emphasizes the provider's priorities for how the program uses data.

Finally, teachers and staff regularly see data being used not just to measure performance, but also to catalyze action to better support children and families. For example, Acelero Learning's *Shine Insight*, which houses all of the program's data, will alert teachers and staff if certain high-priority data need attention — for example, if a child requires an urgent health follow-up or if there is a family crisis that must be addressed. Those data do not reflect performance; rather, they spur interventions that children and families need in real time. Teachers and staff are exposed to data as a supportive, rather than punitive, tool that is crucial for meeting children's and families' needs.

⑥ Commit resources to building staff capacity in collecting, analyzing, and using data.

Rigorous and continuous quality improvement efforts require that program staff have the capacity to collect, analyze, and use the data. Exemplary programs build staff capacity by funding dedicated research and analysis staff, by providing trainings and professional development on analyzing and using data, and by granting staff access to program data.

Many of the steps that programs put in place to build a culture of curiosity rather than fear around data also play a crucial role in building staff capacity to use data to inform their own practice and improvement. As discussed, everyone in these programs — teachers, coaches, site leaders, and program leaders — is responsible for understanding and using data that is relevant to their daily work. Exemplary programs build the capacity of all staff to analyze and act on data through regularly scheduled trainings and professional development.

Exemplary programs build the capacity of all staff to analyze and act on data through regularly scheduled trainings and professional development.

Educare Miami-Dade, for instance, has multiple layers of coaching (the executive director coaches the center director, the center director coaches master teachers, and master teachers coach teachers), all of which are based in data, including observations of teacher practice, CLASS scores, attendance, and child outcome data. Data-focused professional learning communities offer another venue through which programs build teachers' and teacher leaders' capacity to use data. In **Acelero Learning's** Assessment Work Group, for example, teachers discuss child-level assessments with each other and their coaches, sharing, for example, how they would score a specific observation. Three times a year, the group dives into their assessment data, building skills through collaborative and supportive analysis of real-time data.

Additionally, **CAP Tulsa**, Acelero Learning, and **Utah Community Action** have invested in staff roles that exclusively focus on internal program analysis and building their colleagues'

capacity to access and use data. The **FCPS** early childhood program receives similar support from the district-level research arm, the Office of Research and Strategic Improvement. Through this partnership, FCPS has the capacity to monitor the effectiveness of the program overall. At the same time, early childhood program managers frequently analyze school-, classroom-, and child-level data to inform their own continuous improvement efforts and answer questions in response to specific program- or site-level needs.

Finally, each of the exemplary programs profiled here build staff capacity with data by regularly sharing both raw data and analyses with all staff. FCPS early childhood program managers run dozens of analyses intended to measure progress towards school readiness goals, compare performance over time and to other counties, and disaggregate performance by schools and subgroups; they share all of this information with teachers and staff. The program is also working to develop a data dashboard so principals and teachers can access a variety of aggregated and disaggregated data — such as attendance, behavior, Pre-K Phonological Awareness Literacy Screening, and Devereux Early Childhood Assessment data — at any time and look at these data over time to identify trends and see progress within their own schools. Similarly, as discussed, CAP Tulsa, Utah Community Action, and Acelero Learning regularly share raw and analyzed program-level data with all staff.

7 **Assess program effectiveness via unbiased research partnerships.**

Programs have developed long-standing partnerships with external, unbiased research organizations. These partnerships are another way that the programs hold themselves accountable and assess their own effectiveness and their impact on children.

Data-informed continuous quality improvement processes within programs are most effective when paired with external checks on program performance. To that end, these programs have developed long-standing partnerships with external, unbiased research organizations. These partnerships are another way that the programs hold themselves accountable and assess their own effectiveness and their impact on children.

Program partnerships operate in different ways and to different ends. **Educare Miami-Dade,** for example, has a recurring partnership with the University of Miami that serves two purposes: The research team from University of Miami both produces analyses that inform program practice in real time and conducts formal evaluations of the program's effectiveness in working with children. **CAP Tulsa**, **Acelero Learning**, and **Utah Community Action**, on the other hand, partner with research universities that serve exclusively as external validators. CAP Tulsa and Utah Community Action have recurring partnerships with one to two institutions that conduct analyses every year to understand the programs' effect on children. Acelero Learning partners with multiple research institutions to explore a variety of different questions aligned to research teams' strengths and areas of expertise. **FCPS** achieves a similar function by working with the district's in-house Office of Research and Strategic Improvement, which routinely conducts independent evaluations of various programmatic offices within the district.

Conclusion

To truly realize the potential of Head Start, programs must invest in and commit to using data for continuous improvement. These programs have developed a unique set of tools, processes, and systems to support this work, resulting in positive learning gains for children. The Head Start Program Performance Standards, training and technical assistance resources, and support from the field provide a degree of guidance, but in order to achieve the types of results that will truly improve children's lives, programs themselves must learn from each other and drive their own improvement.

Appendix

Interviewees

Maralyn Akiyama
Steve Barnett
Melissa Beard
Rebecca Berlin
Laura Bornfreund
Jennifer Brooks
Adia Brown
Amanda Bryans
Donna Bryant
Miriam Calderon
Jeffrey Capizzano
Lydia Carlis
Erin Carroll
Jenna Conway
Amy Cubbage
Marquita Davis
Libby Doggett
Steven Dow
Linda Espinosa
Danielle Ewen
John Fantuzzo
RB Fast
Ellen Frede
Yvettee Sanchez Fuentes
Cathy Garland
Jackie Govan
Sharon Huang
Stephanie Jones
Victoria Jones
Myra Jones-Taylor

Gayle Kelly
Joan Lombardi
Amy Madigan
David Mandell
Jana Martella
Kelly Maxwell
Jim Minervino
Rick Mockler
Barbara Montero
Pamela Morris
Jennifer Park
John Pruette
Craig Ramey
Colleen Rathgab
Monica Roers
Joel Ryan
Aaliyah Samuel
Tom Schultz
Kathy Stack
Lisa Stewart
Cynthia Stringfellow
Abby Thurman
Eric Vaughn
Albert Wat
Sarah Weber
Christina Weiland
Elizabeth Weingartner

Endnotes

1   See, for example, Derrick-Mills, Teresa, "Understanding Data Use for Continuous Quality Improvement in Head Start, Preliminary Findings," Office of Planning, Research, and Evaluation (OPRE), February 2015, https://www.urban.org/sites/default/files/publication/51211/2000216-understanding-data-use-for-continuous-quality-improvement-in-head-start.pdf.

2   U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start, Head Start Program Performance Standards, (September 2016), https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/hspps-appendix.pdf.

Acknowledgments

About the Author

Ashley LiBetti

Ashley LiBetti is an associate partner on the Policy and Evaluation team at Bellwether Education Partners. She can be reached at ashley.libetti@bellwethereducation.org.

About Bellwether Education Partners

Bellwether Education Partners is a national nonprofit focused on dramatically changing education and life outcomes for underserved children. We do this by helping education organizations accelerate their impact and by working to improve policy and practice.

Bellwether envisions a world in which race, ethnicity, and income no longer predict opportunities for students, and the American education system affords all individuals the ability to determine their own path and lead a productive and fulfilling life.