

## Building a Dynamic Learning and Continuously Improving Head Start System

Shelley H. Metzenbaum  
National Head Start Association CQI & Research Luncheon Keynote  
Anaheim, CA  
April 25, 2018

### My Mission Today

I am a woman on a mission today.

I have come to excite you, more than that, to enlist you – in the transformation of Head Start.

Specifically, I hope that, together, we can strengthen the capacity of the Head Start system to learn from experience and from measured trials ways to **improve on multiple dimensions**.

Let's me start by sharing my vision with you: it is my hope that we can create a **dynamic learning and improving system and culture** that achieves greater beneficial impacts than now accomplished – that improves outcomes for those served and yields higher returns on spending so more children are served even better on multiple dimensions. Further, I hope to convince you to adopt the actions needed to transition to **outcomes-focused management** and, beyond that, to manage for **reduced risk, higher ROI, and high-quality** parent or other caregiver and community **experience** at all Head Start centers.

Please note that I distinguish between **improving outcomes** and **improving experience**. Both are important and both should be managed. So should **risks** and **costs**.

Become a vanguard for change in Head Start. In doing so, pave new paths for other early learning programs.

Consider: if you don't, who will?

Consider further: what will happen if you don't? Will you still, ten years from now, be stuck choosing programs and practices for Head Start programs without knowing which are likely to work better for children of different learning inclinations and different life experiences in your classroom as you are today?

I want to share a vision with you that I began to discuss with a smaller group of Head Start folks in 2014. Happily, many of you and the National Head Start Association have picked up the gauntlet to bring this vision to life. NHSA produced the [Moneyball for Head Start](#) report and recommendations, together with three other organizations, one of which I headed at the time. Building on that, NHSA recently launched the [Data Design Initiative](#), including today's program, to bring the recommendations of the Moneyball report to life.

Today, I hope to enlist more of you in this change-making and outcomes-improving effort – to work together to find better approaches that enable Head Start children to thrive throughout their lives.

### What would a dynamic continuous learning and improvement system look like?

Such a system has 7 key characteristics:

1. **Providers, especially the front-line managers, are deeply involved in learning and improvement efforts.**
2. **Outcome data, and other key indicators, are regularly collected.**
3. **Constructive data analyses are done frequently.**
4. **Both data and analyses are broadly shared, with careful attention given to privacy needs but also to identifying key users and determining when, where, and how those users can and should use the analyses.**
5. **Data-informed discussions and decision-making occur regularly and frequently to find what works better, problems needing attention, likely causes of problems, and likely factors contributing to progress. That information then guides action.** These meetings, focused on outcomes, end with clearly established expectations about who needs to do what by when. The meetings may involve individual centers, centers plus delivery partners in their communities, across centers, the National Head Start Association, the U.S. government, researchers in academic and think tanks, foundations, consultants, and others – together or separately.
6. **Providers participate in measured field trials to discover ways to do better that current data do not reveal, trials informed by needs identified during data-informed discussions.**
7. **The public can easily find and understand Head Start’s and each center’s goals, why they were chosen, some key indicators, strategies, why the strategies were chosen, progress made, problems encountered, and planned next steps, short and long term.**

Let me elaborate a bit on these 7 characteristics.

**I. First, Head Start providers play a leading role, creating a continuous learning and improvement system and communities.**

It is time to integrate the collection and application of evidence about outcome changes more seamlessly into operations. By evidence, I mean data about outcomes and other key information, plus results of measured field trials and lab research.

One step toward that is creating stronger links across practitioners, researchers, and government program managers.

Another step is collecting data that is more useful and meaningful for those involved in the system – people working with the children, center managers, parents, others in the community who can contribute to children’s well-being, and funders.

That requires **returning data back to data suppliers and other key information users with value added through analyses**, analyses that help those engaged in program delivery understand what is likely to work better in different situations and alert them about risks that can be prevented. Think how much more attention people will pay to the quality of the data they submit to central offices if the data get returned in ways they find useful.

It also requires sharing field trial findings with folks in the field in a format they can easily find, access, understand, and apply. This should be a core and critical aspect of program operations, not an afterthought.

**To create a continuous learning and improvement system, we also need to change the questions we try to answer with analyses and evaluations.**

For years, Head Start studies have attempted to answer the question: “Does Head Start work?”

I would urge that we focus on a somewhat different question: “What works *better*?” Or, more precisely, “**What works better for whom in what situations?**”

If you frame the question as “Does Head Start work?,” you get one of three answers: yes, no, or maybe. How can you act on that information? If the answer is yes, so what? It does not help you figure out how to do better. If it is no, it sets up a tussle about whether or not to kill the program, but does not suggest how to address the problems or advance the opportunities the program was designed to address. And if it is maybe, then what do you do? In short, the answer to this question, when answered, is not actionable.

Changing the question from “what works” to “what works better?” sets up a whole series of answers, and triggers a set of more focused questions that start to identify better ways to advance the objectives.

That is not to suggest that it is not important to find what works. It is.

But that is just the first question, and it may not be a question that should be applied to a whole program but rather to a practice and collections of practices. As soon as an effective practice or collection of practices is found, we should start asking how to do better. Are there ways to use the same resources to accomplish more and ways to accomplish the same with fewer resources?

## **II. Better Along What Dimensions – Outcomes and Other Indicators**

Whether the question is what works or what works better, you also need to answer the question: “Better along what dimensions?” What are the objectives you seek to advance, and the risks and costs you hope to minimize? Let me urge you not to answer this question in overly complicated ways. Start with a list.

List the good things you want to see grow for those engaged with Head Start. List the bad things you want to slow. Good to grow, bad to slow. Don’t make it more complicated than that to start.

Let me offer a starter list to get you thinking:

### **Good to Grow**

- improve:
  - readiness of at-risk children to enter grade school on multiple domains,
  - subsequent success in school,
  - success post-school,
  - ability to handle social situations,
  - children’s health and safety;
- help home caregivers of Head Start children help their children;
- keep parent’s happy; and
- build good community relations.

### **Bad to Slow**

- Inadequate dental care, doctor’s visits, immunizations

- Unwanted incidents in center (e.g., accidents, infections, fires, poisoning, unreasonable levels of illness, sunburn)
- Chronic child absenteeism
- High teacher absenteeism, high teacher turnover rate
- Complaints
- High or rising costs

Good to grow and bad to slow. Decide what indicators you need to collect and share to move trends in the right direction.

I know I do not have the full list of good-to-grow and bad-to-slow objectives, nor the goal statements exactly right. That is not important for today. What is important is agreement on the need to find and create useful suites of indicators for outcomes of interest, collect and analyze them, and then share and use the analyses to make better decisions and take more informed action leading to better results for the children. It is not essential to collect all the indicators all the time. Also, some useful indicators have long lag times so others will need to collect them. It is, however, important to get started and get better collecting more useful near-term, intermediate, and longer-term outcome indicators than you currently do and then analyzing them and applying the insights of the analyses.

Next, figure out how to “tag” the data you collect to enable greater insights from the data. Note things such as the time, day, and date each measurement is collected. Note other attributes likely to influence outcomes of interest, such as the curricular program used by each child. Tap the rich body of expertise in unwanted incident management, such as transportation fatalities and fires, to find useful ways to characterize and analyze collected data – to slice and dice the data to discover likely progress and possible problems and to decide where deeper dives to search for causal factors that can be influenced should start.

Incredible technical developments have dramatically increased the ease and reduced the cost of collecting, analyzing, and sharing data and analyses. Consider how our smart phones time stamp and geo-code our actions and then pull the data together to help us decide what mode of transportation and route to take. The time is ripe for Head Start transformation.

**Value of Prescribed Practices.** You are probably thinking, “More indicators?” We already collect and report hundreds, maybe even thousands. You want us to collect more?

My first question is: are you learning from all those indicators ways to do better? If yes, does the value of what you are learning exceed the cost? If no, why are you collecting them? Is someone else using indicators you do not find useful, how, for what purpose, and is it working? If not, why continue to collect those?

Some of what you report pertains to monitoring compliance with required practices. In those cases, do you understand why the prescription make sense? For example, when my kids were in our cooperative day care, I recall a number of prescribed practices:

- change diapers only on the changing table,
- wash the table with a 1:10 bleach solution after each diaper change, and
- run the tap water each morning for 20 minutes to flush lead out of the pipes.

I understood the reasons for all of these prescriptions so complying with them did not bother me.

If you don't understand the reason for prescribed practices, start asking why they exist. Offer to participate in an effort to test the validity of specific compliance requirements against the objectives they were established to advance. If you think there is a better way to accomplish the objective, be part of a well-designed trial to test if your proposed alternate approach might produce better outcomes and/or be more cost-effective. If you find a better approach, work with others to figure out how to speed its adoption.

Note that I have chosen to use the word prescribed practices. I use this term intentionally to get people thinking about required practices the same way we think about prescriptions. In medicine, before we start a treatment, we want to know the evidence, or at least the compelling logic, behind them.

### III. Frequent Analysis

Of course, it is not enough to collect outcomes and other data. It is essential that data be analyzed and that analyses get used. It is time to increase the use of data and analyses. Work with government, foundations, researchers, and among yourselves to use data more in constructive ways!

The initial and primary use of data and analyses should be to find ways to do better. Data and analyses can also be used to motivate, but only if done with great care. Punish only the most serious laggards, those who persistently don't use data, analyses, and evaluation findings to find ways to do better, who persistently fare poorly compared to their demographic peers, and who fail to improve along multiple dimensions.

What do you want analyses to look for and how do you want to use the analyses?

Let me answer the second question first. In a continuous learning and improving system, you want analyses to trigger focused-follow up questions that lead you to discover causal factors you can try to influence. Also, you want analyses to identify areas of weaknesses for which effective weakness-reducing practices have been identified. The latter is what Salman Khan, founder of Khan Academy, talked about this morning when he described how the [College Board and Khan Academy have started working together](#) to analyze PSAT scores to identify individual student's areas of weakness and point them to free online lessons to address specific areas of weakness. [Preliminary analyses](#) suggest the Khan Academy courses have produced noteworthy testing (and learning) gains.

Let me also offer a few ideas about kinds of analyses likely to be useful. This is hardly an exhaustive set of ideas, but intended to get your thinking started.

#### **Promising practices: what some people call outliers, positive anomalies, or bright spots.**

Statisticians often look at statistical outliers and toss them out so as not to skew their conclusions. Some times, however, outliers can reveal a promising practice worth testing to see if it can be replicated. For example, you may want to look for centers serving, say, very low-income children of Latin American heritage outperforming other centers serving the same population.

You may also want to look for centers showing significant improvement (on one or more dimensions) not evident in others. When that happens, when you see what looks like a potentially promising practice, use it to trigger the very simple focused follow-up question, "*Why?*" (asked one or more times), to search for causal factors that can be influenced. To answer that question, you might visit or call the center to identify practices the center thinks might explain the difference (recognizing that sometimes an outlier is just a data entry error or statistical anomaly), or you might want to do additional analytics. If a center did something different that might explain its better result, recruit other centers to

try to replicate the practice(s) to see if they, too, show better results and if the practice is worth promoting for broader adoption.

**Patterns and variations across different subsets of what is being measured.** Consider, for example, looking at when accidents happen in Head Start centers. Are accidents, for example, more common in all centers after 2 in the afternoon? Are they higher in just one center? When analysis of the data reveals something like this, it should trigger the focused follow-up question: why is it happening and how do we prevent it?

**Inform choice.** Analyses can also be used to inform parent choice about centers if they are in a geographic location that affords choice or are choosing where to live or center choice of curriculum and other practices. Think about how *Consumer Reports* compares products to inform consumer choice. It realizes that buyers have different preferences. Some will trade off quality for lower price; others care about noise or dust or other peripheral attributes of a product or service other than its precise mission, whether transporting you or cleaning your house or dishes. Head Start parents may care a lot about center hours, for example. When using comparative data to inform choice, it is important to present the information fairly and in a format that is easy to interpret accurately.

#### **IV. Data and analyses are broadly shared**

This gets me to the topic of sharing data and analyses. I think of this as “UPS-ing the data.” Who needs what data when, where, and in what format to make better decisions? The UPS driver needs data to know when and where to deliver package and make decisions about the order and routing of deliveries. The person who loads the truck needs data to know which packages go on which trucks, and where on each truck to load each package. The central office needs data to find ways to increase net revenues; deliver packages quickly, affordably, and in good condition; employ well-qualified, committed employees; and keep UPS employees and the equipment they use safe and in good condition.

As we build a continuous learning and improvement system, we need to think about and answer these kinds of questions, and, of course, build information systems that make data collection, analysis, distribution, and visualization easier and far less costly. Technology advances slashing the costs of information collection, analyses, and communication have made this more feasible than ever, but these technology advances were neither necessary nor essential. The National Highway Traffic Safety Administration figured out how to use data to bring down traffic fatalities long before the technology advances of the last few decades. What is essential is the will and a commitment to think and act this way.

#### **V. Data-informed discussions and action**

It is not enough to collect, analyze, and share data. The information needs to be discussed and insights from it applied regularly and frequently, as Stuart Jones of Acorn Evaluation illustrated just moments ago with his terrific analysis of absentee patterns at one center and his discussion of how center leaders have shared and discussed the data with the front line. The center was able to bring the child absentee rate down, reduce reporting errors, and make the data more meaningful to teachers and other staff. Data-informed brainstorming shares lessons, invites innovation, and affords a ready venue for quickly deciding whether or not to pursue tests of new approaches to reduce problems and advance opportunities.

Data-rich discussions have helped bring about a revolution in policing and baseball. It is the story of Moneyball, started with Billy Beane using data, Sabermetrics, to raise the standing of the Oakland A’s far beyond what its budget would have predicted. For years, scouts used only their instincts to recruit players. Beane added statistics into the picture, with great results.

Data-informed discussions are helping to drive down hospital readmission and infection rates, as well as crime and veteran's homelessness. In states such as Massachusetts where data and benchmarking analyses are shared with teachers, principals, superintendents, and the public, it is helping to increase math and literacy scores and reduce education disparities. Data-rich discussions are also helping [clean up the Charles River in Boston](#).

Data-rich discussions encourage brainstorming to find what works better, problems needing attention, likely causes of problems, and likely factors contributing to progress. They inform action. They also keep folks in the field focused on what is being measured and discussed.

Who needs to be in these discussions and how often should they take place? Think about it. Then, get started and try it. Learn from your experience and share that experience with each other.

*One important note: I am not calling for data-driven action, but data-informed action. Human judgment, instincts, and values also influence decision-making, and should be informed by data, analyses, and well-designed measured trials.*

## **VI. Provider participation in measured trials**

Analyzing data about the current state of affairs will not tell us everything we need to know. Despite the declining costs of data collection, analysis, and sharing, measuring everything all the time is neither possible nor affordable. Moreover, analysis of data about the past only reveals the current and past state of things. It cannot reveal what could be in the future with the adoption of more effective new practices, although it can be useful for predictive purposes, something starting to happen in exciting ways in [building inspections](#), [economic development](#), [policing](#), and many other areas. To test what is possible, however, to accomplish new performance highs, well-designed measured field trials are needed. These can sometimes be complicated, but, much of the time, don't need to be.

Consider, for example, the signs we see in bathrooms encouraging us to wash our hands for at least 20 seconds. Proper hand washing, public health experts have figured out, is one of the easiest and most effective ways to slow the spread of disease. The question is: how effective are the signs in changing our hand-washing behavior? What percentage of folks who don't currently wash their hands before leaving the bathroom start washing their hands after the signs are posted? What percentage who already wash but only for short durations extend their hand-washing time to 20 seconds or longer? Consider how you might assess how well the signs work and whether they are worth the cost of their installation. We, of course, don't want to install cameras in bathrooms, but I am betting that with a little brainstorming, you could come up not only with a reasonable means to figure the value and cost of posting the signs but also whether changes in their timing, placement, design, and location affect their impact and cost-effectiveness.

Whether it is testing different hand-washing encouragement methods to prevent the spread of disease or testing different ways to boost literacy skills when English is not spoken at home, isn't it time to engage providers in a robust, vibrant learning and improving system that helps to identify problems and opportunities, design and test new practices, assess if those found to work in a few locations can work in others, and speed uptake of the most effective, cost-effective approaches?

## **VII. Public accountability and motivational mechanisms**

With all the debate about Head Start, wouldn't it be great if it was easy to find and understand Head Start and each center's goals and why they were chosen, key indicators being tracked, strategies and

why they were chosen, progress made, problems encountered, and planned next steps short and long term?

To my mind, these should be the essential first steps for accountability.

The second should be discussing the goals, data, strategies, and planned next steps with employees, parents, the community, delivery partners, and funders in a way that invites feedback and informs future goals and action. Goals and measurement can motivate when organizations are fairly compared and incentives systems carefully designed, encouraging data use to enlighten and point the way to improvement, not to punish.

### **Motivational mechanisms**

Given the limits of time, I am not going to discuss incentives and other motivational mechanisms today. However, especially given the legislatively mandated competition to identify and weed out the lowest performers in the Head Start program, this subject warrants future discussion as well as further exploration and application of evidence about effective incentive structures. That is a much easier discussion to have if you first establish clarity around objectives and related performance and other indicators, and commit to getting more value from data, analytics, and well-designed trials.

### **Next Steps**

It is not a small thing I am asking you to embrace and bring to life.

But it is incredibly important. Moreover, the alternative is dismal.

Be the change.

Happily, the National Head Start Association has made that commitment with the Moneyball for Head Start report, the Data Design Initiative, and other efforts. Happily, the Office of Head Start has issued new performance standards embracing outcomes-focused management and continuous improvement.

The question I have for you today is: will you be the change, too? Have I excited you about becoming change leaders to create the next generation Head Start program to serve the next generation of high-risk children with hopes that, in the generations that follow, fewer children will be in that category?

I hope the answer is yes. If so, please check out the [Data Design Initiative](#) page of the NHSA website and let [Victoria Jones](#) know you want to be part of this change. We need you, and now is the time.