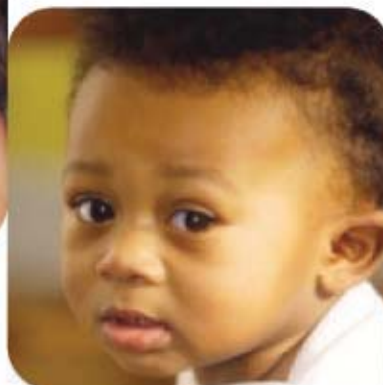


## Preliminary Working Paper



### **RESOURCES FOR MEASURING SERVICES AND OUTCOMES IN HEAD START PROGRAMS SERVING INFANTS AND TODDLERS**

**Prepared by:**

Ellen Eliason Kisker  
Kimberly Boller  
Charles Nagatoshi  
Christine Sciarrino  
Vinita Jethwani  
Teresa Zavitsky  
Melissa Ford  
John M. Love

Mathematica Policy Research, Inc.

**Prepared for:**

Rachel Chazan Cohen

Child Outcomes Research and Evaluation  
Office of Planning, Research, and Evaluation  
Administration for Children and Families  
U.S. Department of Health and Human Services

---

## **Abstract**

This document contains resources to help Head Start programs that serve pregnant women and families with infants and toddlers develop a performance measurement plan and carry out data collection that will support their continuous program improvement efforts. These performance measures activities should build upon existing screening and assessment activities required by the Head Start Program Performance Standards. This document discusses the importance and development of a comprehensive plan and presents profiles of instruments that may be useful to programs. Because we anticipate that it will be used under different circumstances for different purposes, we present the background information in a book format, and the entire document on a compact disk, to enable users to search for the sections and measures that apply to them. This format will also support the continued evolution of the document, which is intended to be a “living” document to which information on new instruments can be added, through which new tools and approaches can be shared, and in which other resources that individual programs find useful can be compiled. These materials will be most useful when used in consultation with an assessment expert.

THE INCLUSION OF AN INSTRUMENT IN THIS RESOURCE DOCUMENT DOES NOT CONSTITUTE ENDORSEMENT OF THE INSTRUMENT BY THE AUTHORS, MATHEMATICA POLICY RESEARCH, OR THE U.S. GOVERNMENT.

# Table of Contents

<b>I. PERFORMANCE MEASUREMENT</b>	<b>1</b>
<b>II. FORMULATING A PERFORMANCE MEASUREMENT PLAN</b>	<b>7</b>
<b>III. INFORMATION INCLUDED FOR EACH INSTRUMENT</b>	<b>21</b>
<b>REFERENCES</b>	<b>33</b>
<b>APPENDIX A:</b>	<b>35</b>
WORKSHEET TO HELP DEVELOP A COMPREHENSIVE PLAN FOR GATHERING AND ANALYZING DATA	
<b>APPENDIX B:</b>	<b>39</b>
PERFORMANCE MEASURES FRAMEWORK FOR HEAD START PROGRAMS SERVING INFANTS AND TODDLERS	
<b>APPENDIX C:</b>	<b>47</b>
LIST OF MEASURES USED IN THE NATIONAL EARLY HEAD START RESEARCH AND EVALUATION PROJECT	





## Performance Measurement

# Section 1



**Performance measurement** includes data collection and aggregation activities that give staff members the opportunity to look at how their program is doing, that is, to determine whether they are providing the services they intend to provide and how children and families are faring. Program staff can use this information in planning for continuous program improvement activities, and it can be shared with stakeholders such as parents and funders.

**H**ead Start, as a national laboratory for early childhood education, has been a leader in developing outcomes-oriented accountability. It began developing performance measures in 1995, partly to respond to the Government Performance and Results Act (GPRA) requirements. In 1997, the Head Start Bureau launched the Family and Child Experiences Survey (FACES) to collect data on child and family outcomes, as well as program services and management systems, for a large nationally representative sample of children and families in Head Start programs (Administration

on Children, Youth and Families 2001b). This study links the development of children and families with their experiences in Head Start. Following the reauthorization of Head Start in 1998, Head Start programs were required to include child outcomes in their self-assessment process by 2003.

Although Early Head Start programs have not been required to report child outcomes, many have started to try to define and measure outcomes, for several reasons. Some programs are doing so in conjunction with performance measurement in Head Start because they operate within Head Start programs. Some are responding to other funders' requirements. Finally, some simply want to improve their services to families with infants and toddlers. In Spring 2001, the Early Head Start Technical Work Group recommended that the Head Start Bureau move forward to develop performance measures for Head Start programs serving infants and toddlers (Early Head Start and Migrant Head Start programs) to support programs' efforts.

The development of new performance measures for Head Start programs serving pregnant women and families with infants and toddlers has two purposes. These are (1) to create performance measures sensitive to the developmental stages from pregnancy to age 3 that can be used at both national and local levels for learning about

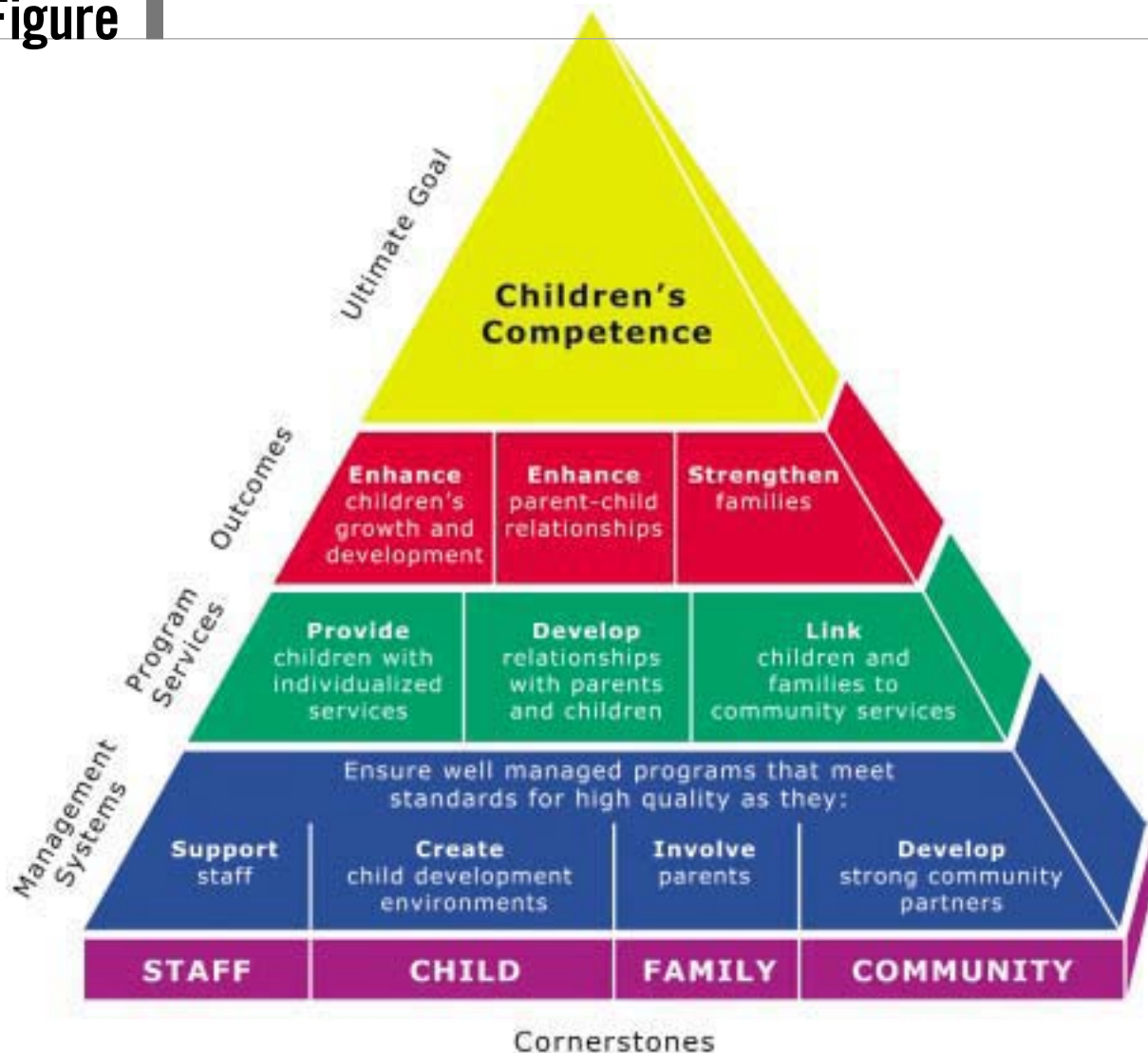
child and family experiences along with program services and management systems; and (2) to provide guidance to local programs in defining and measuring outcomes and using this information for continuous program improvement.

To fulfill these purposes, the Head Start pyramid, representing the Head Start performance measures framework, was expanded to reflect the unique features of programs serving infants and toddlers (see Figure 1 and Appendix C). The ultimate goal of Head Start programs that serve pregnant women and families with infants and toddlers is the same as that of Head Start programs that serve preschool children—enhancing children's competence.<sup>1</sup> The main outcomes supporting this ultimate goal—the blocks in the pyramid—have been expanded to reflect the central role of relationships in supporting attainment of this goal in families with younger children (U.S. Department of Health and Human Services [DHHS] 1994). Enhancing children's growth and development, enhancing parent-child relationships, and strengthening families as primary nurturers of their children are outcomes that support the ultimate goal of enhancing children's competence. Supporting these outcomes are objectives related to program services, including the objectives of providing children with individualized services, developing relationships with parents and children, and linking children and families to

<sup>1</sup> The Head Start goal, traditionally stated as children's "social competence," was shortened to "competence" because the interpretation of social competence in the Head Start pyramid has sometimes been narrower than was intended. Competence is the child's everyday effectiveness in dealing with his or her present environment and later responsibilities in school and life.

**Figure 1**

**Conceptual Framework for Programs Serving Infants and Toddlers**



needed community services. The objective of ensuring well-managed programs that meet standards for high quality in supporting staff, providing early childhood environments, involving parents, and developing community partnerships serves as the foundation of the pyramid. The pyramid rests on the four cornerstones recommended by the Advisory Committee on Services for Families with Infants and Toddlers—staff development, child development, family development, and community building (DHHS 1994).

The national Early Head Start Research and Evaluation project studied early program implementation and impacts on children and families (see Box 1). From this study, we learned that each program operates with its own theory of change—that is, a theory explaining how the services they provide will improve the child and family outcomes they are focusing on to meet the needs of children and families in their community. Although all programs must implement all aspects of the pyramid, the ways that programs



configure their management systems and program services (the bottom levels of the pyramid) vary widely. The specific child and family outcomes they focus on (the upper levels of the pyramid) also vary. As the program staff learns from continuous program improvement activities or as families' needs change, programs' theories of change may evolve. The evaluation showed that in general, patterns of program impacts reflected differences in theories of change. Home-based programs, which emphasized improving parenting and the home environment as an important path to improved child outcomes, had significant favorable impacts on parenting and child outcomes. Center-based programs, which emphasized direct services to children to improve outcomes, had favorable impacts on children and fewer significant impacts on parenting. The study also showed that programs that fully implemented key elements of the Head Start Program Performance Standards (elements related to all levels of the pyramid) had more favorable impacts on a wide range of outcomes than programs that were incompletely implemented (Administration for Children and Families 2002).

A *theory of change* is a belief or set of beliefs about how program services and other factors produce changes in the desired outcomes. The theory of change is the basis for designing and implementing program services. It may be explicit in program documents or staff discussions, or it may be implicit in the decisions of program designers and implementers.

*The pyramid provides a framework to help programs identify their own theory of change. Together, the pyramid and theory of change provide guidance for examining links between program services and outcomes to inform continuous program improvement activities.* The performance measures framework is general and does not identify specific program services and outcomes that should be measured. Individual programs must identify specific services and outcomes based on their own theory of change and select appropriate instruments to fit their available resources.

We do not recommend that programs collect data related to every element of the pyramid (at least to start with). We recommend that programs collect data related to each level of the pyramid (management systems, services, and outcomes), because having such information allows staff to link information on program characteristics and services with outcomes for children and families and to learn about both how well children and families are doing and how services might be improved to promote better outcomes.

*Selecting appropriate measures is an important and complex process. This document is intended to be a resource for programs undertaking this process.* It includes both screening instruments and assessment instruments that may be useful to programs as they explore how to approach performance measurement. We attempted to identify a wide range of instruments and select those likely to be most useful to Head Start programs serving pregnant women and families with infants and toddlers, but there are useful

tools that are not included in this review. For example, some kinds of assessment that are useful for individualizing services, such as portfolio assessment, are not included (Martin 1999). In addition, new measures are constantly emerging, and very recent measures may not be found here. Some programs may wish to use this resource document in consultation with an expert on screening and assessment to help them develop a plan and select instruments. (For more explanation of how we selected instruments for this resource document, see page 17.)

***This resource document will be most useful to programs if it is used in conjunction with a comprehensive plan for performance measurement.*** In the next section, we discuss the importance of developing a plan and the elements that make up such a plan. In the final section, we describe how instruments were selected for this resource document, explain the information provided for each instrument, and present profiles of instruments that may be useful to Head Start programs that serve pregnant women and families with infants and toddlers.

# Box 1

## MAKING A DIFFERENCE IN THE LIVES OF INFANTS AND TODDLERS AND THEIR FAMILIES: THE IMPACTS OF EARLY HEAD START

A rigorous evaluation of Early Head Start services in 17 programs selected from the first groups of programs funded showed they had significant favorable impacts on a wide range of parent and child outcomes, some with implications for children's later school success. Findings from the study (*Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early Head Start*), using data gathered when children were age 3 and had completed the program, show that the programs sustained and broadened the pattern of impacts reported when children were age 2 (*Building Their Futures: How Early Head Start Programs Are Enhancing the Lives of Infants and Toddlers in Low-Income Families*, 2001). All Early Head Start evaluation reports are available online at [[http://www.acf.hhs.gov/programs/core/ongoing\\_research/ehs/ehs\\_intro.html](http://www.acf.hhs.gov/programs/core/ongoing_research/ehs/ehs_intro.html)].

**Early Head Start Improved Outcomes.** The national evaluation conducted by Mathematica Policy Research, Inc. and Columbia University's Center for Children and Families at Teachers College, in collaboration with the Early Head Start Research Consortium, reported that 3-year-old Early Head Start children performed significantly better on measures of cognitive, language, and social-emotional development than a randomly assigned control group. While children who participated in Early Head Start performed better than their peers who did not receive Early Head Start on all aspects of development that were assessed, both groups lagged behind 2- and 3-year-old children nationally. This may, in part, be due to the fact that Early Head Start programs must reserve at least 10 percent of their slots for children with disabilities, including those with developmental delays who score at the lower end of the distribution. Parents in Early Head Start scored significantly better than control group parents on measures of many aspects of the home environment and parenting behav-

ior. Furthermore, Early Head Start programs enhanced parents' progress toward self-sufficiency years. Early Head Start fathers benefited as well. Although the overall impacts of Early Head Start were generally modest, the pattern of favorable findings across outcomes in a wide range of key domains is promising.

**Full Implementation Was Important.** The impacts on children and parents were consistent with the substantial difference the programs made in families' receipt of services. Early Head Start families were significantly more likely than control families to receive a wide variety of services, but especially to receive intensive services, and to receive intensive services that focused on child development and parenting.

Implementing the Head Start Performance Standards early and well is important for maximizing impacts on children and families. The research programs were systematically rated according to the extent to which they implemented key elements of the Performance Standards. Those that fully implemented the standards demonstrated a broader pattern of significant impacts than did the programs that did not reach full implementation during the evaluation period.

**Patterns of Impacts Were Consistent with Theories of Change.** Analyses that compared the contribution of impacts on parenting when children were age 2 to impacts on child outcomes at age 3 in programs providing home-based or a mix of home- and center-based services generally provided support for the theories of change that staff in those programs described, with some of the effects of the program on 3-year-old children being associated with the effects on parenting when they were age 2.



## FORMULATING A PERFORMANCE MEASUREMENT PLAN

# Section 2



The Head Start Program Performance Standards require programs to conduct screening and assessment activities and emphasize their importance for individualizing services and informing continuous program improvement. The performance standards allow considerable flexibility in how programs meet the requirements. In response to the requirements in the performance standards, Head Start programs serving pregnant women and families with infants and toddlers are already collecting data. Head Start programs serving pregnant women and families with infants and toddlers may use screening and assessment instruments for different purposes, including:

- **To Support Development and Learning.** Child assessment results can tell caregivers and teachers what each child can do and what he or she is ready to learn next. Family or parent assessment results can help program staff identify family or parent strengths, needs, and concerns and tailor services to the family. Over time, assessment results can demonstrate how each child and family is progressing. Assessment results can also help staff communicate with family members

about their children's needs and progress, as well as their own needs and progress. This information can help staff individualize services and improve them over time.

- ***To Identify Special Needs and Concerns.***

Because of the cost of in-depth assessments, screening is usually the first step in identifying special needs. Children or families for whom an in-depth assessment is indicated are often referred to a physician or other expert for a complete evaluation. Screening results provide the information needed for referrals to other agencies, such as Part C agencies, to obtain services for children with disabilities.

- ***To Evaluate the Program and Monitor Trends.***

For this purpose, child and family screening and assessment data may be aggregated and used to inform continuous program improvement efforts. Program assessment data and feedback from Head Start Bureau monitoring may also be used for this purpose. Aggregated screening and assessment results can inform staff about how well the program is meeting child, family, and community needs.

Screening and assessment results for individual children, along with other information from parents and caregivers, are needed to help staff tailor services for those children. When the same instruments are used for all children in a program, aggregating data across families can provide a picture of how children and families in the program are doing overall. When this aggregate information on child and family outcomes is linked to information on services and other pro-

gram characteristics, it can provide insights that are useful to staff members in their continuous program improvement efforts. Aggregating data provides a picture of how children and families and the program are doing (that is, it measures the program's performance). Over time, aggregated data can be used to track changes in child and family functioning, which along with information about changes in program services or characteristics, can help staff learn about what works best for particular types of children and families. The aggregated information on how children and families in the program are doing can also be useful for meeting other funders' reporting requirements.

***Screening.*** Screening is a generic term referring to activities designed to identify individuals who have a high probability of exhibiting delayed, abnormal, or problematic development. The screening is intended to identify problems at an early stage and identify individuals for whom further, in-depth assessment activities are needed.

***Assessment.*** Assessment is a generic term referring to procedures for obtaining systematic information on a child's, parent's, family's, or program's strengths or needs. As noted in Chapter I, the Head Start Program Performance Standards focus on the child and family assessment purposes of identifying "(i) the child's unique strengths and needs and the services appropriate to meet those needs; and (ii) the resources, priorities, and concerns of the family and the supports and services necessary to enhance the family's capacity to meet the developmental needs of their child."

*[See the Early Head Start National Resource Center's Technical Assistance Paper #4 for more detail on screening and assessment activities.]*

**Performance measurement** includes data collection and aggregation activities that give staff members the opportunity to look at how their program is doing, that is, to determine whether they are providing the services they intend to provide and to learn how children and families are faring. Program staff can use this information in planning for continuous program improvement activities, and it can be shared with stakeholders such as parents and funders.

Screening results for individual children and families are useful for deciding whether further assessment is needed. Screening results can also be aggregated to provide information on the extent of potential problems in the population and the need for in-depth assessments among children and families overall. Aggregated screening results and information on referrals can inform a program as to whether these in-depth assessments are happening if the program's tracking system does not provide this information.

Assessment results for individual children can be used for planning services; in addition, they can often be aggregated to provide broader information on child outcomes. Some approaches to assessment are valuable for individualizing services, but cannot be aggregated unless they are translated into another form. For example, portfolio assessment can be extremely useful for individualizing services for children, but unless a systematic

way of coding the information is developed, the results cannot be aggregated.

Programs face both a great opportunity and a significant challenge as they consider performance measurement. They have an opportunity to select instruments and collect data that best meet the needs of their families and their program, and they face the challenge of figuring out just what those instruments should be, how they should be administered, and how the information collected using those instruments should be analyzed. Meeting that challenge effectively requires programs to develop a plan that considers the purpose(s) for which the data will be used, what data are already being collected and additional data that need to be collected, and how data will be aggregated and analyzed for continuous program improvement. Taking the time to develop a comprehensive plan will help ensure that the program's resources for the required ongoing screening and assessment of individual children and families, as well as the new performance measurement and data aggregation activities, are used well and that the activities are as useful as possible.

***To formulate a comprehensive plan, programs need a clear understanding of why performance measurement is needed.*** What does the staff need to know to determine whether the program is achieving its intended outcomes and to individualize or optimize services for children and families? What information is essential? What, in addition, would be good to know? Some questions programs might consider to help reach this understanding are:



















































































































































































































































































































































































































































































