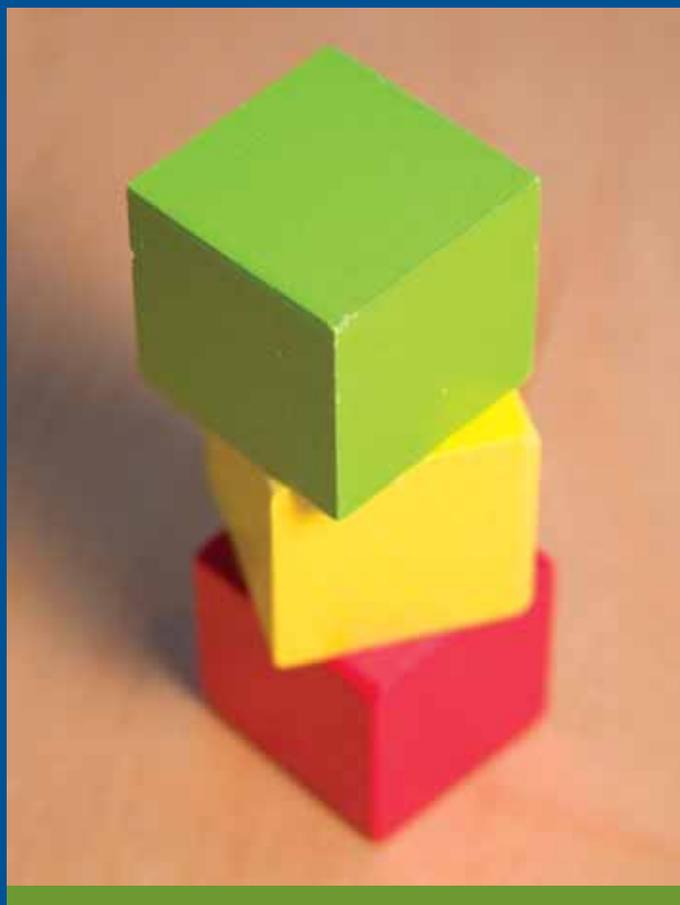


Taking Stock: Assessing and Improving Early Childhood Learning and Program Quality



THE REPORT OF THE NATIONAL EARLY
CHILDHOOD ACCOUNTABILITY TASK FORCE

Table of Contents

Task Force Members	2
Executive Summary	3
Preface	13
Chapter One	
Clarifying the Challenges	17
Chapter Two	
Framing a New Approach	29
Chapter Three	
Designing an Accountability and Improvement System	37
Chapter Four	
Differing Viewpoints on Using Child Assessment Data for Local Agency Accountability	61
Chapter Five	
A Call to Action	69
Additional Comments	75
Appendix A	
Commissioned Papers	81
Appendix B	
Approaches to Child and Program Assessment	83
Appendix C	
Examples of State Assessment and Program Improvement Efforts	87
Notes	95
Acknowledgments	99

Task Force Members and Project Staff

- Dr. Sharon Lynn Kagan, Task Force Chair, Professor and Associate Dean for Policy, Teachers College, Columbia University, New York, NY and Professor Adjunct, Yale University, New Haven, CT
- Dr. Eugene Garcia, Task Force Vice-Chair, Vice President for Education Partnerships, Arizona State University, Tempe, AZ
- Dr. W. Steven Barnett, Director, National Institute for Early Education Research and Professor, Rutgers, The State University of New Jersey, New Brunswick, NJ
- Ms. Barbara Bowman, Director, Office of Early Childhood Education, Chicago Public Schools and Professor, Erikson Institute, Chicago, IL
- Dr. Mary Beth Bruder, Professor of Pediatrics, University of Connecticut, Farmington, CT
- Dr. Lindy Buch, Director, Office of Early Childhood Education and Family Services, Michigan Department of Education, Lansing, MI
- Dr. Maryann Santos de Barona, Professor, College of Education, Arizona State University, Tempe, AZ
- Ms. Harriet Dichter, Deputy Secretary, Office of Child Development and Early Learning, Pennsylvania Departments of Education and Public Welfare, Harrisburg, PA
- Mr. Mark Friedman, Executive Director, The Fiscal Policy Studies Institute, Santa Fe, NM
- Dr. Jacqueline Jones, Assistant Commissioner, Division of Early Childhood Education, New Jersey State Department of Education, Trenton, NJ
- Dr. Joan Lombardi, Director, The Children's Project, Washington, DC
- Dr. Samuel Meisels, President, Erikson Institute, Chicago, IL
- Ms. Marsha Moore, Commissioner, Georgia Department of Early Care and Learning, Atlanta, GA
- Dr. Robert Pianta, Professor, University of Virginia, Charlottesville, VA
- Dr. Donald Rock, Senior Research Scientist, Educational Testing Service, Princeton, NJ
- Dr. Thomas Schultz, Project Director, The Pew Charitable Trusts, Philadelphia, PA
- Ms. Debi Khasnabis, Research Assistant, University of Michigan, Ann Arbor, MI

The Task Force was created through the generous support of The Pew Charitable Trusts, the Foundation for Child Development and the Joyce Foundation. The positions and views expressed in this report do not necessarily reflect those of the funders.

Executive Summary



Unprecedented attention to young children has ushered in a new era for early childhood education. Parents are more concerned than ever before about their children's learning, development, and readiness for school. Early childhood teachers are taking on the challenges of serving all children equitably and well. And policymakers are looking carefully at the outcomes reported for children participating in publicly funded early education programs. Motivated by these concerns and by the growing emphasis on accountability, parents, teachers and policymakers all want more information as they make decisions on how to foster children's early learning and development.

These demands for information come at a time when early childhood educators are uneasy about the effect that increased performance demands may have on young children's development and early childhood practice. At the same time, early educators are aware of the potential of well-designed assessment and evaluation efforts to enhance the credibility of early childhood programs, and support investments in program improvement and expansion. But accountability requires great care. Poorly conceived accountability initiatives can generate misleading feedback, impose onerous burdens, and lead to misguided decisions. And accountability should not stand alone. Linking accountability efforts to program improvement efforts and resources is essential to warrant the effort needed to gather assessment data. Clearly, issues surrounding early childhood accountability and improvement are important, challenging and controversial.

Task Force Charge and Process

Given the importance and timeliness of the issue, the intensity of early educators' concerns, and the imperative for action, the National Early Childhood Accountability Task Force was created to build a comprehensive vision for future accountability efforts. In April 2005, The Pew Charitable Trusts, the Foundation for Child Development, and the Joyce Foundation charged the Task Force to

Develop recommendations for a state accountability system for early education programs for prekindergarten children and for linking such efforts to standards-based assessment efforts in kindergarten and the primary grades.

The charge called upon the Task Force to review current state and federal initiatives, to provide guidance on tough issues and controversial questions, and to be forthright in recommending steps that states should and should not take as they embark on accountability and program improvement efforts.

To accomplish this task, the Task Force assembled a team of well-respected individuals, including leading experts on early childhood assessment, program evaluation, early childhood pedagogy and practice, and advocacy. We focused on designing the best possible approaches to using assessments to strengthen the early childhood field. Our report offers an across-the-board, thoughtful, long-term vision. We hope that our recommendations will combat the tendency for assessment and accountability mandates to proliferate in haphazard, uncoordinated, and potentially counterproductive ways.



The Challenges

Before presenting the recommendations of the Task Force, it is important to offer a clear definition of an early childhood accountability system and acknowledge the specific challenges the recommendations seek to address.

The Task Force defined a State Early Childhood Accountability and Improvement System as

A system of standards-based assessments of (a) children's development and learning and (b) program quality, designed to inform state policy decisions, investments, and improvement efforts for early education programs for three- and four-year-old children, linked to a continuum of kindergarten through third grade standards, curriculum, assessments, and program improvement efforts.

Establishing such a system is daunting in light of the challenges that face American early childhood education. In particular, our report highlights four clusters of challenges:

- **Structural challenges** related to the fragmented “non-system” of programs for preschool-aged children and disjointed early childhood and public education policies.
- **Conceptual challenges** related to long-standing approaches to early childhood assessment, program evaluation, and program management.
- **Technical challenges** related to the need for tools and methods appropriate for assessing increasingly diverse populations of young children and varied types of programs.
- **Resource challenges** related to limitations and inequities in funding for early childhood programs and infrastructure efforts.

Framing the Work

To fulfill our charge and address these challenges, the Task Force agreed on five framing recommendations:

1. States should develop a unified system of early childhood education that includes a single, coherent system of standards, assessments, data, and professional development efforts across all categorical programs and funding streams.
2. States should align high-quality and comprehensive standards, curriculum, and assessments as a continuum from prekindergarten through grade 3.
3. States should assure that all child and program assessments are valid and reliable, meet high psychometric standards, and are well suited for their intended purpose. Data analysis and reporting methods should incorporate state-of-the-art methods to accurately and fairly document the performance of programs, including, where feasible, information from assessments of children and program quality together:
 - Data from assessments of children should not be reported without data on the programs that serve them.
 - Reporting on program quality should highlight attributes of classroom quality, instructional practices, and teacher-child interactions that are most highly correlated with enhancing children’s progress in learning and development.
 - Reporting on child assessments should highlight children’s progress over time (or the “value-added” contributions of programs) as well as their end-of-program status.

4. States should support the full inclusion of all children in accountability and improvement efforts, with appropriate accommodation of assessments to fully document their progress and status:
 - Young English Language Learners should be evaluated in both their primary language and the language(s) of instruction.
 - Adaptations in child assessment tools and procedures should be made to allow children with disabilities to participate in the same assessments as their peers and to allow a valid assessment of their knowledge and abilities.
5. States should provide adequate resources to enable programs to meet performance standards, and to support accurate, credible, and useful assessments and effective program improvement efforts.



Recommended System

Based on these framing recommendations, the Task Force designed an Early Childhood Accountability and Improvement System that is powerful and flexible enough to allow any state to adapt it to meet its priorities and needs. The overall design comprises three primary building blocks:

- **System Infrastructure.** The design begins with the vital supports needed to ensure high-quality assessments, timely, accurate reporting, and appropriate understanding and use of assessment data.
- **Assessment/Program Improvement Approaches.** Recognizing the vast diversity among the states, the plan provides multiple approaches for assessing and improving early childhood programs. Each approach is designed to meet different state needs, so that states can select or combine approaches as they deem appropriate.
- **Steps toward Coherent PreK-Grade 3 Accountability Efforts.** Finally, the design proposes that states work to align and integrate standards and curricula from prekindergarten through grade 3, thereby fostering the continuity of instruction, assessment, and program improvement efforts throughout children's early years.

Infrastructure

The proposed system begins with an Infrastructure comprising four interconnected parts. These parts work together to support successful assessments, accurate reporting, and effective program improvement efforts:

- **Early Learning and Program Quality Standards.** Comprehensive, well-articulated standards for children’s learning and program quality should be aligned with curricula and assessments. These standards provide a context for understanding assessment information and guiding program improvement efforts.
- **Program Rating and Improvement System.** An inclusive Program Rating and Improvement System is needed to assess, document and improve the quality of all forms of early education programs. A Program Rating and Improvement System would guide on-site monitoring of the quality of teaching and learning opportunities and document local programs’ adherence to program quality standards. The system would provide information to parents and consumers, offer resources to improve quality, and could provide incentives to reward higher levels of quality.
- **Professional Development System.** The aim of a Professional Development System is to create a consistent, accessible approach to professional development for everyone in a state who works with young children. In particular, the Professional Development System would help teachers and program managers administer assessments, interpret data, and use child and program assessment data for program improvement.
- **Data Management and Reporting System.** A coherent accountability and improvement system hinges on a well-maintained, integrated, user-friendly database on children’s characteristics, staff and program characteristics, and assessment information. A Data Management and Reporting System would provide for quality assurance and integrity of data and generate reports suited to varied audiences and purposes. It would make it possible for leaders and planners to look at data at many levels, including statewide, for different state early education programs, for local communities and for individual local agencies. A unified system of child identification numbers would allow tracking of children’s program experiences and progress in learning and development across the preK-grade 3 years.

The Approaches

Every state in the nation is faced with the challenge of designing an approach to early childhood accountability and improvement. Their approaches to accountability, like the services they offer, differ markedly. States vary in terms of what they want to know, how they plan to use the information, and in the human and fiscal resources they can allocate to the development of an accountability and improvement system. The Task Force therefore developed several approaches that respond to the diverse circumstances, needs and interests of states. In the full report, each approach is described in depth, including specific policy questions it addresses, what data are to be collected, designs for data collection, how assessment information can be used to improve programs, and, most importantly, key challenges, concerns and safeguards regarding potential misuse of assessment information. The Task Force recognizes that several of these approaches may be difficult to implement immediately and would require careful planning, as well as new state investments, in order to work as intended.

- **Child Population.** What is the status of all children in the state? Based on assessing representative samples of children, this approach provides information on the developmental status and demographic characteristics of *all young children* in a state, regardless of whether they participate in early childhood services. This information can pinpoint areas where children are in need of supports, and therefore can guide overall systems planning; these data can also inform decision-making on state investments and inform public/private and interagency initiatives. It also can provide baseline information for public education planning efforts. It is important to note that the Child Population Approach does not report data on individual children.
- **Program Population.** What is the quality of all early education programs? This approach provides information on the quality of *all forms of early education services* in a state as well as data on the early childhood workforce. As with the Child Population Approach, this information can guide planning for overall systems of services for all young children, inform decision-making on state investments, and inform public/private and interagency initiatives. It also can provide baseline information for public education planning efforts.
- **State Program Evaluation.** What is the quality of and how well are children progressing in specific state early childhood programs? This approach would apply rigorous evaluation methods to report on program quality and child assessments for a *specific type of state program* (e.g., a state prekindergarten program). Information from these evaluations can guide efforts to strengthen state programs or refine program policies. Results can inform state decisions about funding

different types of programs. For many, this is the preferred approach because it combines data on children and on programs.

- **Local Agency Quality.** What is the quality of services in local agencies? This approach provides information on program quality at the level of *individual local agencies*, whether states elect to fund school districts, local governmental units, for-profit, nonprofit, or faith-based agencies. This information can guide decisions about targeting technical assistance to strengthen quality at specific sites, awarding incentives to recognize program improvement, and funding decisions by state agencies.

The Task Force also discussed at length an additional approach that would examine both program quality and children’s learning and development in local agencies, but we reached no consensus on its feasibility or desirability. Deliberations focused on the question of whether child assessment data should be used for local agency accountability. Among the serious concerns raised are high costs, the lack of appropriate assessment tools, and, most seriously, the potential misuse of data.

Creating Coherent PreK-Grade 3 Accountability Efforts

To conclude our proposed design, we offer recommendations to enhance continuity in accountability and improvement efforts from the preschool years through grade 3. Our proposal is intended to better align high-quality, comprehensive standards and support efforts to study the progress of children and the quality of learning opportunities across the preK-grade 3 years. State leadership can set the stage to enable early childhood and elementary school educators to work together in reviewing assessment data and using the findings to strengthen teaching, learning, and professional development efforts.



What are the Benefits?

The Task Force recommendations, taken together, provide a durable, adaptable framework that any state can use to create an effective accountability and improvement system tailored to its needs and priorities. The energies and investments required to accomplish this are substantial, but the anticipated benefits are even greater. The benefits include

- More relevant and credible data to guide state investments in services for young children;
- More targeted efforts to strengthen equity, informed by richer and more accurate evidence on the extent to which early childhood programs are providing quality services and helping subgroups of children progress;
- Enhanced credibility for the early childhood profession based on expanded public awareness of how early childhood services contribute to the public welfare;
- Stronger partnerships between early childhood programs and public schools to build a preK-grade 3 continuum of enriched learning opportunities;
- Higher quality learning experiences for children, as states support well resourced, evidence-based program improvement and professional development efforts;
- Improved outcomes for all children as accountability and program improvement efforts help states build overall systems of high-quality early education services.

A Call to Action

A task force can recommend, but realizing its vision requires leadership, collaboration, and investment. To create strong early childhood accountability and improvement systems, people and organizations need to engage in new ways of working across categorical programs, invest in quality assessment and program improvement efforts, and advance their own learning about issues of assessment tools, data, and analysis.

- **Governors and state legislators** should invest in high-quality early childhood programs and a strong accountability and improvement system. We recommend allocating from 2 to 5 percent of all program funding to support our recommended infrastructure of standards, program rating, professional development and data management efforts and to implement varied assessment and program improvement approaches.

- **State agencies** should develop and implement a strategic plan for a coherent early childhood accountability and program improvement system; support local preK-grade 3 partnerships; and work toward a robust, positive, and rigorous culture for early childhood accountability efforts.
- **Federal government agencies** should carry out a data harmonization initiative to allow information systems for child care, Head Start, and early childhood special education services to mesh with each other and with data generated from state early education programs. They should also invest in research and development efforts to support more coherent and effective state accountability and program improvement systems.
- **Local early childhood agencies** should create opportunities for teachers and managers to study and discuss child and program assessments, and to use the data to enrich opportunities for children's learning and development. In addition, they should initiate dialogue with local school districts around child assessment spanning preK to grade 3, and they should share data on demographic trends and learning opportunities.

A Compelling Need

People and organizations across the nation are already hard at work creating systems of high-quality and accountable early learning opportunities for America's young children. But all too often, their work is hampered by organizational fragmentation, infrastructure gaps, and inadequate tools and methods. The National Early Childhood Accountability Task Force has sought to help move our nation toward more coherent and effective early childhood accountability and improvement systems.

We know that readers may not agree with every approach or every recommendation. Indeed, Task Force members did not agree with every detail of the approaches presented. We also know that some readers approach accountability efforts in the realm of early childhood education with deep misgivings. While we understand these concerns, members of the Task Force concur that when accountability efforts are of high quality, when they safeguard children, and when they are used in the service of program improvement, they can contribute powerfully to make America's early education fair, just, and equitable, and among the best in the world. Our work is aimed toward that end, and it is dedicated to those who have worked, and continue to work, toward realizing that vision.

Preface



More than ever before, Americans appreciate the significance of the learning that occurs in the early years. In fact, early childhood is now often considered to be the most critical period of human development, and early education is widely regarded as a fundamental accelerator of children’s educational and life trajectories. A growing body of hard evidence bolsters this view:

- High-quality learning opportunities, beginning as early as possible and sustained over time, increase the odds that children will succeed in school and become productive citizens.¹
- Substantial achievement gaps exist between subgroups of children at the time they enter kindergarten and earlier.²
- Most of the centers and settings where young children are cared for and educated are not providing high-quality levels of teaching, learning and developmental supports, typically due to inadequate resources and a workforce with low levels of formal training and compensation, and high rates of turnover.³

This evidence has led states to introduce a variety of leadership initiatives to improve early learning opportunities for all children, including the following approaches:

Building overall systems of services for all young children. These efforts include organizing new units of state government to coordinate a variety of categorical programs for young children, developing new public-private investment strategies, and sponsoring comprehensive planning efforts at the state and community levels.⁴

Developing and funding specific programs. Most states have focused on efforts to provide high-quality, voluntary prekindergarten programs for all young children. In 2005–06, 38 states sponsored programs for more than 940,000 children at a total cost of nearly \$3.3 billion. Over the last four years, these investments increased prekindergarten funding by \$1.2 billion and swelled enrollments by more than 260,000 children.⁵

Providing technical assistance to improve quality and outcomes in local early childhood agencies. State agencies are responsible for overseeing, reviewing and mounting efforts to improve the quality of services to young children in a diverse set of local agencies, including school districts, for-profit, and nonprofit organizations.

Connecting early childhood, elementary, secondary and higher education into a more seamless P-21 continuum.⁶ A number of states are leading efforts to connect children’s educational and life experiences across the developmental spectrum. These efforts increasingly include a focus on linking early childhood programs and primary-grade schooling.



Focus on Accountability

These new state initiatives have led to expanded efforts to use assessment data to improve early childhood program effectiveness and enhance children’s learning and development. States are engaged in a wide array of new efforts to establish standards for and assess early childhood program quality and children’s learning and development.

Forty-eight states and the District of Columbia have developed early learning guidelines specifying what preschool children should know and be able to do, and the two remaining states are in the process of creating similar documents.⁷ Many states have gone further: 13 states are engaged in assessments of prekindergarten children, and 17 states have child assessment efforts that involve kindergarten children.⁸

Many states are focusing on programs, as well. For example, program evaluation, a key approach to accountability and program improvement, is widespread. At least 25 states have conducted evaluations of the quality, outcomes, and effectiveness of their prekindergarten programs.⁹ All 38 states that sponsor prekindergarten programs have developed minimum standards for the quality of program services, and 30 states are monitoring programs’ adherence to these standards.¹⁰ Seventeen states joined in a recent

collaborative effort to develop a common framework of school readiness indicators.¹¹ Finally, at least 42 states are developing or implementing Quality Rating Systems to assess child care agencies against multiple tiers of program standards, tied to financial incentives and the provision of information on program quality to parents and the public.¹²

Complementing these state initiatives—and complicating them—are significant federal policy changes in the areas of early childhood assessment and program accountability. These changes include major, new, and some quite contentious child assessment initiatives in Head Start, Even Start and programs for young children with disabilities.¹³ As explained in Chapter One of this report, this plethora of efforts is hard to manage and coordinate. The complexities are not limited to management. As they conceptualize and plan assessment and accountability initiatives involving young children, states encounter many views—often passionately held. Among those with strong opinions are early childhood leaders and experts who often express deep concern about the potential misuse of child assessment data and, in particular, the use of results to reward or punish local programs on the basis of their children’s test scores.¹⁴ In short, issues surrounding standards-based assessments and early childhood accountability policies are important, challenging and controversial.

Despite controversy, America has a vital national interest in bolstering the capacity of early childhood programs to prepare children to succeed in school and to narrow the achievement gaps that have been shown to begin early in life. Well-designed accountability efforts can enhance the credibility of early childhood programs, guide program improvements and generate data to support continued public investment. On the other hand, poorly conceived accountability initiatives can produce misleading feedback, impose onerous burdens, and lead to misguided decisions. Our challenge is to apply the lessons from past and current efforts to develop **a fresh, comprehensive, long-term vision for early childhood accountability and program improvement.**

Task Force Charge and Process

Given the importance and timeliness of these issues, the National Early Childhood Accountability Task Force was created by The Pew Charitable Trusts, the Foundation for Child Development, and the Joyce Foundation and charged to

Develop recommendations for a state accountability system for early education programs for prekindergarten children and for linking such efforts to standards-based assessment efforts in kindergarten and the primary grades.



The charge called upon the Task Force to review current state and federal initiatives, provide guidance on tough issues and controversial questions, and to be forthright in recommending steps that states should and should not take as they embark on improvement and accountability efforts.

To accomplish this task, the Task Force assembled a team of well-respected individuals, including leading experts on early childhood assessment, program evaluation, early childhood pedagogy and practice, and advocacy. A diverse group, these individuals came to the Task Force from academia, city and state departments of education, think tanks, and research and evaluation firms. From 2005 to 2007, the Task Force met five times and delved deeply into the conceptual and technical challenges of early childhood accountability and program improvement. Focus groups, interviews, 13 commissioned papers, and presentations from invited experts informed the deliberations.¹⁵

From the start, we eschewed polarizing debates about whether accountability has any place in early education or whether it offers a “silver bullet” solution for improving the care and education of American’s young children. Instead, we focused on designing the most powerful, effective and efficient approaches to strengthen the early childhood field. We hope that our recommendations will combat the tendency for assessment and accountability mandates to proliferate in haphazard, uncoordinated, and potentially counterproductive ways.

Our goal was to develop a consensus position while honoring the diverse perspectives of Task Force members. Indeed, although not all members agreed with every detail, the Task Force reached consensus on an overall framework to guide state accountability planning and implementation efforts. In a few areas where opinions were especially disparate, our report reflects the different views that surfaced in our deliberations. We approached our work with humility, guided by three core commitments: we are committed to responding to the questions of policymakers and the public, enhancing the effectiveness of early education programs, and, above all, safeguarding and advancing the interests of children.

Clarifying the Challenges



As they design and implement early childhood accountability and program improvement efforts, states face crucial challenges:

- **Structural challenges** related to the fragmented non-system of programs for preschool-aged children and disjointed early childhood and public education policies.
- **Conceptual challenges** related to historic approaches to assessment of young children and early childhood programs.
- **Technical challenges** related to the need for tools and methods appropriate for assessing increasingly diverse populations of young children and varied types of local agencies.
- **Resource challenges** related to limitations and inequities in funding for early childhood programs, personnel and key support efforts.

This chapter takes a close look at each of these four areas of challenges.

Structural Challenges: Moving toward Coherence and Continuity

States face two key structural challenges: contending with multiple early childhood accountability mandates, and addressing discontinuities between prekindergarten and elementary (K-3) education.

Contending with Multiple Mandates

Today's early childhood accountability and assessment efforts are fragmented and uncoordinated for reasons that are largely historical. Virtually all of these efforts flow from four major—and historically separate—categorical programs: Head Start, child care, early childhood special education, and state-funded prekindergarten. Each is governed by specific legislative and regulatory requirements. These programs take different approaches to child and program standards, assessment and reporting requirements, program monitoring, technical assistance, and professional development.

In short, the programs have diverse approaches to standards and assessments. Moreover, they exist in a context that is highly dynamic. In the last five years alone, the early childhood field has seen new state early learning guidelines; rapidly growing state efforts to establish Quality Rating Systems to assess and improve program quality; and new child assessment initiatives in states, Head Start, and early childhood special education. At the same time, the nation's largest early childhood professional organization, the National Association for the Education of Young Children, has restructured its voluntary Program Accreditation system.

Figure 1

Current Early Childhood Standards and Assessments¹⁶

	Child Care	Head Start	State-Funded PreK	Special Education
Program Quality Standards	State Licensing Standards (49 states) State Quality Rating & Improvement Systems (QRIS) (14 states + 29 pilots) ¹⁷	Federal Program Performance Standards	State Program Standards (39 states)	Federal IDEA regulations State Program Standards
Assessing Local Program Quality	State Licensing Visits State QRIS Assessments	Federal PRISM Monitoring	State Program Monitoring (30 states)	State Program Reviews
Standards for Children's Learning	State Early Learning Guidelines (49 states)	Federal Child Outcomes Framework	State Early Learning Guidelines (49 states)	3 functional goals (Federal)
Child Assessments	No current requirements	Federal National Reporting System ¹⁸ (projected to be suspended)	State PreK Assessments (13 states) State Kindergarten Assessments (17 states)	States report % of children in 5 categories on 3 goals
Research/Program Evaluations	Yes	Yes	Yes	Yes
↓ ↓ ↓ ↓				
Kindergarten-to-Grade 3 Standards, Assessments, Data				

For state and local managers and practitioners, this fragmentation, coupled with inadequate funding to meet quality standards, is a daily fact of life. Figure 1 represents the maze they must negotiate as they work to understand and implement many new initiatives related to standards, assessments and the program improvement efforts they inform. This policy labyrinth creates many complications:

- **Responding to multiple funders' requirements.** Many school districts and other local agencies receive funding from several of these funding streams, and therefore struggle to respond to multiple assessments, standards, reporting requirements, and monitoring reviews. Program managers are understandably concerned about not only the rewards and sanctions of these varied accountability initiatives, but also about their costs and burdens. Teachers may be required to administer several different (and often changing) assessments to their children. They must try to align their instructional practices to multiple sets of state and federal standards.¹⁹
- **Managing fragmentation.** States are responsible for managing these varied systems and approaches (with the exception of Head Start, which operates through a federal-to-local management structure). They must ensure compliance with multiple sets of standards, and they typically work with multiple forms of data on the performance of children and programs.²⁰
- **Making sense of multiple forms of feedback.** Audiences for data on the status of young children and the performance of early childhood programs face the challenge of interpreting multiple forms of feedback based on differing standards and varied assessment tools.

Addressing Discontinuities between Early Education and Elementary Schools

A second structural challenge is the gap between early education and K-grade 3 standards, assessments, data, and related instructional and professional development efforts. As Figure 1 also highlights, the vast majority of young children move from disparate forms of early childhood education into a common and universal public education system. Linkages between programs for preschoolers and the public education system remain limited and sporadic, despite decades of research showing the benefits to children of continuity in preK-grade 3 learning.²¹ Early childhood programs collect valuable data on children's accomplishments, characteristics, and the sequence of their early learning experiences. However available technology that could help transmit these records to kindergarten and primary-grade teachers is not used. The communities where educators can view children's progress across the preK-grade 3 continuum are few and far between.

Similarly, most localities do not coordinate preschool and elementary curricula and professional development.

And yet, we know that early childhood and elementary school educators share common goals and rely on each other's efforts. Elementary schools teachers have a better shot at helping children meet high standards when those children have had the benefit of intensive, enriched, and comprehensive early childhood learning opportunities. And early childhood professionals' hard work is more likely to pay off, in terms of later academic success, when children move from their classrooms into a well-planned continuum of high-quality learning experiences in kindergarten and the primary grades. These natural allies should share responsibility for children's development and learning outcomes, but current policies and organizational structures get in the way, and instead, all too often, they often find themselves at odds: early childhood educators blame schools for failing to build on the progress children have made in their preschool classrooms, while elementary teachers focus on the shortcomings of the children who come to them from early childhood settings.

In summary, the existence of deep structural challenges underscores the need for an effective accountability and improvement system for early childhood education. Fragmentation among diverse programs for preschoolers, as well as fissures between early education and K-3 schooling, affect all of the other challenges discussed below.



Conceptual Challenges: Reconciling Competing Constructs

States also face the challenge of reconciling long-held concepts—and the practices based on them—about assessment of children's learning, and evaluation and management of the programs that serve them.

Long-standing Child Assessment Constructs

Early childhood educators are trained to base their practice on the observed needs and interests of the young child. Building on decades of theory about how young children develop and learn, preparation programs for early educators have emphasized that the child's interests are paramount and must frame what is taught. Unlike elementary and secondary educators, who are often bound to teaching highly prescriptive curricula, early educators are taught to take their cues primarily from children. Early educators are expected to develop or adapt curricula to meet young children's rapidly changing interests and needs.

In order to incorporate this flexibility and individualization into early childhood programs, early educators are trained to be master observers and recorders of children's behavior. In early education, data collection is ongoing and authentic. Teachers routinely observe children's behavior and accomplishments in all domains of development: physical, social, emotional, language, and cognitive. Data are collected in many ways: as informal notes, according to predetermined checklists, or, increasingly, by hand-held computerized devices. The data from these observations are used in a variety of ways: to plan learning activities for an individual child; to plan an entire learning unit; to communicate children's progress to families; or to determine which, if any, children may need additional learning support. It is important to note, however, that typically, these data are not formally reported to funding or accountability agents. In short, using assessments to tailor instruction is a hallmark of quality early education and must be continued.

This approach, often called *instructional assessment*, however, is a far cry from using these data as the criteria for determining program efficacy or for public reporting on the outcomes of early education programs. Indeed, such high stakes uses of data on children's capacities collected for the purpose of instructional improvement are ill regarded. Reconciling the use of authentic assessments for instructional improvement with requests for accountability poses very serious practical, scientific, and conceptual challenges, which are addressed later in this report.

Long-standing Program Evaluation Constructs

The second conceptual challenge is based on beliefs and practices in early childhood program evaluation and management. The performance of early childhood programs has generally been measured by evaluating the quality of services and inputs, including staff credentials. This approach to gauging quality has been so ingrained that, despite the structural challenges described above, all local agencies programs required to meet minimum regulatory standards in order to obtain a license to operate. Other programs have had more stringent program accountability requirements (e.g., the Head Start Performance Standards). Still others have opted to document and improve programs by undertaking voluntary accreditation. But in all of these cases, program quality, practices and inputs, not child outcomes, comprised the required and accepted metric. Inscribed in policy, program quality assessment has a long and rich tradition, in contrast to the accountability approach in K-12 education, which accords less value to program/school inputs and practices and more to student performance outcomes.

While a focus on program quality in early education is the prevailing ethos, this focus does not dismiss the importance of child outcomes. In fact, decades of empirical research studies have documented the impacts of varied early childhood program strategies on children's learning and development. In turn, these programs were studied carefully to

determine the variables that contribute to children's success. The intent was to carefully study small samples to discern the programmatic correlates of children's success; once identified, these items could be incorporated into program regulations for all programs. Highly valuable, these small-scale studies provided useful information to program administrators and staff while informing policymakers about the efficacy of high-quality early childhood interventions.

While the early education field has been well attuned to program evaluation, it has less experience with large-scale, ongoing, standards-based child assessments. It also has less experience in making the connection between the proliferating early learning standards (which are rarely validated) and child outcomes. The field has accepted certain historic correlates of quality (e.g., group size, ratios), but as new programs standards and expectations for children emerge, the connections between these child outcomes and program standards need to be re-established. The challenge, then, and the reason for this report, is to respect early education's heritage while devising mechanisms that will provide policymakers, practitioners, and parents the information they need to make decisions that will improve the quality and outcomes of early education.



Technical Challenges: Assuring Credibility, Inclusiveness, and Equity

As they design and implement accountability and improvement efforts, state officials and planners face two sets of technical hurdles: the first relates to selecting or developing appropriate assessment tools and methods; the second relates to including all children in assessment initiatives.

Selecting or Developing Appropriate Assessment Tools and Methods

High-quality, accurate, credible assessment tools and methods are crucial to accountability and program improvement efforts. Without them, data on the performance of programs and children will be unreliable or distorted, policy decisions and program improvement efforts will be misguided, and public confidence in policies and programs will be eroded. As they select or develop assessment tools and operationalize assessment methods, states face difficult choices and daunting challenges. Their work must take full account of young children's distinctive attributes and research-based understanding of children's early developmental and learning needs:

- **Variability in development and learning.** The rate of early development is swift and extremely variable. From infancy to age five, young children grow physically, emotionally, and socially at a pace more rapid than at any other time in life. As children develop skills and knowledge in an episodic, uneven fashion, an assessment at any single point in time may overestimate or underestimate their true level of development and learning.²²
- **Assessing progress across the developmental spectrum.** The dimensions of early learning and development are broad and diverse. While elementary and secondary school assessments typically limit their focus to cognitive skills and knowledge in academic curriculum areas, assessing the progress of preschool children must encompass the domains of social/emotional development, physical development, approaches to learning, and language development as well as cognitive development. Each of these components contributes to children's readiness for school and should be represented in preschool curricula and learning opportunities. Because young children demonstrate developmental progress in so many domains and in so many ways, it is difficult to design suitable early childhood assessment tools.²³
- **Administering assessments appropriately.** Young children are not capable of responding to group-administered paper-and-pencil tests. They lack the motor skills needed to grip and control a pencil, and they do not yet have the skills needed to understand and respond to written questions. Moreover, they are unfamiliar with large-group testing procedures and settings. Consequently, assessments of young children must be individualized, with an adult assessor or teacher recording data or generating observational ratings. However, the need for one-on-one assessment creates significant feasibility issues for large-scale accountability initiatives. Relatively large numbers of assessors must be trained and supervised. Quality assurance is another major challenge: the consistency, credibility and integrity of child assessment reports must be established and monitored.
- **Moving beyond yesterday's yardsticks.** The child assessment tools that are currently available were developed for purposes that differ from the parameters of today's state accountability and program improvement initiatives. Most were created for two purposes: instructional planning by teachers, or use in research and evaluation studies, in which assessments are typically administered by highly trained teams of external assessors. They were not designed for today's large-scale, ongoing state initiatives, which often call for reporting on the developmental and learning profiles of a state's overall population of young children or those participating in a specific program.

States also must assess the quality of preschool programs, and here, too, selecting or developing suitable tools is demanding. States need tools and methods to ascertain whether local agencies are meeting or exceeding specific state standards related to class size, teacher credentials, and other program measures. They also are seeking tools to gauge what are often called the “dynamic attributes” of program quality: classroom environments, teaching practices, and learning opportunities provided to children. The aim is to identify practices and opportunities that most powerfully and directly foster optimal learning and development.²⁴

States face similar challenges as they design methods for collecting, analyzing, and reporting data. For example, given young children’s rapid growth and diverse starting points at program entry, states must make provision for reporting on their progress over time, rather than simply measuring accomplishments at a program’s end. Sometimes called “value-added,” this approach is a more accurate and fair metric for documenting the contribution of programs to children’s learning and development. Finally, since it is not affordable to conduct comprehensive, high-quality assessments on every child and every classroom, states must determine sampling methods and the timing for cycles of assessments to provide the most accurate gauge of how well programs are working and the most powerful data for informing program improvement efforts.

Including All Children

A second set of technical challenges relates to enabling early childhood accountability and program improvement systems to include all children. This is not yet reality. Currently, not only are individual children omitted from accountability assessments, but large sub-groups of children have been excluded from several of the largest, purportedly state-of-the-art national surveys and research efforts. When accountability systems fail to examine programs’ success in fostering the development and learning of all children, policymakers and the public are denied the full, accurate picture they need, and program improvement efforts are undermined. Moreover, such omissions tend to marginalize precisely the groups who may need the most attention, notably young English Language Learners and young children with disabilities.

Assessing Young English Language Learners. The number of young children in the United States who speak languages other than English is growing significantly. In 2004, 23 percent of births in the United States were to Hispanic mothers, an increase from 15 percent of births in 1990.²⁵ Reflecting the immigration patterns of Hispanic families, these young children come from diverse social, cultural, and linguistic backgrounds. They vary in terms of the age at which they are exposed to English, in their fluency in their native language and English, and in the types of family and school resources they encounter. At the same time, as a group they face many common challenges. Non-English-proficient

children in grades K through 5 are roughly twice as likely to live in poverty as English-proficient children, and only about 50 percent have parents who completed high school.²⁶ On average, their academic achievement scores and high school completion rates are substantially below those of white, English-speaking children.

English Language Learners have much to gain from high-quality early learning opportunities, as they begin to acquire English while continuing the development of their home language skills.²⁷ Becoming fully proficient in a second language is a complex process that takes many years, and children vary enormously in the rate at which they learn a first and second language, based on their characteristics and learning environments. The dynamics of this learning process create challenges for early childhood teachers and programs, and can present technical hurdles for accountability-related assessment efforts.²⁸

As things stand, little systematic information is available about the learning and development of English Language Learners in early education programs because they are often excluded from assessments or assessed only in English. Some assessment tools are available for gauging the progress of young English Language Learners, but they are often inadequate. In some cases, assessment tools offer alternative versions in Spanish or other languages, but items in the two languages do not always have comparable content or levels of difficulty. Norm-referenced tests can be especially problematic. Often, the norming process has not sufficiently represented the actual linguistic diversity of the children who will be assessed. In many cases, children are assessed by staff who do not share or understand the child's language and culture, or who lack training in the process of second language acquisition. Moreover, standards for program quality and accompanying tools rarely include criteria that are specifically geared to fostering optimal learning environments and teaching strategies for young English Language Learners.²⁹

Assessing Young Children with Disabilities. Assessing young children with disabilities also creates challenges for a variety of reasons. First, like their English Language Learner counterparts, children with disabilities comprise a growing proportion of all the children receiving early childhood services. Data from the 2002 school year show that 647,420 children with disabilities were being served in preschool programs through preschool special education under the Individuals with Disabilities Education Act (IDEA).³⁰ This includes children with a wide range of special needs that cause developmental delays, including mental retardation, hearing impairments (including deafness), speech or language impairments, visual impairments (including blindness), serious emotional disturbance, orthopedic impairments, autism, specific learning disabilities, or developmental delays in the areas of physical development, cognitive development, communication development, social or emotional development, or adaptive development.

A second challenge is that there are few assessment tools that capture the full range of children's abilities. A third challenge is the fact that most children with disabilities are usually assessed multiple times, through multiple methodologies, as a function of their need for intervention services (in contrast to English Language Learners, who are often under-assessed). This often begins with initial screenings to identify children in need of supportive services and then moves on to more comprehensive diagnostic assessments aimed at determining a child's specific needs and developmental delays. Additional assessments are conducted to plan and monitor progress in relation to each child's Individualized Education Plan (IEP), which specifies goals and objectives for the child as well as services to assist in meeting those goals. In addition, in 2007, for the first time, all programs receiving preschool IDEA funds are participating in a national child assessment reporting effort through the Office of Special Education, U.S. Department of Education.³¹ This effort requires states to report annually on children's performance in three broad functional areas: improved positive social-emotional skills (including social relationships); acquisition and use of knowledge and skills (including early language/communication and early literacy); and use of appropriate behavior to meet needs. The bottom line: as states plan their overall approach to accountability-related child assessments, they need to take into consideration the variety of assessments already being administered to children with disabilities.



Resource Challenges: Assuring Quality, Capacity, and Fairness

An additional cluster of challenges facing policymakers and planners involves the resources accorded the field.

Investments in some forms of early education programs are proliferating, enabling the expansion of services to young children and their families. On the other hand, policymakers, state leaders, parents and early childhood advocates are

concerned about the quality of these programs and their ability to produce the kinds of gains expected. Such concerns are justified; a 2006 report by the National Institute for Early Education Research noted that only two states' prekindergarten programs met all 10 of the report's quality benchmarks. Not only are quality issues of concern; there is grave concern about the capacity of the field. Generally low staff qualifications coupled with limited investment in professional development delimit personnel capacity. Finally, there is grave concern about which children have access to early education and the adequacy and equity of the distribution of services.

Against this backdrop of serious resource challenges to the field overall, there are three specific connections between issues of investment and development of sound and fair approaches to assessment and accountability: sufficient investment to support improvement, sufficient investment to support accountability, and reporting that reflects resource levels.

Ensuring Sufficient Investment to Support Improvement

States must ensure that programs have sufficient resources to make required improvements. Accountability systems cannot assure quality when states set high performance standards but deny local agencies the resources needed to reach them. When that happens, programs may be punished or sanctioned for failing to meet standards that they cannot possibly reach, despite their best intentions and efforts.

Ensuring Sufficient Investment to Support Accountability Functions

States and the federal government must commit to making the investments needed to support accountability efforts, including the design and implementation of appropriate assessments and the collection, analysis, and dissemination of data. No program should be expected to fund such efforts within its current budget. When this happens, program quality is likely to suffer as administrators divert staff time away from services to children and families in order to comply with assessment and reporting mandates.

Ensuring that Reporting on Program Outcomes Reflects Resource Levels

Finally, reports on programs' performance must take into account the level of resources available to them. Given the great funding disparities that exist among different programs, such inequities must be considered to ensure appropriate judgments and prevent further inequity.

Summary of the Challenges

In summary, states are contending with substantial structural, conceptual, technical, and resource challenges as they develop new approaches to assessing children and programs and as they use assessment data to plan, oversee, and improve overall early childhood service systems, programs, and local provider agencies.

Given all the complexities, it is not surprising that discussions and decisions on accountability policies bump up against sensitive and contentious issues. Local agencies are justly concerned about the criteria and tools used to generate evaluative judgments about how performance data are made available to the public and how state officials use assessment data in decision-making. There also are strong concerns about the risks of unintended negative consequences as local agencies and teachers respond to new accountability mandates.

The recommendations of the Task Force aim to help states respond to these concerns and overcome the challenges highlighted in this review. Despite the difficulties, we are obligated to effectively confront the issues because

- Children miss out if we overlook opportunities to document the positive impacts of high-quality early childhood programs, and thereby to support accelerated investments;
- Children miss out when there are limited resources to support quality programs and well qualified teachers;
- Children miss out if resources are wasted in duplicative assessment and reporting efforts;
- Children miss out if teachers are bewildered by multiple sets of learning goals and objectives;
- Children miss out if teachers do not teach what children need to know and be able to do to succeed in kindergarten;
- Children miss out if valuable assessment data cannot be shared between preschools and elementary schools;
- Children miss out if assessment approaches generate misleading feedback on programs and we cannot learn from and replicate effective strategies;
- Children miss out if poorly functioning local agencies are not identified and strengthened.

Framing a New Approach



Having surveyed today's early education landscape, highlighting the disparate challenges facing accountability and improvement efforts, we turn to the future. *The overall goal of any accountability effort in early education is to improve the quality and equity of services to America's children.* With this goal firmly in mind, in this chapter we present (1) framing beliefs, (2) framing recommendations, and (3) a reframed approach to accountability and improvement that builds on the strengths of the early childhood field while advancing its linkages with elementary education.

Framing Beliefs

Any systematic reform must proceed from a set of beliefs that guides its framers. The Task Force worked hard to articulate four fundamental beliefs and to assure that they were reflected in our recommendations:

- We believe all publicly funded early childhood programs should be held to performance standards, commensurate with levels of public investment, and that performance expectations should be documented and verified.
- We believe that information from such well-designed and carefully implemented performance assessments can and should inform state planning and decision-making and be tied to a comprehensive program improvement strategy.
- We believe, however, that assessments alone will not yield desired results for young children unless they are coupled with sufficient state investments to support high-quality services and ongoing technical assistance and professional development efforts to improve the quality and effectiveness of all programs.

- We believe that one approach to assessment and accountability will not meet the needs of all states nor will it address all the questions being posed by policymakers, parents, and practitioners. To that end, a variety of options must exist. In the end, the approach or combination of approaches used must conform to the questions being addressed, the resources and capacities of the state.

Given these framing beliefs, what should an early childhood accountability and improvement system look like? The Task Force embraced the obligation and opportunity to address this question, delineating how such an effort will work and what leadership efforts are needed to bring it to fruition.

We recognize that the accountability and improvement strategy we shape will not look like what we have, nor will it look the same in all states. It must address not only the framing beliefs listed above, but also the challenges set out in the previous chapter. It must also offer a fresh approach that will serve the public interest, advance the practice of early childhood education, improve program effectiveness, and enhance success for children.

The strategy has taken shape within the context of dramatic changes in American early childhood education, and our framing recommendations therefore take full account of the rapid expansion of early childhood services, multiple efforts to establish an early childhood system, a host of pressing program and workforce quality challenges, and the need to link children's preschool experiences with their early years in elementary school.

Framing Recommendations

We have called our proposals *framing recommendations* to suggest their scope: they identify broad efforts that society must make if we are to respond to the structural, conceptual, technical, and resource challenges discussed in Chapter One. These framing recommendations support our nation's emerging early childhood system, and provide a context for the specific assessment and program improvement

approaches presented in Chapter Three. As such, these recommendations constitute an enduring framework for a dynamic early childhood accountability and improvement system.

1. States should develop a single unified and coherent system of standards, assessments, data and professional development efforts across early childhood categorical programs and funding streams.

We urge states to plan and promote unifying structures and strategies for assessing and strengthening *all* forms of early childhood programs. This recommendation is based upon



a nuanced appreciation of the legitimate reasons for distinctive standards and approaches to assessment and accountability. Many programs were designed to target specific populations of children or families with special needs and characteristics. Given different goals and disparate levels of investment, understandable differences emerged both in what was assessed and how assessments were conducted. Indeed, as long as categorical programs continue, some specific reporting to document program-specific goals and requirements will be necessary.

However, many programs do share common goals: advancing children's physical, social, emotional, and cognitive development; providing nourishing environments for them; and optimizing their chances for success in school. Early childhood programs also share a common research heritage that delineates the attributes of effective programs and best practice pedagogy. Therefore, within the limits of legislative and regulatory mandates, states should strategize to

- Create a single set of early learning standards that would encompass all early childhood programs in the state.
- Simplify the standards, assessments, and reporting requirements mandated by different funding sources for services to preschool-aged children, and make them more consistent.
- Minimize the burdens of multiple assessments for children, teachers and programs.
- Combine resources and supports across program types in the areas of technical assistance, training, professional development and program improvement.

2. States should align high quality and comprehensive standards, curriculum, and assessments as a continuum from prekindergarten through grade 3.

An integrated approach to early childhood services hinges on the creation of high-quality, comprehensive developmental and learning standards and their alignment with both program content and assessments. We must be sure children in preschool programs are assessed on what has been taught and that what is taught aligns with the expectations held by the state. If this is not the case, then children, teachers and programs are being held to unfair expectations.

If this alignment were achieved tomorrow, it would be an immense step forward, but it would not be enough. Future accountability efforts must go further, aligning the standards, curriculum, and assessments for preschool children with those for kindergarteners and primary-grade children. High-quality early childhood programs provide a significant boost in children's school readiness, but alone they cannot guarantee continued success in

school. Young children need a continuous pathway of high-quality teaching and learning experiences, beginning as early as possible and extending through kindergarten and the primary grades and beyond. Early childhood and public school educators share responsibility for building these connections and pathways.

A first step in these efforts is to ensure “vertical” alignment of standards and compatibility of standards, curriculum and assessments between preschools and the first years of elementary school. Vertical alignment strengthens the continuity of expectations, curricula, and assessments between preschool and kindergarten, kindergarten and first grade, first and second grades, and second and third grades. The process of vertical alignment should reflect and respect age-appropriate expectations and children’s natural process of development and learning. It should be approached as an opportunity for reciprocal learning between early childhood and elementary educators and experts. That is, the alignment process should not equate to a “top-down” translation of outcomes from higher grade levels to grade/program levels for younger children. Rather, states should build on the strengths of standards emerging from both early childhood and elementary school communities, to enhance “backward and forward” compatibility of expectations. For instance, early learning guidelines may highlight the significance of diverse domains of child development, while kindergarten/primary grade standards may stress learning goals in academic content areas. Vertical integration can examine inconsistencies and provide for an age-appropriate continuum of standards, curricula, and assessments across the early childhood, kindergarten, and primary-grade years.

3. States should assure that all assessments used are valid and reliable, meet high psychometric standards, and are well suited for their intended purpose. Data analysis and reporting methods should incorporate state-of-the-art methods to accurately and fairly document the performance of children and programs.

Assessing Children’s Growth. As noted earlier, developing and using appropriate assessment tools for children is critical to the effectiveness of any accountability and improvement system. In particular, an assessment system (including the tools selected for use) should meet these criteria:

- Alignment with the full range of a state’s goals for young children’s learning and development.
- Capacity/sensitivity to capture significant changes in children’s learning over time.
- Respect for the unique developmental characteristics of young learners and particularly those from diverse populations.

- Psychometric properties with adequate documentation regarding appropriateness for use with diverse populations of young children.
- Reporting practices that highlight children’s progress over time (reflecting programs’ value-added contributions) as well as their end-of-program status.

Assessing Program Quality. Turning from approaches that assess children to those that measure program quality, we recommend that states target assessment tools and methods to ascertain whether local agencies are meeting or exceeding specific state standards related to class size, teacher credentials, and other criteria. In addition, states should use standardized observational tools to gauge what are often called the “dynamic attributes” of program quality: classroom environments, teaching practices, and learning opportunities provided to children. The aim is to target attention to practices and relationships that most powerfully and directly foster children’s learning and development. In reporting on program quality, we recommend that states highlight performance levels that are typical for state programs as a whole, as well as documenting the variability in performance levels across local agencies, centers and classrooms.

Appendix B provides additional discussion of the strengths and limitations of several approaches to child and program assessment.



4. States should support the full inclusion of all children in accountability and improvement efforts, with appropriate accommodation of assessments to fully document their progress and status.

The children and families entering early care and education programs represent many kinds of diversity. While we are primarily addressing two subgroups (English Language Learners and children with disabilities) we recognize that state leaders and planners must consider multiple dimensions of diversity and tailor their standards, assessments and program improvement efforts accordingly.

What are appropriate accommodations for English Language Learners and children with disabilities? Figure 2 describes the kind of practices that support appropriate reporting on the developmental and academic progress and accomplishments of these groups of children, as well as the context of instructional practices, program quality, and teacher characteristics they experience.

Figure 2

Assessing the Progress of English Language Learners	Assessing the Progress of Children with Disabilities
<ul style="list-style-type: none"> - Document children’s early language experiences in the home. - Evaluate young English Language Learners both in the primary language and the language(s) of instruction. Dual-language assessment will generate data on the trajectories of children’s development in both languages, reflecting the goals of helping children progress towards English acquisition as well as supporting ongoing development in their home language. Assessors should be trained to understand the process and stages of second-language acquisition in order to accurately interpret children’s responses. - Include in assessments of program quality information on instructional practices, including the balance of teaching and interactions in each language, as well as the language abilities of staff members. - Draw attention, in assessment reporting, to children’s progress in their native languages and English and connections with the type, quality and proportion of instruction provided in each language. Guidance should be provided in interpreting child assessment data, given the stages of developing competence in two languages, children’s variable rates of progress, and the influence of language abilities on other aspects of children’s learning and development. 	<ul style="list-style-type: none"> - Identify opportunities to utilize information from assessments carried out by local programs, and in conjunction with data reported in the new federal reporting system. - Provide for adaptations in assessment tools and procedures to allow children with disabilities to participate and to ensure valid measurement of their knowledge and abilities. Adaptations may include augmenting or providing alternative communication systems; providing alternative modes for written language; providing visual support, assistive equipment or devices; functional positioning; sensory supports; and alternative response modes. - Consider involving family members as reporters of child behavior and development. Family involvement expands the validity of any assessment results to home and community applications and allows a window into a child’s cultural and linguistic background. - Use assessment reporting to educate policymakers and the public on the diverse forms of disabilities among young children, as well as ways to strengthen program responsiveness and effectiveness for these children.

5. States should provide adequate resources to enable programs to meet performance standards, and to support accurate, credible and useful assessments and effective program improvement efforts.

Initiatives to hold early childhood programs “accountable” for performance should not proceed unless states have first met their obligations to provide sufficient resources to allow programs to meet standards. For example, if a state standard calls for programs to employ teachers with a specific level of education or credentials, funding rates should support compensation levels sufficient to attract and retain a workforce with this level of training and preparation. The principle of adequate resources highlights the need for accountability expectations to be commensurate with levels of public investment.

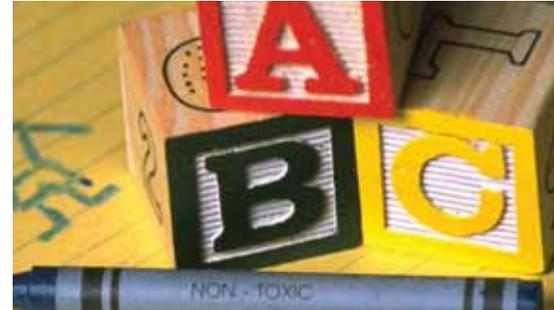
This principle also means that states should provide resources to support the full costs of any and all assessment and reporting mandates, including the training and oversight of

staff members or consultants to administer child or program quality assessments. In addition, data analysis and reporting of assessment results should include contextual information on the levels of investment in the relevant program unit.

A Reframed Approach to Early Childhood Accountability and Improvement

Addressing all five framing recommendations and bringing them to life demands an entirely new approach to early childhood accountability and improvement. No longer can the field be content to allow assessments being prescribed before the purposes and specific uses of the assessment have been spelled out. Nor can data on young children alone be used to determine the merit of a program without full consideration of the nature of the population being served, the resources allocated to the program that serves them, and the quality of the program, including the quality of the teacher-child interactions.

Instead, what is needed is a completely reframed, hand-and-glove approach that combines information on children and on programs to provide a clear picture of the performance of both. Toward that goal, we advocate

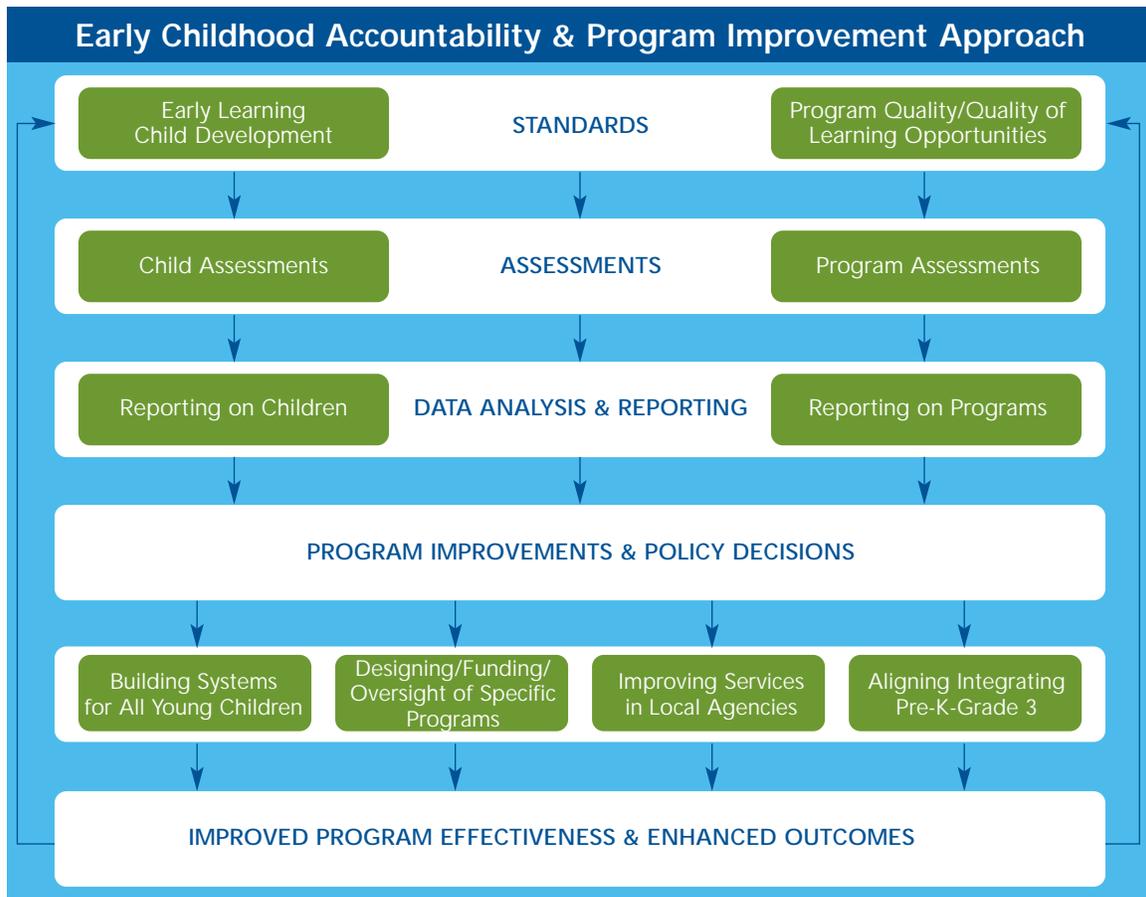


A system of standards-based assessments of (a) children’s development and learning and (b) program quality, designed to inform state policy decisions, investments, and improvement efforts for early education programs for three- and four-year-old children, linked to a continuum of kindergarten through third grade standards, curriculum, assessments, and program improvement efforts.

Figure 3 represents this approach graphically, displaying a standards-based and purpose-driven system, including the following:

- Standards for both children’s learning and development and for program quality.
- Appropriate assessments that are based on these standards.
- Data analysis and reporting methods that assure reliability, validity, and accuracy of the data and safeguard the rights of individual children.
- Specific plans for using the data to guide and motivate program improvement initiatives and policy decisions.
- An ultimate goal of enhancing program effectiveness and positive outcomes for all young children.

Figure 3



In summary, the Task Force believes that a systemic approach to early childhood accountability and improvement must be implemented. This approach builds on the past, rests on a clearly articulated set of beliefs, and incorporates five reframing recommendations. When implemented, it will yield a system that is integrated, inclusive, purpose-driven and adaptive. It will respect children and will honor the diversity of states. It will respect the heritage of early education while responding to the evolving needs of policymakers. It will link early childhood and K-12 education in new ways. Finally, it will provide the basis for the improved delivery of continuous, high-quality services for young children and their families. How can all of this actually take hold? It is to these questions we turn in Chapter Three, where we present a specific operational design, replete with many approaches for states to consider as they develop their early childhood accountability and improvement systems.

Designing an Early Childhood Accountability and Improvement System

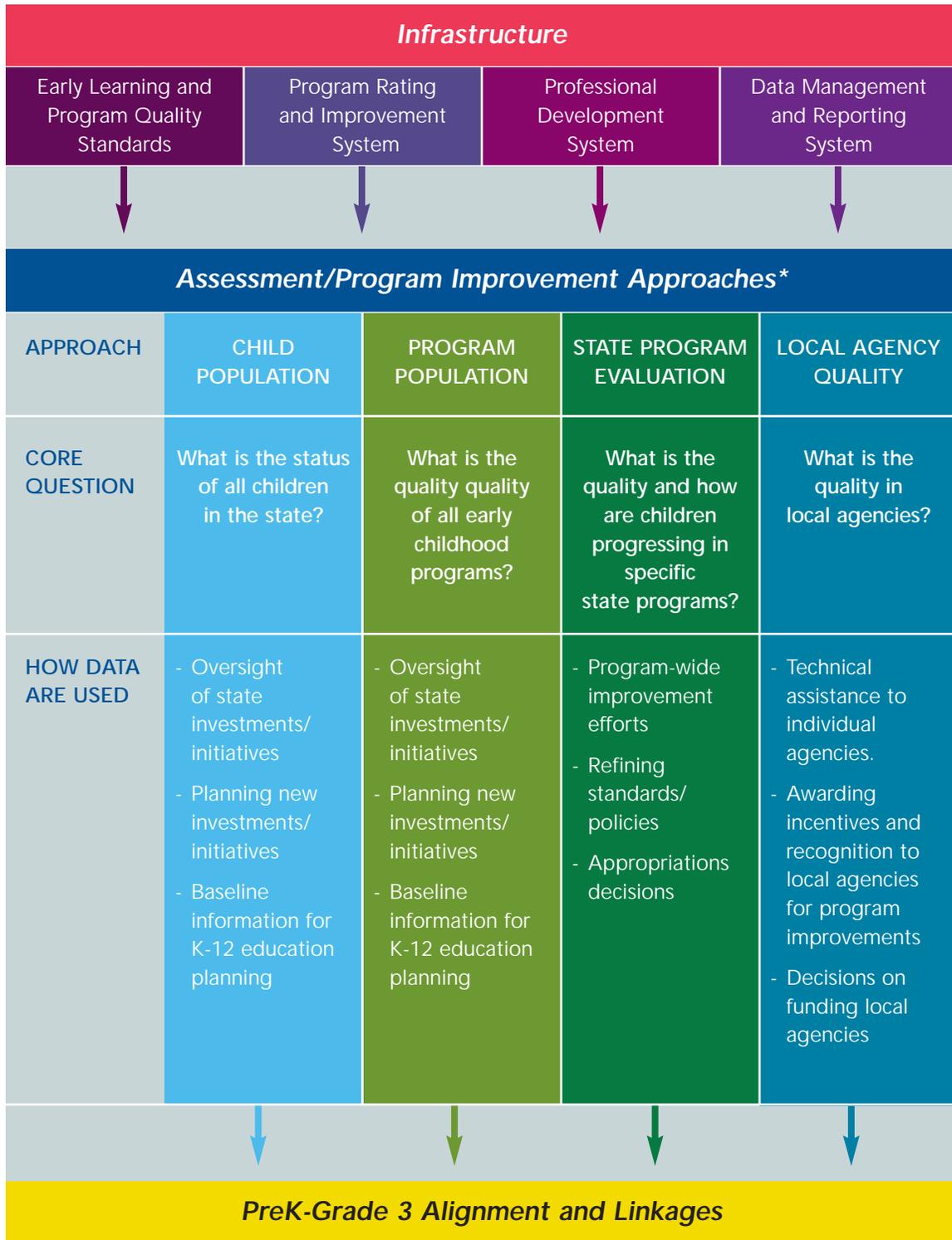


This chapter builds on the recommendations and approach described in Chapter Two, offering an operational design for fair, just, and durable state accountability and improvement efforts. As Figure 4 shows, the overall design comprises three primary building blocks:

- **System Infrastructure.** The design begins with the vital supports needed to ensure high-quality assessments, timely, accurate reporting, and appropriate understanding and use of assessment data. These include
 - Early Learning and Program Quality Standards
 - Program Rating and Improvement System
 - Professional Development System
 - Data Management and Reporting System
- **Assessment/Program Improvement Approaches.** Recognizing the vast diversity among the states, the design provides four approaches for assessing and improving early childhood programs. Each approach responds to distinctive questions and leads to specific program improvement efforts. Each state can select the approach or combination of approaches that most closely meets its programmatic needs and priorities. In addition, some members of the task force wished to consider an approach that would involve reporting on children’s progress in learning and development as well as program quality in individual local agencies. Highly controversial, this approach is discussed in Chapter Four, with diverse positions and safeguards noted.
- **Steps toward a Coherent PreK-Grade 3 Accountability System.** Finally, the design proposes measures aimed at increasing the continuity of standards and flow of data across prekindergarten to grade 3, to enable local efforts to track the progress of children and the quality of learning opportunities across these years.

Figure 4

State Early Childhood Accountability and Improvement System Design



* The Task Force also discussed but failed to reach consensus on an additional approach to assess program quality and children’s learning and development in local agencies. The varied views of Task Force members on this approach and the controversial issue of using child assessment data for local agency accountability are summarized in Chapter Four.

Building the Infrastructure

The Task Force recognizes that to function effectively, all of assessment and program improvement approaches requires the same basic infrastructure. A sturdy infrastructure does more than just support assessment and improvement activities. It unifies supports for all forms of early care and education services, and in the process helps to build an overall early education system, rather than perpetuating a separate support structure for each source of state and federal funding. Each key element of the infrastructure is discussed below.

Early Learning and Program Quality Standards

Any accountability effort must begin with the development of standards that define desired outcomes for program participants as well as criteria for the quality of program services. All states have invested in developing early learning guidelines for young children as well as standards for prekindergarten program quality, child care licensing and, in a growing number of states, Quality Rating Systems. Taken together, these standards reflect the state's values and vision with respect to its young children and the early education services it supports and oversees. Standards serve many functions: they anchor and provide the rationale for determining adequate funding rates; they guide professional development and program improvement initiatives; and they serve as a framework for child and program assessments and as criteria for interpreting assessment information.

But having standards is not enough. Once early learning standards are developed and validated, states must take five important steps toward an integrated accountability and improvement system:

1. **Alignment with curriculum.** The content of what children are taught—curriculum—should align with early learning standards.
2. **Alignment with assessments.** Any and all assessments of children's progress should be linked with early learning standards.
3. **Alignment with K-3 standards, curricula, and assessments.** Sustaining the progress children make in early education programs requires vertical integration—linking early learning standards, curricula, and assessments from age to age and grade to grade.
4. **Alignment across program structures and funding streams.** Early learning and program standards should be aligned across state and federal program structures and funding streams. Toward this end, states should map and cross-walk current standards, to highlight (a) shared goals and priorities, (b) instances where different categorical programs have unique approaches to standards, and (c) criteria in which different programs have more or less stringent or prescriptive requirements.

5. **Alignment of early learning and program standards.** Learning standards that specify what children should know and be able to do must be linked with the program quality standards, so that programs provide the supports needed for children to achieve expected outcomes. Similarly, program quality standards should be validated through studies to establish that they are correlated with positive outcomes and enhanced rates of progress for children.

Standards, be they for the children or the program, are not static. We recommend that states periodically revisit their early learning and program quality standards to ensure that they reflect the changing context of state policy goals as well as emerging research on young children, program characteristics and teaching practices. For example, recent research highlights the importance of specific attributes of classroom environments, teaching practices and child-staff interactions in defining program quality. A second example of the need to revisit standards frameworks is the growing number of children entering early childhood programs who are English Language Learners. Program quality and early learning standards should be reviewed and amended to incorporate provisions related to teaching and serving such children well.

Program Rating and Improvement System

The second Infrastructure element, a Program Rating and Improvement System, has three primary goals:

- Improving the overall quality of early education programs.
- Raising public and consumer awareness about program quality.
- Providing increased funding to encourage and reward programs that provide higher quality early education.

A state Program Rating and Improvement System provides tools and resources that enable local programs to self-assess their quality in relation to relevant quality standards. The state then provides an on-site external review, in which trained evaluators visit the program and use a formalized measure of the program's quality to verify the self-assessment. Assessment data allow states to recognize local agencies based on the level of quality they achieve. The recognition is critical because it is made available to parents and the general public. This motivates programs to engage in the Program Rating and Improvement System and to achieve recognition for providing higher levels of quality services.

We note that these goals and features are similar to Quality Rating Systems currently being implemented through state child care agencies in the majority of states. We chose to use the title "Program Rating and Improvement System" because we are recommending a number of policies and approaches to assessing and improving program

quality which are not currently evident in Quality Rating System efforts. For example, we recommend that participation in Program Rating and Improvement System reviews be mandatory for all publicly funded early education programs. We also recommend that states develop linkages between Program Rating and Improvement System quality criteria and the program standards for all state and federal early education programs. In this approach, systems of standards that are now disparate would be connected through a single comparable set of quality rating levels. Parents, providers and public officials would benefit from the resulting single system of ratings. Moreover, we recommend efforts to simplify and consolidate separate systems of program monitoring and licensing reviews to eliminate duplicative assessments. States could develop reciprocity agreements between monitoring efforts to reduce the incidence of multiple reviews of the same aspect of program quality, or conduct a consolidated review of agencies receiving funding from multiple state and federal programs.

We also recommend that states examine their approach to the timing and targeting of on-site Program Rating and Improvement System reviews. For example, one approach could include unannounced on-site external reviews once every three years for all providers. At the time of the verification visit, data collectors would randomly sample and observe 15 percent of the classrooms in each center, stratifying the sample according to the ages of the children served. Child and employee files would be checked for accuracy. If the findings did not comport with those in the center's Program Rating and Improvement System rating, the center would need to re-certify itself with the Program Rating and Improvement System.

A second approach rests on the conviction that more attention should be accorded to new and "vulnerable" centers. In this approach, the state would determine, on an annual basis, which centers need verification visits and which do not. New centers or those of questionable quality would be given priority for a verification visit. Other programs would receive on-site reviews at least once every five years. This strategy concentrates resources on those most in need of support and those most likely to need or experience change.

Data from Program Rating and Improvement System assessments may be used in many ways. For example:

- **Local program managers** can use the data to identify priorities for program improvement as well as to document strengths and achievements.
- **Families** can use the data in choosing among various provider agencies for their young children.
- **School administrators** can become familiar with the quality of agencies that serve families and young children in various school attendance areas.

- **State program managers** can use the data to help them manage, oversee, and strengthen diverse local provider agencies. They can use this information to set priorities for improving individual agencies, or create technical assistance support for clusters of providers with similar profiles. The assessments also allow state managers to identify consistently high-performing local agencies; examine their characteristics, practices, and strategies; and use their findings as they provide technical assistance to other programs.

We recommend an improvement approach that begins with the state using assessment data to identify lower-performing local agencies for further investigation and analysis. This leads to a deeper review of the context of the agency, levels of resources, characteristics of its workforce and the children and community it serves. This review leads to a technical assistance and program improvement plan, supported by state resources. Follow-up assessments then document whether recommended changes are implemented.

We expect that this compact of shared responsibility, technical assistance support and Program Rating and Improvement System incentives will allow a high proportion of agencies to succeed in making needed improvements. However, states also are responsible for deciding when to discontinue funding if—after receiving multiple opportunities and substantial assistance from the state—an agency is unable to rectify substantial deficiencies.

Such decisions are by definition “high stakes” as they affect the funding of community agencies and the careers of managers and teachers. There also are costs associated with identifying a new agency, transferring funding and starting up services. Therefore, such decisions generate the highest level of concern about the accuracy of assessments and procedural fairness. In such circumstances we recommend the following safeguards:

- The state must provide local agencies with reasonable time, adequate resources, appropriate technical assistance and the opportunity to implement improvements, before taking any steps to withdraw funding.
- In cases where funding is ultimately withdrawn, the state must develop a system for shifting resources to other community-based agencies so there is no reduction or discontinuity in services to children and families.

In summary, a Program Rating and Improvement System provides a central stream of assessment information on the quality of program services, crucial for a number of different audiences and uses. By providing local provider agencies with recognition, incentives and technical assistance, a Program Rating and Improvement System serves as a linchpin of the program improvement infrastructure.

Professional Development System

Professional Development Systems recognize that all adults need ongoing opportunities to improve their skills and competencies as they carry out their roles and responsibilities. The aim of a Professional Development System is to create a consistent, accessible approach to professional development for everyone in a state who works with young children.

Professional Development Systems link informal training with formal education, improve the quality of training content through a training approval process, provide incentives (including compensation) for training, and offer training passports or career registries that chronicle the cumulative training and education individuals receive. Elements of a Professional Development System have been identified by different scholars. One such definition comprises 10 elements:

1. Core knowledge
2. Career path
3. Professional development delivery mechanism
4. Quality approval and assurance system
5. Qualifications and credentials
6. Incentives for professional development
7. Access and outreach
8. Financing
9. Governance
10. Evaluation³²

No matter which definition is embraced, all elements of the Professional Development System must work together and reinforce each other.

Professional Development Systems are essential to the successful implementation of any and all forms of early childhood assessments. Managers and staff responsible for assessment need ongoing opportunities to master early learning and program standards, assessment administration, database management, and other elements of accountability systems. Teachers and local program directors need to understand and make the best possible use of their own instructional assessments of children and the wide range of assessment data, reports and analyses flowing from state and federal accountability efforts. Consumers of assessment data, who must grapple with the implications of child assessments, program assessments and other forms of data, can benefit from supports designed to enhance their “assessment literacy.”

Of course, assessment is not an end unto itself. Assessment guides program improvement—and here too, professional development is critical. Managers and practitioners need to know how to put the data to work—using assessments to plan and

implement improvement efforts and to put into place best practice strategies. As states disseminate accountability reports, they must provide an array of opportunities for adult learning and on-site technical assistance so that front-line educators can use assessment data to address quality gaps and strengthen learning and teaching.

Data Management and Reporting System

The third element in the infrastructure is a comprehensive system for managing and reporting information on children and early childhood programs. Once developed, this repository should contain, in one place, information on

- **Children**, including demographic characteristics, which early childhood programs they participate in, and any available assessment information.
- **Programs**, including funding sources, and results of reviews of program quality.
- **Workforce**, including levels of education, credentials and experience.

To date, early childhood information systems have been developed by state and federal program offices to provide information about and to their individual grantees. This “silo” approach necessitates the duplication of data entry and reporting, and it precludes effective linkages with public school data systems. By contrast, the Data Management and Reporting System would transcend state and federal program structures and apply to all early education programs in the state. In the near term, states will face constraints in moving to a single system. However, states can take initial steps to consolidate and simplify parallel reporting requirements and explore ways of linking separate information systems.³³

Creating a common store of management information also enables states to plan for the future and to invest resources more strategically. Bringing together existing data bases also will help to identify gaps in information and allow for a step-by-step plan to gather additional data. For example, current information systems are not well designed to identify the numbers of children who participate in different combinations of programs, and they often fail to incorporate accurate data on children’s attendance.

A key element in building a Data Management and Reporting System is assigning unique identification numbers to children when they enter early childhood programs, and then linking these numbers to student identification numbers assigned by public school districts. This feature will enable states to connect data from the early childhood years to public education data systems. It will allow children’s progress to be followed over time as they move among programs, schools and communities. Early education programs will be able to get feedback on how their graduates progress when they enter school.

The Data Management and Reporting unit also can take the lead on quality assurance, examining the consistency, integrity, reliability and validity of child and program assessments. Examples of quality assurance efforts include the following:

- Providing clear definitions of data elements and procedures for data entry.
- Designing training programs and materials related to child or program assessment, such as administering assessments or submitting information and reports.
- Designing systems for making random spot-checks of data accuracy.
- Making site visits or studies to audit the validity and reliability of assessments.³⁴

Considering Assessment Approaches

The previous section described the major building blocks of the infrastructure needed to create and sustain any early childhood accountability and improvement system. Now it is time to turn to the four approaches developed by the Task Force and presented in Figure 4. Each approach addresses different questions; each may be used independently or in conjunction with the others. For each approach, we address the following questions:



- What specific questions can it address?
- What data are collected?
- What are recommended designs for data collection?
- How is assessment information used to guide program improvements?
- What are key challenges and cautions?

Child Population Approach: What is the Status of All Children in the State?

The Child Population Approach provides information on all young children in a state, based on studies of representative samples of children, including those who are not enrolled in any preschool effort. States that choose this approach will use data to enhance the awareness of the public and policy leaders about young children's status, demographic characteristics, and patterns of learning progress and attainment. This assessment effort reflects the public interest in the well-being and learning potential of every child, whether or not he or she is enrolled in a publicly supported early childhood program.

What Questions Can the Child Population Approach Address? Depending on specific interests of a state and how information is collected, the Child Population Approach can address such questions as these:

- What is the learning and developmental status of all the state's children at a point in time?
- How are young children progressing over time in their learning and development?
- What are patterns of outcomes/progress for key sub-groups of children?
- Are trends over time moving in the right direction in terms of children's rates of progress and levels of achievement?
- Are there significant demographic changes in our population of young children?

What Child Population Data are Collected? The key data collected in this approach are assessments of a representative sample of all young children in the state, to document their progress on all domains of learning and development addressed in state Early Learning Guidelines. Data also should include the characteristics of the children and their families (e.g., name, birth date, address, health status, any diagnosed disabilities, parental data, family socio-economic status data); preschool enrollment (if applicable), with center identification numbers. States also may make use of other available demographic data to highlight significant trends related to such characteristics as children's health status, parental employment, education and income, family mobility, numbers of recent immigrant families, and families who speak a language other than English. While data are collected from individual children and their families, the data are aggregated for reporting purposes; data on individual children are not reported in this approach.

What are Designs for Collecting Child Population Data? States that choose the Child Population Approach generalize about all children based on data collected about representative samples. Using a sample dramatically reduces the costs of administering assessments, training assessors, data management and analysis. The sampling plan can be designed to provide data on the overall population of young children, or enable reporting on specific subgroups of children. Expanded sampling plans could generate information at the levels of local communities or school districts.

There are three possible designs for sampling: assessing children at the beginning of kindergarten; tracking children's status at ages three, four, and five; and following a sample of children over an extended period. All three approaches can generate reports on trends over time if representative samples of children are tracked annually or on a periodic basis.

Assessing samples of new kindergartners. This design asks teachers to gather observational data and/or administer assessments to representative samples of children at

the beginning of kindergarten. Each year, or at periodic intervals, another representative sample of new kindergartners will be assessed, providing information on trends over time. Although some states are assessing all kindergarteners, we recommend a sampling approach because it offers a more feasible and less costly way to obtain information.

Collecting data at kindergarten entrance is suggested for several reasons. First, because the majority of children attend public kindergartens, it is easier to locate them than when they are in preschool settings. Second, collecting data at kindergarten entry is advantageous in that administering assessments at the beginning of the year reduces the risk that teachers will be pressured to “teach to the test” or inflate test results, because the information is not a reflection of the kindergarten program’s effectiveness. Third, kindergarten entry represents a pivotal transition point for children and represents the intersection of early childhood and public education services, standards and assessments. Reporting such assessments can help elementary school educators focus on children’s strengths and needs at the very beginning of their schooling, and avoid any delay in marshalling needed learning supports. Early childhood educators also benefit from ongoing, global feedback on the performance of young children as they enter kindergarten.

Assessing samples of children at ages three, four, and five. This approach calls for the assessment of representative samples of children at ages three, four, and five. The advantage of this approach is that it allows state officials, educators, and the public at large to understand typical patterns of early childhood development and learning. By comparing data for successive cohorts of children, states can observe changes in patterns of achievement and development for all children. Compared to assessing samples of new kindergartners, this approach increases the costs and challenges of obtaining high-quality assessment data. In particular, it will be challenging for states to locate and complete assessments for three- and four-year old children who are not enrolled in a formal early education program.

Assessing one sample of children over an extended period. This approach resembles a national initiative called the Early Childhood Longitudinal Survey. States that choose this approach identify a representative sample of children, including children who stay at home as well as those who participate in formal or informal early care and education programs and track their progress over an extended period. States can determine the age span they wish to cover, collecting data about the children at different points in time as they move through that span. This approach offers educators, policymakers and the public insight into the trajectories of children’s learning and development over time. However, it requires considerable effort to identify and access children who are not enrolled in any program, and to maintain contact with parents over time to avoid high rates of children dropping out of the sample.

How are Child Population Data Used to Improve Programs? More states are recognizing the human and financial costs of fragmentation and the benefits of a more coherent approach to serving young children and their families. They are reviewing existing services and planning or creating systems of support, bringing together resources and investments across a variety of programs. Toward this end, many states are developing new structures and strategies, including public-private entities, interagency mechanisms, organizational units, and legislative oversight efforts. When states have a broad perspective on early childhood—that is, when they can discern patterns of development and learning for all young children in the state—they are in a good position to support and guide comprehensive early childhood service systems as they emerge.

This kind of information can draw attention to problem areas or patterns of challenges for particular regions or specific types of communities. It can help states track, over time, relationships between patterns of investment and patterns of children’s development and learning. It can be shared with the public through periodic “report cards” to increase awareness of the state’s youngest residents—their characteristics, needs and trends. Finally, this big-picture view of the characteristics and capabilities of all young children can serve as a valuable tool that public education leaders can use to strengthen learning opportunities for children as they move from early childhood settings into kindergarten and the primary grades.

What are Key Challenges and Cautions in the Child Population Approach?

As noted earlier, the Child Population Approach poses the problem of locating and gaining access to preschool-aged children who do not attend any publicly sponsored early childhood program, and who may move in and out of several early care and education programs and settings.

Assessments conducted under Child Population Approach should not be used to make inferences about the quality of children’s preschool experiences. This means that assessments administered to answer questions about how all children are faring should not be used retrospectively to evaluate state programs or local agencies. In particular, using assessments from a single point in time (e.g., at kindergarten entrance) to evaluate local agencies can lead to inaccurate and unfair judgments. Without knowing about prior patterns of children’s skills and knowledge, it is impossible to know whether outcomes result from program participation, or whether they reflect what children already knew or could do before entering the program.

The Child Population Approach addresses some very important issues, but there are several questions it can not answer. It cannot discern why certain conditions exist or why certain changes have taken place. It may happen, for instance, that assessments reflect significant progress in children’s learning at a time when investments in early childhood

programs have increased. But have the investments caused the progress? This approach offers no conclusive proof.

Program Population Approach: What is the Quality of Services in All Early Education Programs in the State?

States can use this approach to report data on all of its licensed or publicly funded early education programs. By documenting the characteristics and quality of all programs, a state can gain insight into how well it is serving its young children. The big plus for this approach: it equips states to survey the entire early childhood landscape, looking at services and supports across different funding streams and types of provider agencies. We recommend reporting this information in conjunction with data on the early childhood workforce and state fiscal investments in services for young children. This provides a big picture of the status of all centers, settings, and staff members that educate and serve young children as well as the overall context of public funding in different types of programs and support efforts. It also can draw attention to trends over time with regard to the state's early childhood enterprise.

What Questions Can the Program Population Approach Address?

This approach addresses several important questions:

- How well is the state's overall early education system working, in terms of delivering quality services?
- Which percentage of local agencies has attained which levels of quality?
- Is the quality of services improving over time, as evidenced by reduced percentages of centers at lower tiers of quality and higher percentages in upper tiers?
- Are programs of high quality accessible on an equitable basis across regions and communities and for different populations of children and families?
- What is the status of the early education workforce (e.g., training, qualifications, compensation, turnover, availability of bilingual teachers) and what are the trends?

What Program Population Data are Collected? Data are collected about program quality, workforce and public investments.

- Program quality data.** States collect data on the characteristics of the center, classroom or setting, its staff, and the population it serves. Program quality assessments should not only document a program's resources and inputs; they also should include direct observation of the quality of learning environments, curriculum implementation, and staff-child interactions.

- **Workforce data.** States would focus data collection on the education, credentials, experience and demographic characteristics of teachers and other staff members and managers. This information draws attention to the people who make early childhood settings into learning environments—those managers and staff members on the front lines of implementing early learning and program quality standards.
- **Public investments.** Tracking data on overall state early childhood investments highlights the multiple sources of funding (state, federal, and private) that support services across different state agencies.

What are Designs for Collecting Program Population Data? Data from the state’s Program Improvement and Rating System will provide the core information on program quality and workforce characteristics. Unlike the Child Population Approach, this approach does not involve representative sampling. The Program Population Approach uses data from the Program Rating and Improvement System on all centers, without exception.

How are Program Population Data Used to Improve Programs? Information on the quality of all programs is especially helpful to policymakers who want a broad understanding of the quality of early childhood services across categorical funding streams and across different types of provider agencies. As public officials, legislators and private-sector representatives work together toward the goal of seamless early childhood systems, they can use this feedback on progress, shortfalls, and priorities to plan improvements and investments. State leaders also can use the Program Population Approach to plan targeted investments and program improvement efforts in cases where equity is in question. Special analyses of sub-sets of data on program quality and workforce characteristics (e.g., by geographic locale, population served, or program type) can highlight disparities in the provision of high-quality learning opportunities across the state. Public education leaders can also use this information to understand the range of learning environments and quality of program experiences provided to children before they enter kindergarten.

As noted earlier, we recommend that program quality data be reported in conjunction with workforce data and data on public investments. When policymakers and planners have made use of data on these key components of quality—programs, workforce, and investments—they have generally focused on them separately. Considering them together, and examining the relationships among them, highlights the state’s responsibility for and progress in providing the resources needed to achieve high-quality programs and a well-qualified workforce. Creating a single consistent mechanism for assessing and reporting program quality will make it easier for these groups to understand the strengths and limitations of current policies and investments.

What are Key Challenges and Cautions in the Program Population

Approach? First, as states report on the quality of all programs, they must ensure the appropriate interpretation and use of the information. The data tell a story, but it is not the whole story. It says that certain conditions exist, but not why they exist. They indicate whether changes have taken place, but not why they have come about.

In addition, when reports show differences or changes in levels of quality, they need to guide their audiences about the significance of those differences or changes. What levels of quality are crucial to protect the health and safety of children and ensure optimal opportunities for learning?

Finally, states need to guard against comparing quality levels for different classes of programs without taking into account differing levels of public investment and other resources available to different groups of local agencies.

State Program Evaluation Approach: What Is the Quality and How Are Children Progressing in Specific State Programs?

The Child Population and Program Population approaches discussed above provide panoramic or population-level data regarding the status of a state's young children and of its early childhood programs and services. State Program Evaluations offers assessment strategies to document the performance and effectiveness of a particular type of state program or investment.

The State Program Evaluation's design offers a number of unique strengths in comparison to other approaches. First, distinguished from the other three approaches in this chapter, it draws together child and program data, enabling state officials—if the evaluation is done well—to understand the relationship among program inputs, practices, quality and the child outcomes. Moreover, because assessments and data analysis within evaluation studies must meet rigorous scientific standards, the results are highly credible. Finally, new evaluation methodologies are being developed and can provide inventive approaches to data collection and analysis.

State Program Evaluation efforts generate data about the overall performance and typical outcomes of a state initiative, categorical program, or funding stream. This approach allows a state to meet its responsibility to report on the performance of specific legislatively mandated programs and investments. It does not, however, answer questions about the performance of specific local agencies, nor differences in quality or outcomes across different types of local agencies.

What Questions Can the State Program Evaluation Approach Address?

This approach gauges how well a given program or initiative is working and answers these questions:

- Is the program being implemented as intended? That is, are local sites adhering to key legislative mandates (e.g., serving targeted groups of eligible children, providing particular forms of services, or meeting standards for teacher certification)?
- What is the typical level of program quality?
- How are the enrolled children progressing, on average, based on the program's standards or expectations for them?
- What program aspects are contributing to the success or lack thereof on the part of the children?
- Under what conditions are different forms of a program effective for different types of young children?
- How well is this particular program working in view of its purposes?

More elaborate designs with more extensive and extended data collection can answer questions such as the cost-effectiveness of state programs, and the extent to which program impacts on children are sustained over time.

What State Program Evaluation Data are Collected? The data collected include information on a sample of centers that represent the universe of the programs being studied, their staff, the population served, as well as data on the children's background, learning, and development. Adhering to the principles enunciated above, the data on the children would include information on all domains of development; data on the centers would be consistent and comprehensive, with information being collected on teachers, staff, and classroom practices. Data on program quality should be gathered in relation to specific mandates and standards of the state program or funding stream, which is usually accomplished by administering standardized observational measures of program quality. Information on levels of financial resources and how funds are utilized also should be incorporated into data collection plans and reports.

What are Designs for Collecting State Program Evaluation Data? The design for collection and analysis of data can take two forms:

Evaluation design. This approach to collecting and analyzing data is modeled on a standard program evaluation design and provides stronger, more authoritative evidence as to whether programs are making a difference. It builds on long-standing program evaluation methodology and draws upon the expertise of organizations that have

experience evaluating a wide range of public and private programs and institutions. It must be carried out with the scientific rigor that such evaluations demand. States can enhance the credibility of reports by contracting with independent organizations (not associated with providing a program) to conduct the evaluation.

The process is straightforward. The first step is identifying the exact state program to be evaluated. That sounds simple, but in reality many local agencies share funding streams or administrative structures, making them harder to isolate and study. Next, the audiences for the evaluation and their key policy questions must be clarified. Are the audiences interested in fiscal data? Do they want to know how children fare over time so they can judge whether program impacts are sustained? Do they want to focus on particular sub-populations? The research questions and outcomes of interest will determine the scope and duration of data collection on children and programs.

At a bare minimum, pre- and post-program data on children and quality data on programs need to be collected. Looking at both program and child data allows consumers of reports to see connections and make judgments about the program. Program evaluation can answer a wide range of other kinds of questions as well. For example, if policymakers want to know about long-term program impacts, they can follow children's progress into kindergarten and beyond, tracking grades, achievement test scores, attendance, retention in grade, and placement in special education. Based on the program goals and scope of services, data may be collected on parent involvement and family outcomes, or on children's health.

Program evaluations may make use of randomly selected comparison or control groups of children who are not enrolled in the state program that is being evaluated. States where a large proportion of children attend the program in question may find this requirement hard to fulfill. States that offer widespread enrollment opportunities but adhere to a uniform age eligibility standard may use regression discontinuity designs.

Analysis of Program Rating and Improvement System quality data and child assessment data. This design is a reasonable choice for states where funding for an independent evaluation may be a challenge. It may also work well for states that are already collecting data on program quality and child assessments on a regular basis for other purposes. If existing data sets include a sufficiently representative sample of local sites and children for a given state program or funding stream, the relevant child and program indicators can be analyzed and reported on a regular cycle.

Key advantages of this approach are its capacity to generate ongoing data cycles at lower cost and provide more timely feedback. However, if a state wants to establish whether participation in the program leads to particular outcomes, this design is less helpful. Reports should acknowledge when analyses rely on comparisons that do not meet the

“gold standard” (i.e., random assignment). Less-than-gold-standard analyses need to also take into consideration entry data on children’s developmental status, their families, and their home languages; child assessment data; program characteristics; and data on program quality. Metrics for combining these varied types of data are not readily available. Indeed, states that want to use integrated data from the three sources (home context, child, and program) may need to undertake methodological prototype work.

How are State Program Evaluation Data Used to Improve Programs? Data reported from State Program Evaluation can inform the public, state policymakers and state program administrators about quality, outcomes and impact of a particular program strategy. These data enable policymakers to make some connections between program inputs and outcomes; they also offer insight into why certain conditions may exist or why changes may or may not have taken place.

These results, if positive, could be used to confirm that a program is effective and justify greater program investments. Data analysis also might point to aspects of the program or specific groups of participants that need more attention. If program results are negative, the data could be used to design targeted program improvement efforts. If successive reports show that the program as a whole is not addressing key problems or improving, the information could be used to redirect state investments or terminate a specific intervention strategy.

Depending on state leaders’ interests and the design of data collection, reports can also address questions that are more fine-grained than those related to overall program effectiveness. For example, states may want to explore which forms of a program are effective, under what conditions, for different types of young children. They may want to compare full-day with part-day programs, or programs serving both three- and four-year-old children with those serving only four-year-olds. They may want information on the effectiveness of different mixes of staffing in terms of credentials and experience. Such findings could lead to policy changes to increase program effectiveness.

What are Key Challenges and Cautions in the State Program Evaluation Approach? Several characteristics of state early childhood initiatives complicate the use of conventional program evaluation methodologies. The Task Force therefore notes four key challenges:

- It is difficult to know with certainty whether a specific program is responsible for particular effects because children may participate in more than one program. They may benefit from more than one funding stream because a number of local agencies avail themselves of multiple funding streams to support a variety of services for children and families.

- It is difficult to know whether children have received the full benefit of a program because different sites may implement the model to different extents, and because children may not have attended regularly or experienced the full course of funded services. Evaluations need to take these factors into account.
- As noted earlier, many states find it difficult to identify appropriate control or comparison groups though new and well-accepted methodological approaches are replacing this method.
- Program start-up presents special challenge that can skew results. States should not begin a full-fledged evaluation of a new program until typical start-up challenges have been overcome and there is reasonable evidence that local agencies are implementing the program in conformity with its purpose and guidelines.

Local Agency Quality Approach: What is the Quality of Services in Local Agencies?

The first three approaches provide state level information: Child Population assessment provides data on the overall status of the state's young children; Program Population assessment provides data on the overall quality of all programs that serve young children in the state; and State Program Evaluation provides information on the overall effectiveness of a particular type of statewide program. In contrast, Local Agency Quality assessment responds to public interest in the quality of services provided by specific local agencies. This form of reporting provides more detailed feedback about how well agencies in particular communities or neighborhoods are meeting quality standards. Local Agency Quality assessments also allow states to report on the variability in program quality across local agencies.

What Questions Can the Local Agency Quality Approach Address? Three questions in particular are addressed through this approach:

- What are levels of quality for individual local agencies in relation to program standards?
- Are individual agencies maintaining or improving their program quality over time?
- Can we identify providers with persistently high or low levels of performance in terms of program quality?

What Local Agency Quality are Collected? In this approach, data are collected on program quality in relation to state program standards as well as standardized observational ratings of the quality of teaching and learning opportunities.

What Are Designs for Collecting Local Agency Quality Data? The recommended approach is to use data from the state’s Program Rating and Improvement System as the basis for reporting and improvement efforts.

How are Local Agency Quality Data Used to Improve Programs? Managers of local program agencies can use feedback on the quality of their services to identify priorities for program improvement, reporting to parents, and sharing with information with local elementary schools and school districts.

Assessments of Local Agency Quality also can inform state program managers as they oversee and improve the quality of services offered by local provider agencies. State officials can use this information to set priorities for improving individual agencies, or create technical assistance support for clusters of providers with similar profiles. As discussed in the presentation of the Program Rating and Improvement System on page 40, the state’s intent and expectation would be to help all programs meet quality standards and move to higher levels of quality over time. However, if a local agency persistently fails to meet standards, in spite of numerous program improvement supports and opportunities, states may wish to withdraw funding and allocate those resources to another community-based agency.

What are Key Challenges and Cautions in the Local Agency Quality Approach? We note four particular challenges and cautions for this approach:

- Managing a provider-level assessment and improvement initiative places considerable demands on state agencies. States must ensure that quality standards are clear and reasonable, and that assessments are accurate and consistent across large numbers of diverse provider agencies.
- States must invest sufficient resources to train and support a cadre of staff or consultants to conduct on-site reviews.
- States must develop technically defensible benchmarks for determining when a provider’s level of performance is exceptional, typical/expected, or inadequate.
- States must create technical assistance approaches to address varied patterns of deficiencies in program quality.

Creating Coherent PreK-Grade 3 Accountability Efforts

To conclude our system design, we offer recommendations aimed at greater continuity in program accountability and improvement from the preschool years through grade 3. Our proposal is intended to better align standards and connect data, and to enable local efforts to track the progress of children and the quality of learning opportunities across the preK-grade 3 years. State leadership is needed to facilitate local partnerships between early childhood and elementary school educators so that they can collaboratively review assessment data and use the findings to strengthen teaching, learning and professional development.



Our call to link and align preK-grade3 accountability efforts builds on the Infrastructure recommendations outlined in this report. Specifically, it echoes the call for state leaders to create a vertically aligned framework of standards for both child learning and program/classroom quality across preK-grade 3, as well as a unified system of child identification numbers that would allow tracking of children’s demographic characteristics, program experiences, and assessment information across those years.

These supports can enable a variety of new partnership efforts between early childhood and elementary school educators, such as the following:

- Develop aligned assessments and methods of reporting to draw attention to trajectories of children’s progress across preschool to grade 3.
- Develop aligned assessments and reporting efforts on the quality of learning environments/learning opportunities for children across preschool to grade 3.
- Develop “vertical” teams of teachers and administrators from each grade/age level to review assessment information and jointly plan to adjust, enrich, and offer new forms of learning experiences and teaching strategies to support children’s continuous progress.
- Develop joint professional development efforts for preK-grade 3 programs and practitioners.
- Develop a stronger sense of shared responsibility for children’s success across the preK-grade 3 continuum.

Building effective preK-grade 3 collaborative teams requires sensitivity to differences in status, credentials, and compensation of early childhood and public education educators.

Yet as the differences are acknowledged, these partnerships can unleash new ways for practitioners to teach and learn from each other. For example, at present, elementary educators have more extended experience in analyzing and using child assessment data, while early childhood educators have more familiarity and expertise in working with quality standards and assessments of program quality, which potentially opens the way for shared responsibility for children's success.



Looking Ahead: Putting the Three Tiers Together to Improve Early Education

In advancing three tiers of work—infrastructure, assessment and program improvement approaches, and prekindergarten-to-primary-school linkages—the Task Force has established a very ambitious agenda. Looking ahead, we understand that full state implementation cannot happen immediately, but we suggest that when efforts in each of the tiers are undertaken, there will be cumulative benefits for children and programs. In this way, all three tiers of the proposed system are complementary.

While effort will need to be expended to bring these elements to reality, we suspect that, over time, implementation of the infrastructure will strengthen the quality and credibility of state assessment and program improvement efforts. As standards for program quality and children's learning are validated, aligned, consolidated and simplified, early childhood educators will be more confident in the basis for state assessment efforts. A single, consistent mechanism for rating program quality will allow states to pinpoint the strengths and weaknesses of different types of programs and local agencies. Investments in program improvement and professional development will allow states to work with local agencies to build on those strengths and remedy those identified weaknesses. A single place to organize and report on all early childhood data and assessments will cut out wasteful duplication and better inform all audiences who care about young children and early education.

As states carry out different combinations of child and program assessments, the results will fill in gaps in our understanding of how children are doing and how well programs are working. Population-level studies of all children and all programs will set the context for interpreting data from more targeted studies of specific programs and local agencies. Rigorous evaluation studies will answer the toughest questions about program impact and cost effectiveness and establish realistic benchmarks for interpreting assessments and guiding improvement efforts. Ongoing reporting on the quality of individual local agencies will inform parents, educators, and citizens about how well children in specific neighborhoods and communities are being educated.

These efforts will also enable, energize and inform more productive partnerships between early childhood and elementary school educators. When states establish a single support system for early childhood standards, assessments, professional development, and data, it will be much easier for elementary school educators to understand and work with their local early childhood colleagues. Enriched, expanded, and higher quality assessments from the early education sector will be an invaluable resource for prekindergarten-grade 3 teamwork.

While the Task Force is quite convinced of the critical need for all states to implement all aspects of our recommendations, we fully expect states to begin in different places, with some spending time and resources on some elements of the infrastructure, others focusing more on a single or multiple approaches, and still others electing to focus on the linkages. In fact, we fully anticipate that states may create additional innovative designs for other infrastructure elements, and that new approaches to standards-based assessments will unfold.

No matter what pathways individual states elect to follow in their accountability and improvement efforts, this movement will significantly contribute to building a high-quality early education system for all children. Inherent and fundamental in our many recommendations, then, is a vision of accountability that is positive and productive and that contributes to the advancement of early education in general.

Differing Viewpoints on Using Child Assessment Data for Local Agency Accountability



In Chapter Three, we presented a comprehensive and adaptable design for an early childhood accountability and improvement system. The design includes three tiers of recommendations: (1) a recommended infrastructure to support accountability and improvement efforts as well as emerging early childhood systems; (2) four distinctive yet complementary approaches for collecting and using assessment data; and (3) recommendations on how to link and integrate early childhood and elementary education standards, assessments, curricula and professional development efforts.

Many members of the task force support these recommendations, and, in particular, the four assessment/program improvement approaches, as a complete response to the challenge of early childhood accountability. Other members argued that the needs of policymakers, program administrators, and the public are not fully addressed by the proposed plan. They suggested that, particularly as publicly funded programs expand, states will want to answer the question of how children are learning and developing in individual local agencies. Accordingly they recommended an additional assessment/program improvement approach for states to consider: to examine both program quality and children's learning and development in local agencies. Consideration of this approach led the Task Force to extensive deliberations around the question of whether child assessment data should be used in local agency accountability efforts.

Amongst knowledgeable individuals, differences in perspective on the future course of early childhood assessment and accountability are understandable. Accordingly, we felt it honest and wise to include a full analysis of these differing views as a prelude to outlining an approach to assessing and improving local agency quality and child outcomes. It is important to note that no consensus was reached on these issues: this chapter is a report of deliberations, rather than a set of recommendations.



Perspectives on Using Child Assessment Data for Local Agency Accountability

There are two major perspectives on the central issue: those who strongly oppose using child assessment data for local agency accountability and those who favor it. Each position will be discussed in turn.

Highly controversial for decades, the use of formal assessments of young children has been strongly criticized for reasons noted in earlier chapters. When states contemplate public reporting of child assessment results for individual local agencies and use of this data in accountability efforts, the level of concern soars dramatically. There are multiple reasons for this heightened level of concern: issues related to the adequacy of the instruments, the challenges of large-scale data collection, and most importantly, the risks of misuse of data. As an example of the scale of these challenges, Task Force members cited the multiple problems in Head Start's heavily criticized and projected to be suspended effort to administer a common child assessment in all local programs, the Head Start National Reporting System.³⁵

More specifically, those who recommend against using local agency-level child assessment data for accountability purposes cite four major reasons:

- First, there are concerns about the adequacy of child assessment tools. Outside of the context of a scientifically designed program evaluations, currently available child assessment tools are largely inadequate. They do not cover all domains of development and cannot capture normal fluctuations in children's development. They do not recognize that young children are unreliable test takers because they often have not been trained to understand the verbal cues or adapt to the situational conditions associated with formal testing.
- Second, there are concerns about implementing large-scale child assessment efforts, including threats to the integrity of data, and challenges in data analysis and reporting. Gathering child assessment information for all local agencies in a state would require a massive effort to train assessors. Establishing and maintaining consistency in assessment procedures and recording data would also be a significant challenge. It would be particularly difficult to ensure the quality and credibility of assessments if local agency staff are the assessors and if the data are used for high stakes decisions. For example, if the data were used to impose sanctions on programs for poor performance, inadvertent coaching of children on assessment items in an effort to show more positive outcomes could result. There

are also challenges associated with data analysis and reporting. There is a concern that child assessment reports for agencies serving small numbers of children may exhibit substantial year-to-year fluctuations in outcomes due to changes in the characteristics of a very few children. This may unduly bias perceptions of how well these agencies are performing.

- Third, there are concerns about the high costs of developing and implementing a local agency-level child assessment system. It is argued that more benefits would come from investing to remedy inequities and deficiencies in program quality, staff training, and compensation, rather than using resources for an expensive and expansive child assessment effort.
- Fourth and most important, there are strong concerns that using child assessment data for high-stakes decisions will lead to serious negative consequences for children. There is the risk that programs, wanting to show good performance, would narrow their curriculum and “teach to the test” in ways that limit and misdirect the quality of children’s learning opportunities. These risks would be particularly acute if the assessment tool addresses only a limited segment of a state’s learning and developmental goals for children. In addition, there are concerns that programs may shift recruitment practices to enroll more children from advantaged backgrounds to increase their chances of showing higher levels of outcomes or better rates of progress. Moreover, children could be harmed if states apply sanctions or reduce funding for agencies based on child assessment data, even if the programs were given sufficient time, resources, or assistance to implement improvements in their staffing, curriculum and learning opportunities.

For all these reasons, advocates of this position strongly opposed any state or federal initiative to collect, report, or use child assessment data at the level of individual local agencies.

On the other hand, Task Force members who favor the use of child assessment data for local agency accountability hold the view that child assessment data can be a significant additional resource in state efforts to understand and improve local agency performance. In their view, questions about how children in specific agencies are progressing are legitimate to ask and appropriate to answer. They see benefits in using child assessment data as an additional factor in targeting resources for more intensive program improvement efforts. They contend that the data from a state-initiated child assessment effort could have positive influences on instructional practices and learning opportunities. This could occur by heightening awareness among teachers and local agency leaders of the needs of groups of children who are not progressing at expected rates, and by marshalling stronger support for more intentional teaching strategies. These Task Force

members are also concerned that accountability strategies that are limited to assessing only the program quality of local agencies may not be sufficiently powerful or sensitive to improve learning outcomes for all children.

Advocates of this perspective acknowledge the substantial technical, logistical, and financial challenges that must be overcome in building an accountability approach that incorporates local agency child assessment information. However, their view is that these obstacles can be addressed through state leadership, investment, and careful planning and management. They share the deep concerns of their colleagues regarding the risks of potential misuse of agency-level child assessment information, but they are persuaded that diligent use of safeguards can minimize such risks. Moreover, they contend that understanding these risks thoroughly is the best means to overcome them.



A Possible Local Agency and Child Outcomes Approach: What is the Quality of and How are Children Progressing in Local Agencies?

Although there are highly divergent views on the viability of using child assessment data for local agency accountability, we felt it important to share how such an approach might be developed if it were desired. To that end, we discuss what such an approach might entail, what might be done to accomplish it and key safeguards that are strongly recommended for any state considering such an approach.

What Questions Can a Local Agency Quality and Child Outcomes Approach Address?

In addition to those questions listed for the Local Agency Quality Approach, the inclusion of child data would also enable the following questions to be addressed:

- What are the patterns of children’s learning progress and end-of-program accomplishments in local agencies? How do these assessment results compare to state early learning and development standards for young children?
- What is the range and variability in children’s performance across local agencies?
- Can we identify local agencies with consistently high or low rates of progress or levels of end-of-program-year performance by children?
- Are there relationships between program quality measures, levels of public investment and local agency-level child assessment data?

Assuming that this approach would operate on an ongoing basis, states could also review trends on each of the above questions.

How Would a Local Agency Quality and Child Outcomes Approach Work?

Data on the quality of local agencies should include structural characteristics of the setting (class size, ratios) and dynamic characteristics that examine how teachers and children interact with one another, the nature of the curriculum and how it is implemented.

Additionally, information should be collected on the teachers (experience, training) and on the families served by the agency.

Information on the children should include data on their prior out-of-home program experiences, primary language, and any identified special needs. To complement these data, states would design a child assessment effort to document the status and progress of children's knowledge, skills, and behaviors. To do so, the state would develop or select a child assessment tool or tools aligned to its early learning guidelines, covering the full range of domains of learning and development. States should consider the approaches of observational, direct, and adaptive direct forms of child assessment as discussed in Appendix B.

Some task force members recommend the use of adaptive direct assessments, based on Item Response Theory (IRT), that describe levels or patterns of children's growth, ability or developmental achievement. Used as individually administered assessments, they can provide information on a child's relative position on a developmental path. Moreover, such assessments can be administered without teaching to the test because different items that assess the same construct can be used with different children. Adaptive assessment strategies that reduce stress on children and assessors could be used; these often involve a two-stage design wherein children take a brief routing assessment that helps determine their general level of performance and routes them on to more detailed and appropriate assessments.

In addition, there are different designs to the overall data collection strategy. States would use data on local agency quality and program characteristics from their Program Rating and Improvement System. In terms of child assessments, it is critical to collect child data at more than one point in time, so child assessment information could be collected at the beginning, end, and, if feasible and affordable, mid-point in the year. Assessing all children is likely to be far too costly to do annually, so representative samples of children should be considered. Moreover, "staggered" strategies could be employed wherein data could be collected in selected geographic regions or on specific populations of children on a regular, rotating basis.

How are Local Agency Quality and Child Outcomes Data Used to Improve Programs?

This approach would follow the strategies for reporting and using data outlined in both the Local Agency Quality Approach *and* in the Program Rating and Improvement System discussed in Chapter Three. However, both program quality assessments and aggregated data on children's performance and progress would be reported to various audiences and used in program improvement efforts. It is important to note that **no information would be reported on individual children**. Reports about groups of children would help policymakers understand relationships between the quality of local agencies and how groups of children are progressing. Contextual information on the characteristics of children, teachers, and program resources will help to enrich these analyses.

Managers of local agencies can use the combination of program quality and child assessment data to identify priorities for program improvement and report to parents and local elementary schools. Families could receive aggregated data on all the overall performance of all the children in the agency as well as on the agency's performance on quality assessments. Local agencies would use information from their instructional assessments to inform parents about their own child's performance.

State program managers can use the data to set priorities for improving individual agencies, or create technical assistance support for clusters of providers with similar profiles in terms of quality and child assessments. These efforts could be informed by studying local agencies that demonstrate high levels of program quality and consistent success in fostering children's learning and development.

A final and controversial use of the data is in decisions to defund persistently low-performing local agencies. This possible use of the data is the source of the strongest concerns expressed regarding unintended negative consequences. States with an interest in exploring this use are advised to carefully consider the safeguards presented in the next section.

What Safeguards are Recommended in a Local Agency Quality and Child Outcomes Approach?

As noted above, implementing this approach is highly controversial, with some suggesting that it should never be considered. For others, implementation represents a possibility to be considered, assuming states would incorporate the following six safeguards to minimize the risks of misuse of assessment information:

1. States should use broad-based child assessments rather than assessments limited to a narrow set of learning goals for children.

2. Child assessment reporting should document children's progress as well as status, so as to not unfairly judge agencies serving children from less advantaged families.
3. If states elect to use direct assessments, the suggestions to employ IRT analyses in developing assessment items and adapted direct assessment formats should increase the accuracy of results and reduce the risks of "teaching to the test."
4. Collection and reporting of contextual information on children, teachers and programs should reduce the possibilities of simplistic or erroneous interpretation of child assessment results. Indeed, this overall approach, pairing quality assessments with child assessment data, would preclude making decisions or framing public perceptions of agencies solely on the basis of child assessment results.
5. Child assessment results should never be the sole criterion for determining rewards or sanctions, or for making funding decisions for local agencies.
6. States should consider developing differentiated performance benchmarks and methods of reporting data that take account of the trajectories of development of different sub-groups of children. This approach would help ensure accurate perceptions and fairness in treatment for local agencies in light of the characteristics of children they are serving.



In addition to these safeguards, some Task Force members who support the approach advise states to follow a carefully planned, step-by-step implementation approach. States would begin with "no stakes" uses of agency-level child assessments and move through low and moderate stakes uses before contemplating high stakes uses. This approach could involve stages such as the following:

Development and Validation. States would develop and validate comprehensive and appropriate assessment tools geared to their child and program standards. Procedures and materials for training staff to administer the assessments would be developed. Methods of analyzing and reporting data, based on technically defensible benchmarks, would be planned. Finally, technical assistance strategies and resources would be procured so the state can work to strengthen agencies with different patterns of child and program quality results.

Pilot Studies. States would implement an initial pilot effort to collect and analyze child and program quality data from a limited number of local agencies with varied characteristics. Careful study of this data would include examining the reliability, validity and credibility of the assessments when implemented in real-world conditions.

Using Assessments for Program Improvement. States would implement the assessment efforts statewide, but assessments would not be reported to the public. State agencies would use the results solely for technical assistance and professional development efforts.

Public Reporting of Assessments. Given concerns about how child assessment information in particular can influence public and parental perceptions, public reporting of results would be delayed until the state and early childhood leaders and practitioners have confidence in the accuracy and credibility of reports.

Using Assessments to Determine Incentives and Funding. Only after incorporating the safeguards outlined above and after considerable experience in using this information for program improvement would states use child assessment information as an additional factor in funding decisions for local agencies.

A Complex Issue, Multiple Views

Our goal in presenting this detailed discussion of the issues and a possible approach regarding the use of child data at the local agency level has been two fold. First, we wanted the reader to understand the tenacity and the complexity of the issues. That the Task Force grappled with this issue over the course of two years of deliberations and came to no consensus is telling. Clearly, controversy remains high. Second, we wanted the reader to understand how some individuals are thinking about inventive ways to collect and use child data as an additional resource along with program quality data in efforts to strengthen local early education agencies.

A Call to Action



The Task Force recommendations, taken together, provide a durable, adaptable framework that any state can use to create an effective accountability and improvement system tailored to its needs and priorities. By creating such a system, a state can move from decisions based on best guesses to policies founded on solid evidence; from improvement efforts based on best intentions to initiatives rooted in empirical analyses; from public opinion efforts based on powers of persuasion to linking parents and citizens to a rich and continually updated storehouse of data on young children and early learning programs. The energies and investments required to accomplish this are substantial, but the anticipated benefits are great. They include the following:

- More relevant and credible data to guide state investments in services for young children.
- More targeted efforts to strengthen equity, informed by richer and more accurate evidence on the extent to which early childhood programs are providing quality services and helping all subgroups of children progress.
- Enhanced credibility for the early childhood profession based on expanded public awareness of how early childhood services contribute to the public welfare.
- Stronger partnerships between early childhood programs and public schools to build a preK-grade 3 continuum of enriched learning opportunities.
- Higher quality learning experiences for children, as states support well resourced, evidence-based program improvement and professional development efforts.
- Improved outcomes for *all* children as accountability and program improvement efforts help states build overall systems of high-quality early education services.

A task force can recommend, but realizing its vision requires leadership, collaboration, and investment across institutions and sectors. To create strong early childhood accountability and improvement systems and practices, people and organizations need to engage in new ways of working across categorical programs, to invest in quality assessment and program improvement efforts, and to advance their own learning about issues of assessment tools, data and analysis.

Accordingly, in this chapter we offer specific recommendations to four groups that are especially well positioned to contribute to accountability systems:

- Governors and state legislators
- State agencies
- Federal government agencies
- Early childhood provider organizations and school districts



What Should Governors and State Legislators Do?

Governors and state legislators have primary responsibility for setting policies for, funding, and overseeing state early childhood programs and are a key audience for the information produced by early childhood accountability systems.

Invest in the infrastructure needed to provide timely, useful, high-quality assessment data.

- Appropriate funding for the costs of enhanced child and program assessment efforts, data management systems and professional development initiatives. As a guideline, we recommend allocating from 2 to 5 percent of all program funding to support state oversight, program management, assessment and program improvement initiatives. Note: The Task Force did not develop detailed cost estimates for more specific elements or approaches in our recommended design.

What Should State Agencies Do?

Leadership from state agencies is crucial because they are responsible for managing public funding for early childhood services, for promulgating and interpreting standards and regulations, and for designing and implementing assessment and program improvement systems.

Develop and implement a strategic plan for the development and implementation of a coherent early childhood accountability and improvement system.

- Conduct an early childhood accountability system audit to map and analyze current policies, standards, assessment mandates, monitoring and licensing efforts, reporting requirements, program evaluation efforts and data system capacity. The audit also should provide a cumulative picture of the costs of current mandates and systems, including infrastructure maintained by the state as well as fiscal and personnel costs incurred by local providers. Use the audit findings to pinpoint opportunities to create a more coherent set of standards, assessments, and other accountability elements across state and federal programs.
- Create a comprehensive long-term plan for an early childhood accountability and improvement system and lead implementation efforts to build an integrated infrastructure, strengthen the quality of current assessment and improvement efforts, and implement new assessment and improvement approaches as appropriate.

Support local preK-grade 3 partnerships

- Explore the current relationships between the state's early childhood assessment and accountability efforts and K-3 education reform and child assessment strategies.
- Examine the alignment of state standards for children's learning from preschool through grade 3.
- Review current practices regarding how child and program assessment data are shared and utilized as children progress from preschool through grade 3.

Work toward a robust, positive and rigorous culture for early childhood accountability efforts.

- Communicate clearly the values of rigor, transparency, and reciprocal responsibility between local providers and state government.
- Articulate and emphasize the aim of ensuring that both policymakers and early educators view the system as fair, appropriate, and relevant to the shared commitment to healthy development and success in school for all children.
- Pursue strategies that support the internalization by educators of the content of standards and the determination to demonstrate ongoing improvement in program effectiveness. These strategies may include collaborative inquiry and planning; an

ongoing public information campaign; and a commitment to refining policies over time based on feedback from pilot programs, formative evaluation, perspectives from the field, and audiences who use the data from various assessment efforts.



What Should Federal Government Agencies Do?

Major funding streams for early childhood services come from the federal government. Federal agencies set policies for these funding streams; they also fund and guide program evaluation and research and development efforts. In particular, the federal government has provided leadership and resources to state efforts aimed at managing early childhood initiatives, strengthening the infrastructure needed for high-quality services, and fostering collaboration across categorical programs. These initiatives include but are not limited to the Head Start Collaboration Office network, the Office of Special Education Program's General Supervision Enhancement Grants, the Maternal and Child Health Bureau's Early Childhood Systems initiative, and the Good Start-Grow Smart Initiative.

Carry out a data harmonization initiative.

- Pursue strategies that would allow information systems for child care, Head Start, and early childhood special education services to mesh with each other and with data generated from state prekindergarten programs.

Invest in research and development that support progress toward coherent state accountability and program improvement systems.

- Support research and development related to child and program assessment tools with an emphasis on tools appropriate for assessing the full range of domains of school readiness; assessment instruments for English Language Learners; and assessment mechanisms suitable for ongoing use in large-scale state systems, such as tools geared to matrix sampling efforts or computer-assisted adaptive measures.
- Establish a national clearinghouse of information on the technical properties, costs, and training requirements of existing child and program assessment tools.
- Establish baseline information for state accountability initiatives by supporting ongoing, longitudinal studies of young children's development and learning as well as the quality of early childhood and kindergarten-grade 3 school services.

What Should Local Early Childhood Agencies Do?

Local early childhood agencies establish learning environments for children and working (and learning) environments for the professionals entrusted with their care. They are the subjects of varied monitoring and quality review efforts, recipients of state technical assistance, and partners in program improvement efforts. Local provider agencies are responsible for generating required information, assessments, and reports for state accountability systems.



Create opportunities for teachers and managers to review, study, and discuss child and program assessments, and to use these data to adapt, refine, and enrich opportunities for children’s learning and development.

- Engage in an ongoing cycle of program self-assessment and continuous improvement, in which feedback from various assessments is combined with analysis of information from ongoing instructional assessments and observations of children and program services.
- Based on this feedback and analysis, identify patterns of strength and weakness and plan appropriate improvement efforts.
- Support collaborative inquiry by providing time as well as expert support and facilitation for collaborative study of assessment data. On a regular basis, convene teaching teams, supervisory staff, specialists in working with young children with disabilities, and family support personnel to analyze ongoing assessments, observations, and other forms of feedback and discuss implications for their work with groups of children and individual children.

Initiate dialogue and schedule information-sharing sessions with local school districts around child assessment, spanning preschool to grade 3, and related data on demographic trends and learning opportunities.

- Work with local school districts to develop mechanisms for reviewing information on the progress of all children, with special attention to the trajectories of learning for key groups such as English Language Learners, young children with disabilities, children from low-income families, African American children, and children of recent immigrant families.
- Identify opportunities to share assessment information on an ongoing basis.

- Contribute to collaborative planning to improve learning opportunities, in particular for children most at risk of school failure, across preschool to grade 3 programs.

A Compelling Need

People and organizations across the nation are already hard at work creating systems of accountable, high-quality early learning opportunities for America's young children. Their efforts are advanced by innovation and research arising from many sources; their determination is inspired by the children. But all too often, their work is hampered by organizational fragmentation, infrastructure gaps, and inadequate tools and methods.

The National Early Childhood Accountability Task Force has sought, in our two years of work together, to help move our nation toward more coherent, effective, sustainable early childhood accountability and improvement systems. Our intent was to provide a set of actionable recommendations and approaches for leaders and planners to consider as they design and implement a system that makes sense for their specific states. Throughout this report, we have offered an honest, self-critical analysis of the challenges to, and prospects for, carrying out our recommendations. Fully recognizing the difficulties inherent in such an undertaking, we have tried to temper ambition with understanding, hope with realism, and vision with fact. We know that readers may not agree with every approach or every recommendation.

We also know that some readers approach accountability efforts in the realm of early childhood education with deep misgivings. While we understand these concerns, members of the Task Force concur that when accountability efforts are of high quality, when they safeguard children, and when they are used in the service of program improvement, they can contribute powerfully to make America's early education fair, just, and equitable, and among the best in the world. Our work is toward that end, and it is dedicated to those who have worked, and continue to work, toward realizing that vision.

Additional Comments

Joan Lombardi

Early childhood policies are at an important juncture. Never before has the field faced such important decisions about how to best serve the needs of young children and their families. As investments grow, pressure increases to assure accountability for public funds.

The recommendations for common infrastructure and the four approaches to accountability and improvement outlined in this report hold much promise for developing a comprehensive system of early childhood development. Of particular importance is the recommended approach to develop a system of review and quality improvement for ALL programs serving young children, regardless of their title or funding stream.

The most challenging issues confronting the task force were questions raised regarding the appropriate use of child outcome data when judging the effectiveness of local programs. Throughout the deliberations I have expressed four overarching concerns:

- Early development is distinct from later development in the degree to which the domains are integrated and the wide variance that occurs in the early years. This variance makes it particularly difficult to use traditional assessment methods or to follow accountability practices as currently defined in education.
- Early childhood development is deeply rooted in a framework that integrates health, parenting support and early education. Service goals include outcomes for parents and community as well as children. Again, early childhood development does not lend itself to the traditional assessment and accountability measures used in elementary education.
- The measurement tools currently available are less than adequate, particularly given the integration of developmental domains and the pace of development. This point is well documented in the National Research Council report, *Eager to Learn*, and was reinforced by Congress in its bipartisan call for an expert panel to look more deeply into the appropriate use of assessment following concern over the National Reporting System developed for Head Start.

- The diversity of funding in the early childhood field and the overall lack of resources to assure quality presents a unique set of challenges when attempting to take a systems-wide approach to the issue of accountability. While there is a need for additional resources at all levels of education, the under-funding of early education, and the inequities in access to quality programs, is much more pronounced during this period of education.

As outlined in Chapter Four of the report, some task force members, including myself, remain concerned with the use of child outcome data for accountability at the local program level. While supporting traditional evaluation methods that have safeguards built in, using child outcome data to make high-stakes decisions about individual programs raises serious concerns. Although this issue has received much attention in education reform debates, the use of this approach is even more sensitive for younger children and among programs that are dramatically under-resourced and without a history of adequate monitoring and support to meet quality standards.

Positive developmental outcomes for children are important goals for early childhood programs. However, policies must be in place to assure the conditions for learning that contribute to such outcomes. Once resources are provided, programs should be held accountable for assuring such conditions are met. Child assessment data should be used to guide curriculum and plan and improve programs, not to make program decisions that can lead to pressure on children, inappropriate teaching practices, and results that are not reliable or valid.

Several issues need further discussion, specifically:

- What do we really mean by “vertical alignment” of assessment from prekindergarten through grade 3? How can we avoid a push down of assessment practices that do not fit the developmental realities of younger children?
- If child outcome data are used in a program review, how much weight does this information receive, compared to information collected on how the program is meeting quality standards? What safeguards are in place to understand how children entered the program and the influence of the home and community environment?
- What are the challenges and potential misuses of developing “student identifiers” and tracking child outcomes from a particular preschool to school?
- How do we avoid the misuse of data collected at kindergarten entry to judge preschool programs retrospectively?

- In the growing number of situations where programs use multiple funding streams, how can we assure that evaluations focused on a single funding stream attribute outcomes appropriately?
- How do we move to a system of common standards, and still meet the more comprehensive service needs of at risk children?
- Where do the quality of infant toddler services fit into our efforts to improve the quality of all early childhood programs?
- How do we assure that investments of time and resources addressing the issue of assessment are not made at the expense of investments in qualified teachers and outreach and supports to families?

In conclusion, the strength of this report can be found in those sections that call for the development of infrastructure to assure quality improvement, particularly around professional development and support, unified standards and data collection. However, serious concerns remain regarding the appropriate uses of child outcome data given the variance in development, the current state of measures, the shortfall in resources and the multiple factors that affect child well-being, particularly in the early years.

Samuel J. Meisels

I am one of the members of the Task Force who is opposed to using child assessment data at the local agency level for accountability purposes. Although I am supportive of many of the main arguments of the report, I believe that test-based accountability in preschool should be rejected in favor of comprehensive program evaluation.

In its simplest terms, accountability asks, “Did the program work?” Evaluations inquire, “How well did the program work and how can we improve it?” The major questions posed by the approaches described in Chapter Three of this report—How are all children doing? How are all programs doing? How is a specific state program doing?—are precisely the questions posed by program evaluations, not accountability systems. These questions cannot be answered effectively or accurately simply by reporting child assessment data. Such data are impaired by numerous shortcomings, including the following, all of which are mentioned in this report:

- Young children are developmentally poor test takers.
- Children’s backgrounds and opportunities to learn are vastly different in early childhood.
- Early development is marked by variability and change.
- Available tests are narrow and unresponsive to individual needs.
- Teaching to the test can have significant negative effects on teaching and learning.

The report does not ignore these and other problems; nevertheless it still presents the “alternative” of using test-based systems of accountability in early childhood in Chapter Four. My contention is that this effort at evenhandedness is inconsistent with what we know about young children and testing. Both our experience with the Head Start National Reporting System, which led Congress to recommend outlawing high stakes testing in Head Start, and No Child Left Behind, which recent research has shown to have limited if not negative effects on student learning, supports the rejection of test-based systems of accountability, particularly in early childhood.

The major argument this report presents is the rationale for program evaluation. This approach enables policymakers to assess the quality and effectiveness of programs without generating the potential stigma or negative consequences of high-stakes testing. But critical to conducting a meaningful program evaluation is the provision of program support in advance of the evaluation.

Common sense tells us that if we want to know how well a program is performing, we need to give the program an opportunity to be successful. Typically, this means providing

it with the human, material, temporal, and procedural resources needed for achieving its stated goals. Here it is worthwhile reminding ourselves of the advice of Donald Campbell, the dean of program evaluation and quasi-experimental research design. He said, “Evaluate no program until it is proud.” Hold off on formal evaluations until there has been an opportunity to create something that works.

Campbell also suggested that judgments and decisions about programs should be based on an accumulation of knowledge. No single study, no single set of test scores, no single piece of evidence should be the basis of decisions to fund or not to fund, to abandon or to replicate a project. Judgments about what works should be founded on a thoughtful appraisal of the many kinds of evidence available. That means relying not only on quantitative but also on qualitative information; not only on evaluations by “objective” outsiders but on the experiences of committed practitioners; not on isolated discoveries but on understanding how consistent the findings are with other knowledge.

Given the state of the early care and education field and the paucity of resources available to it, this advice, which is consistent with the program evaluation approach presented in this report, is of signal importance. Test-based accountability models are inherently narrow and inappropriate for young children. Program evaluations, taking into account process variables about teacher-child relationships, demographic and structural information about who the children are and what the teachers’ backgrounds are, as well as meaningful child outcomes, will give us the comprehensive information we need to determine if a program is functioning as we say it should. Narrow, test-based approaches will not accomplish this.

Commissioned Papers

The following papers were commissioned by the project to inform the work of the Task Force. The papers are available on the project Web site, at www.earlyedaccountability.org.

Papers on Current Accountability Policies and Systems

Assessment in a Continuous Improvement Cycle: New Jersey’s Abbott Preschool Program. Ellen Frede, The College of New Jersey.

Evaluating Early Care and Education Programs: A Review of Research Methods and Findings. Carol Horton, Erickson Institute.

Federal Early Childhood Program Accountability Mandates and Systems. Debi Khasnabis and Thomas Schultz, The Pew Charitable Trusts.

Maryland Model for School Readiness Kindergarten Assessment. Rolf Grafwallner, Maryland Department of Education.

Supporting School Success: Ohio’s Early Learning Accountability System. Sandra Miller and Dawn Denno, Ohio Department of Education.

Standards and Assessment Systems for K-Grade 3 Children. Kristi Kauerz, Teachers College, Columbia University.

State Early Learning Standards and Assessments. Catherine Scott-Little, University of North Carolina at Greensboro and Jana Martella, Council of Chief State School Officers (jointly supported through the Council of Chief State School Officers Advancing PreK for All Project).

Papers on Early Childhood Assessment Tools

A Compendium of Measures for the Assessment of Young English Language Learners. Sandra Barrueco, Catholic University and Michael Lopez, National Center for Latino Child and Family Research (jointly funded by First 5 LA).

The Head Start National Reporting System as a Model For Systems Aimed at Assessing and Monitoring the Performance of Preschool Programs. Nick Zill, Westat, Inc.

Assessment Considerations for Young English Language Learners Across Different Levels of Accountability. Linda Espinosa, University of Missouri, Michael Lopez, National Center for Latino Child and Family Research, and Sandra Barrueco, Catholic University (jointly funded by First 5 LA).

Monitoring Children’s Development From Preschool Through Third Grade. Sally Atkins-Burnette, Mathematica Policy Research.

Papers on Uses of Assessment Data

Rush to Assessment: The Role Private Early Childhood Service Providers Play in Assessing Child and Program Outcomes. Roger Neugebauer, *Exchange Magazine*, and Larry Macmillan, early childhood consultant.

Uses of Data on Child Outcomes and Program Processes in Early Childhood Accountability Systems: Assumptions, Challenges and Consequences. John Love, Mathematica Policy Research.

Child and Program Assessment Approaches

This appendix provides additional discussion of the strengths and limitations of several types of assessment tools for young children and early education programs.

Child Assessment Tools

States may consider the properties, strengths and limitations of two basic approaches to child assessment: observational tools and standardized “direct” or “on-demand” tools.

Observational Tools

These tools are used by teachers or other adults to generate ratings or estimates of children’s knowledge, skills, or abilities based on their performance, behavior, or work in the classroom or other settings. Such tools are commonly used by teachers to collect and systematize ongoing information on the progress of individual children, for use in planning instruction and communicating with parents. In many cases observational tools were developed to accompany a specific curriculum; in other cases they are designed for general use.

Generally, this approach is associated with formats that are criterion-referenced, allowing comparison of children’s performance against criteria for what children their age should know. In many instances, developers have aligned the content of their assessments to state or federal early learning standards. Some observational assessments allow data to be recorded electronically, including via hand-held devices, and can quickly generate a wide variety of reporting formats and analyses.

Standardized “Direct” or “On-Demand” Instruments

These tools involve an assessor presenting a common set of questions or tasks for individual children and recording their responses. A modified version of the direct/on-demand approach, called the adapted direct approach, uses a two-stage method to tailor the level of difficulty of assessment questions to accommodate children’s level of knowledge of competence, based on their responses to an initial set of items. This approach reduces the risk of children becoming frustrated/discouraged by too many difficult items or bored by too many items that are too easy. It also enables more accurate estimation of children’s level of functioning with fewer items and in less time. An additional

advantage is that it reduces the risks of teachers coaching children because different children receive somewhat different sets of questions.

Figure 5 provides an additional overview of some of the strengths and limitations of these two child assessment approaches, in the context of usage in state-managed assessment and evaluation initiatives. Of note is the fact that both forms of assessment require ongoing state investment in training, oversight, and quality assurance mechanisms.

Figure 5

Child Assessment Approaches

<i>Observational Tools</i>	<i>Direct/Adapted Direct Child Assessment Tools</i>
<p style="text-align: center;">Strengths</p> <ul style="list-style-type: none"> - Available tools cover all domains of child development and learning. - Assessors have benefit of multiple opportunities to observe the children over time in a variety of contexts to confirm their ratings. - Assessment process is unobtrusive and does not require removing children from their classroom or interrupting learning activities. - Since teachers use this assessment format for instructional purposes, if data are aggregated for reporting in accountability/evaluation systems, additional costs/burdens are minimal. - Risks of coaching/“teaching to the test” are minimized because assessments are not composed of individual questions. 	<p style="text-align: center;">Strengths</p> <ul style="list-style-type: none"> - Due to the structured nature of questions and method of eliciting direct responses from children, there are lower risks of errors based on the assessor’s judgment. - Use of a common set of questions, similar to standardized assessments used with older children, creates the perception that results are more objective than ratings generated by observers. - The scope, depth and costs of training are typically lower than training for observational tools. - Allows programs to compare the performance of children to norms for nationally representative samples of similar-aged children.
<p style="text-align: center;">Limitations</p> <ul style="list-style-type: none"> - Assessors must be well trained in order to carefully observe and analyze children’s behavior, discourse, work samples and other evidence and generate consistent and accurate ratings. - Bias in teacher ratings can occur if they don’t share the same cultural and linguistic background as the children they are assessing. - The accuracy of teacher ratings can decline over time or suffer from “drift.” - There are risks of teachers inflating ratings to show more rapid progress or higher end-of-program outcomes if assessors perceive that results will influence the reputation of their local agencies or lead to changes in funding levels. 	<p style="text-align: center;">Limitations</p> <ul style="list-style-type: none"> - Assessors must be trained to ensure consistent administration of questions, recording of responses, developing rapport with children and addressing behavioral challenges. - Some children may be distracted or may not perform at their best if they aren’t comfortable with the assessor. - Children must be able to process language well. - Cultural differences among children and a program’s pedagogical practices may influence how children respond to questions or tasks. - Requires removing children from their classroom and assigning and/or compensating staff to administer the assessment. - This type of tool is not appropriate for assessing some important goals, notably social and emotional development. - Reliance on a specific set of questions creates the risks that, if items become known, teachers can coach children on the questions to inflate outcomes.

Program Assessment Tools

States may consider three approaches to standardized rating tools for assessing the quality of early education programs: global assessment of program environment, a focus on teaching strategies and adult-child interactions, and a focus on program administration and management systems.

Global Assessment of Program Environment

The first type of tool is designed to generate a global assessment of a program environment. They document properties associated the setting (e.g., availability of age-appropriate materials and a safe physical environment) as well as such characteristics as predictable routines, a balanced set of activities, and adults who are supportive and available to children. Results from these tools are easy to interpret because they use anchored rating scales, and have norming samples against which to compare ratings. These forms of assessment are widely used in major research studies, emerging state Quality Rating Systems and by many provider agencies for self-evaluation purposes; consequently, they are increasingly familiar to policymakers and the public. There are widespread support systems in place to train assessors and to assist local programs in moving from lower to higher rating levels. One example of this assessment approach is the Early Childhood Environmental Rating Scale-Revised.³⁶

Focus on Teaching Strategies and Adult-Child Interactions

A second approach to assessing program quality focuses more closely on specific aspects of teaching strategies and the quality of adult-child interactions. These tools examine how teachers provide learning opportunities in specific areas of curriculum content, their classroom management practices, and their emotional support for children. Results from these assessments can support teachers' professional growth because they focus on tangible behaviors and teaching strategies. This approach can be used to gauge the extent to which technical assistance or professional development efforts lead to actual improvement in classroom interactions and teaching strategies. Some research shows these process features of classrooms are stronger predictors of child outcomes, when compared to structural features examined in more global ratings.³⁷ Two examples of this assessment approach are the Early Language and Literacy Classroom Observation (ELLCO) and the Classroom Assessment Scoring System.³⁸

Focus on Program Administration and Management Systems

A third approach to assessing program quality focuses on the leadership and management of local provider agencies, highlighting components such as human resource management, fiscal management, program planning and evaluation. This assessment approach highlights the importance of management systems as the linchpin to implementing program quality

standards, identifying and rectifying shortfalls in program performance and creating the working conditions (including ongoing professional development) to enable teachers to foster children's progress in learning and development. One example of this assessment approach is the Program Administration Scale.

Examples of State Assessment and Program Improvement Efforts

The following seven examples illustrate existing state initiatives in standards-based assessments of early education programs and young children and associated program improvement efforts. While they involve a number of the core questions and program improvement efforts described in Chapters Three and Four, they are not presented as specific examples of the Task Force's recommendations.

Maryland

The Maryland State Department of Education (MSDE) is entering its sixth year of reporting on a statewide assessment of all young children.³⁹ More than 2000 kindergarten teachers administer a modified, shortened version of the Work Sampling System (WSS) observational assessment to more than 54,000 children during the month of November. Reporting highlights trends over time in the overall levels of “readiness” of children, as documented by teacher ratings on 30 indicators across seven domains of learning and development. For example, the 2006–7 assessments show that 67 percent of all children scored in the “full readiness” category, (defined on the basis of a state-developed benchmarking system) up from only 49 percent of all children in 2001–2. Assessment data also are reported for various subgroups of children. This information reveals, for example, that the achievement “gap” has narrowed since 2001 for low-income and African American children but has not been reduced for children classified as Limited English Speaking or those in Special Education. Data also are provided for children reported as attending five different types of early care and education programs (child care centers, family child care, Head Start, prekindergarten and non-public nursery schools) as well as home-based/informal care arrangements. Reports also are provided for each local school district, allowing counties and cities to compare results for their children to overall state averages and trends.

Assessment data are reviewed by state legislative committees as part of a bipartisan effort to implement results-based decision-making. For example the previous governor and the General Assembly provided \$1.8 million to expand a state Early Childhood Mental Health Consultation Project to address the social and emotional needs of preschoolers, based, in part, on the kindergarten assessment data. School districts use the information to develop programmatic initiatives and instructional strategies. Kindergarten teachers also are

encouraged to use the WSS assessment on an ongoing basis to monitor the progress of children and in report cards to parents. The MSDE invests close to \$1 million in ongoing funding for professional development to prepare teachers to administer the assessment.

Pennsylvania

Pennsylvania's Keystone STARS is a statewide effort to improve the quality of early learning programs through an integrated system of program quality standards, professional development, assessments, financial incentives and public recognition. Since its inception in 2002, voluntary participation in Keystone STARS has expanded to encompass more than 4,300 local early childhood agencies serving more than 153,000 children. This represents 68 percent of all state-regulated child care centers.

The state's program quality standards encompass staff qualifications and professional development, early learning program, partnerships with family and community and leadership and management. Four levels of quality are recognized through the system. A key element in the Keystone STARS assessment is the use of a standardized rating tool, the ECERS-R or equivalent tools for family child care and other settings/age groups. Agencies seeking to qualify for a "STAR 2" rating must administer the standardized tool in all classrooms and develop improvement plans to address any subscale scores below a 3.0 (on a scale that runs from 1.0-7.0). External assessors are used with applicants for STAR 3 and 4 ratings. They rate a sample of one-third of an agency's classrooms against criteria that incorporate a minimum overall average rating for the agency as a whole (4.25 for STAR 3 and 5.25 for STAR 4) as well as minimum scores and subscale scales for each sampled classroom.

Pennsylvania invests more than \$46 million to support Keystone STARS, including more than \$22.5 million in funding to participating local agencies. Based on their level of quality attainment, providers receive STAR Merit Awards, incentive funding for education and retention of managers and staff members, and STARS Supports to assist agencies in making improvements and moving to higher levels in the rating system.

The Keystone STARS program incorporates references to Head Start's performance standards and program monitoring effort, as well as recognition of voluntary accreditation systems such as those of the National Association for the Education of Young Children, the National Association of Family Child Care and other organizations. For example, local agencies that attain NAEYC accreditation can qualify for a 4 Star-level rating in Keystone STARS if they complete an abbreviated review on a limited set of criteria. The state is considering how Keystone STARS can benefit public school and Department of Education licensed preschool programs as well. As Keystone STARS evolves, the state also reexamines quality criteria and other elements of the system. For example, a recent enhancement of criteria in the area of early childhood education requires agencies seeking

a STAR 3 or higher rating to utilize a curriculum that is aligned to the state's early learning guidelines.

Michigan

Michigan has invested in a series of program evaluations of its Michigan School Readiness Program (MSRP), currently funded at nearly \$85 million per year and serving more than 21,500 children. Beginning in 1995, a state-initiated longitudinal evaluation study documented the quality of services in a sample of six local programs and assessed 338 participating children and a comparison group of 258 children with similar background characteristics who did not have a preschool program experience. The evaluation team assessed program quality in a sample of classrooms via the High/Scope Program Quality Assessment, a standardized rating tool including a total of ten elements ranging from parent involvement, instructional staff and administration, to curriculum and learning environment. Children were assessed in preschool with the High/Scope Child Observation Record, a broad-based observational tool and, from kindergarten through grade 4 via the School Readiness Rating Scale and the Michigan Educational Assessment Program. The evaluation also reviewed data on grade retention, referrals to special education services and school attendance. The study revealed positive and sustained impacts on children's learning from kindergarten through fourth grade, and lower rates of grade retention. This same data base also was used to study other policy questions, such as an analysis comparing the quality and outcomes of full-day vs. part-day preschool programs.

Michigan also participated in a recently completed evaluation of preschool programs in five states. By assessing an overall sample of more than 5000 young children, the study produced reports on each of the five state programs, as well as documenting patterns of outcomes across states. This study utilized a series of standardized, direct assessment tools to examine vocabulary, early mathematics, print concepts and phonological awareness, and a regression discontinuity design. The study found significant positive impacts of the MSRP on children's academic skills at entrance to kindergarten. While this evaluation did not assess the quality of program services, it provided corroboration of the benefits of Michigan's program with a new sample of children, a more rigorous evaluation design, and a battery of direct assessment tools, in contrast to the use of observational child assessment instruments in the earlier longitudinal evaluation.

State officials credit these evaluation studies with creating evidence on the overall effectiveness and positive impact of the MSRP which helped convince state legislators to maintain funding for the program during an era of fiscal stringency in which virtually all state programs and agencies were subjected to budget cuts. In addition, the initial data from the longitudinal study contributed to validating the state's program quality standards and use of the PQA as the state's assessment tool for program quality.

New Jersey

New Jersey's State Department of Education collects and reports several types of data on the quality of services in local agencies funded through the Abbott Preschool Program, currently serving 43,000 children at a budget of over \$450,000,000. Three standardized tools are used for rating classroom quality:

- The Early Childhood Environmental Rating Scale-Revised provides an overview of the quality of classrooms and has the capacity to compare scores with results from national research projects as well as to track trends over time by using data from past state evaluation studies.
- The Support for Early Literacy Assessment (SELA), a state-developed rating tool focuses on specific instructional practices supporting children's early language and literacy skills.
- The Preschool Classroom Mathematics Inventory (PCMI), a state-developed tool focuses on the quality of materials and instructional practices to support mathematics learning.

A team of early childhood faculty from higher education institutions administer the rating tools in more than 300 classrooms each year, chosen to comprise a random sample of classrooms in each school district that receives Abbott funding. Data are used by the state in planning professional development efforts. Reports show improvement over time in the average ratings on all three tools on a statewide basis. In addition, state staff members review the results for individual school districts and use the information in developing plans for local technical assistance and professional development efforts. The state also requires Master Teachers in each local school district to assess the quality of every classroom. This information is used in planning training for individual teachers as well as in evaluating the contracts for funding private child care providers and Head Start agencies. Contracted classrooms must meet or exceed a minimum score in order to continue to be eligible for funding.

In addition, each local program conducts an annual program self-assessment utilizing a state-developed tool, the Self Assessment Validation System (SAVS), a 45-item rating scale based on state program implementation guidelines. State staff members conduct validation visits to roughly one-third of all school districts annually to verify the documentation and validity of the self-assessment data. This information also is used linked to program improvement and professional development efforts. The department also carries out fiscal reviews for more than 500 private providers who contract for services through local school districts. This effort includes tracking reports of enrollments and expenditures and conducting limited audits of approximately 100 providers per year.

Arkansas

In conjunction with a major expansion of Arkansas's Better Chance for School Success (ABC) prekindergarten program, the state departments of human services and education conduct program quality assessments of all provider agencies. State staff members conduct on-site reviews of each local program to examine their compliance with the state's program quality standards, including criteria such as staff qualifications, licensing of the facility and curriculum implementation. In addition, program quality is assessed by consultants who utilize the Early Childhood Environmental Rating Scale-Revised (ECERS-R) in each ABC-funded provider agency at least once per year. All newly funded classrooms and centers are assessed as are one-third of all previously funded settings. Each classroom must achieve an overall level of 5.0 or better on the ECERS-R 7-point scale. Reviewers discuss their ratings with each teacher as well as the local program coordinator, to provide specific feedback on each indicator and any needed improvements. Providers scoring below the minimum level receive technical assistance and are re-evaluated; any program not achieving an acceptable score after a third round of assistance will lose their funding for that center.

Ohio

Ohio has developed a multifaceted accountability approach for its new Early Learning Initiative (ELI) program, established in 2005 to provide early education for children at risk of school failure and meet the child care needs of working families. Currently the program involves 101 local provider agencies serving up to 12,000 children, funded at \$126 million in 2006. Ohio's accountability/program improvement approach uses a combination of measures of program quality and outcomes, including reporting on a common child assessment tool.

Three forms of data on program quality are used to document the performance of local provider agencies:

- External observers assess the quality of a sample of classrooms in each local agency, using a standardized tool, the Early Language and Literacy Classroom Observation instrument (ELLCO) which focuses on literacy and language teaching practices and learning opportunities.
- Local agencies conduct an ongoing program self-assessment using tool based on the state program guidelines. State staff members meet with local program managers to verify evidence of improvement, review scores and to provide professional development and technical assistance support.
- Local provider agencies report a variety of performance data such as teacher credentials, health screening, transition planning, parent education, curriculum alignment. State staff members verify the accuracy of these reports by on-site visits.

Outcome measures include documentation of children’s health status, examining kindergarten retention and special education placement rates, and an assessment of children’s literacy and language skills using Get It, Got It Go!, a brief screening tool that examines vocabulary, rhyming and alliteration. The state recommends that teachers use this tool frequently to monitor children’s progress. Assessment scores are reported to the state in the fall and the spring. To date, Get It, Got It Go! results are shared with local agencies in discussions with state staff members but local agency results are not reported to the public. The state is developing a reporting system to disaggregate data by type of school district and provider agency, and children’s race, English language proficiency and disability status.

State staff members review data from all of the above assessment and reporting efforts to plan program improvement efforts. Data has been used in developing state-wide professional development courses and shared with higher education institutions for their use in enhancing teacher education programs. Local agencies use the information to develop program improvement plans. Individual teachers can meet with an Early Literacy and Language mentor coach to receive specific assistance in areas where s/he would like to improve.

Future plans include expanding the ELI accountability and assessment efforts to state preschool and early childhood special education programs and research and development work to establish performance benchmarks for the child and program assessment tools. Additional effort will be directed to connect ELI data with the state’s Education Management Information System for elementary and secondary education student data. This system is used for reporting the state’s kindergarten entrance assessment, the Kindergarten Readiness Assessment-Literacy (KRA-L), another screening tool focused on literacy and language measures.

New Mexico

The New Mexico PreK Program, initiated in 2005 and currently funded at \$13.5 million, is implementing a new, state-developed child assessment initiative, based on the state’s Early Learning Outcomes. The assessment is primarily designed for use by teachers, but it will also generate data for state oversight and program improvement efforts.

A PreK Assessment Task Force guided the development of an observational child assessment tool incorporating 33 indicators in seven domains of learning and development. Teachers rate all children on all indicators three times per year. In addition, teachers use a portfolio-based tool to provide more detailed narrative descriptions of children’s performance on a smaller set of 5 to 7 indicators. The state sponsors three-day training sessions to improve teachers’ skills in observing children and to enhance the consistency of ratings on the assessment tool. A second training effort is directed to helping teachers use assessment data in curriculum planning.

The primary intended purposes of the state's assessment tool are for planning instruction, reporting to parents, and sharing with receiving schools when children transition to kindergarten. However, each local program also will report their assessments to the state. The state aggregates the data for reporting to the legislature on the progress of all children over the course of the fall, mid-year and spring ratings. State program managers also use the data to set priorities for state training and technical assistance efforts. The state reports overall scores back to each local program and can examine whether patterns of performance for children in different provider agencies are similar to trends for the state program as a whole. The state also provides feedback to local programs on the quality of their portfolio assessment documentation and consistency of ratings. However, to date, assessment information on specific local agencies is not reported to the media, nor has the state established any benchmarks for expected rates of progress or levels of performance on the child assessment tool.

In addition to reviewing child assessment information, state staff members visit each local program twice each year to provide monitoring and technical assistance on the state program quality standards. The state also convenes quarterly meetings of all local programs to clarify expectations and provide assistance in improving program quality and management.

Notes

- 1 Schweinhart, L.J., Martie, J., Xiang, Z, Barnett, W.S., Belfield, C.R. & Nores, M. (2004). *Lifetime effects: The High/Scope Perry Preschool study through age 40*; Ypsilanti, MI: High Scope Press. Campbell, F.A., Ramey, C., Pungello, E.P., Sparling, J. & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*, 6(1), 42-57. NICHD Early Child Care Research Network. (2005). Early care and children's development in the primary grades: Follow-up results from the NICHD Study of Early Child Care. *American Educational Research Journal*, 42(3), 537-570. Reynolds, A.J. & Temple, J.A. (1998). Extended early childhood intervention and school achievement: Age 13 findings from the Chicago Longitudinal Study. *Child Development*, 69(1), 231-246. Loeb, S., Fuller, B., Kagan, S.L., & Carrol, B. (2004). Child care in poor communities; Early learning effects of type, quality and stability. *Child Development*, 75(1), 47-65. Gormley, W.T. & Gayer, T. (2003). *Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-k program*. Washington, DC: Public Policy Institute, Georgetown University. Retrieved from <http://www.crocus.georgetown.edu/oklahoma.html>. Barnett, W.S. (1998). Long-term outcomes of early childhood programs on cognitive and school outcomes. *The Future of Children* 5 (3), 25-50. Retrieved from http://www.futureofchildren.org/usr_doc/vol5no3ART2.pdf. Shonkoff, J.P. & Phillips, D.A., eds. (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press, Chap. 11
- 2 Lee, V.E. & Burkam. D.T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute. Hart, B. & Risley, T.F. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes Publishing. Loeb, S., Bridges, M., Fuller, B., Rumberger, R., & Bassok, D. (2005). *How much is too much? The influence of preschool centers on children's social and cognitive development* (NBER Working Paper No. 11812). Cambridge, MA: National Bureau of Economic Research. Rouse, C., Brooks-Gunne, J. & McLanahan, S. (2005). *The Future of Children School readiness: Closing racial and ethnic gaps*. 15(1). Retrieved from <http://www.futureofchildren.org>.
- 3 Burchinal, M., Yazejian, N., Clifford, R., Culkin, M, Howes, C., Byler, P., Kagan, S., Rustici, J., Bryant, D., Mocan, H., Morris, J., Peisner-Feingold, E., Phillipsen, L. & Zelazo, J. (1995). *Cost, quality, and child outcomes in child care centers. Public report, 2nd ed.* Denver, Co: University of Colorado at Denver, Department of Economics. Pianta, R. C., Howes, C., Burchinal, M., Bryant, D., Clifford, D., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, 9(3), 144-159. Herzenberg, S.; Price, M. & Bradley, D. (2005). *Losing ground in early childhood education*. Washington, DC: Economic Policy Institute.
- 4 Bruner, C., Wright, M.S., Gebhard, B. & Hibbard, S. (2004). *Building an early learning system: The abcs of planning and governance structures*. Des Moines, IA: SECPTAN. Coffman, J., Wright, M.S. & Bruner, C. (2006). *Beyond parallel play: Emerging state and community planning roles in building early learning systems*. Des Moines, IA: SECPTAN.
- 5 PreK-Now. (2006). *Votes count: Legislative action on pre-k, fiscal year 2007*. Washington, D.C.: Pre-K Now. Pre-K Now. (2007). *Leadership matters: Governors' pre-k proposals fiscal year 2008*. Washington, D.C.: Pre-K Now. Barnett, W.S., Hustedt, J.T., Hawkinson, L.E. & Robin, K.B. (2006). *The state of preschool 2006*. Rutgers, NJ: National Institute for Early Education Research, Rutgers University.
- 6 Krueger, C. (2006). P-16 collaboration in the states. *ECS State Notes*. Denver, CO: Education Commission of the States. Retrieved from www.ecs.org/clearinghouse/69/26/6926.pdf.

- 7 Scott-Little, C., Kagan, S.L. & Frelow, V.S. (2003). *Standards for preschool children's learning and development: Who has them, how were they developed and how are they used?* Greensboro, N.C.: SERVE. Neuman, S., Roskos, K., Vukelich, C. & Clements, D. (2003). *The state of state prekindergarten standards in 2003*. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement. Retrieved from <http://www.ciera.org/library/archive/2003-01/index.htm>. Scott-Little, C., Kagan, S.L. & Frelow, V.S. (2005). *Inside the content: The breadth and depth of early learning standards. Creating the conditions for success with early learning standards*. Greensboro, NC: University of North Carolina, SERVE.
- 8 Scott-Little, C. & Martella, J. (2006, April) *Standards-based education, child assessment and evaluation in early childhood programs: A national survey to document state-level policies and practices*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- 9 Gilliam, W. S. & Zigler, E.F. (2004). *State efforts to evaluate the effects of pre-kindergarten: 1977 to 2003*. Retrieved from <http://www.nieer.org/resources/research/StateEfforts.pdf>. Gilliam, W.S. & Zigler, E.F. (2001). A critical meta-analysis of all impact evaluations of state-funded preschool from 1977 to 1998: Implications for policy, service delivery, and program evaluation. *Early Childhood Research Quarterly* 15 (4), 441-473. Barnett, W.S., Lamy, C. & Jung, W. (2005). *The effects of state pre-kindergarten programs on young children's school readiness in five states*. Rutgers, NJ: National Institute for Early Education Research, Rutgers University. Retrieved from <http://www.nieer.org/resources/research/multistate/fullreport.pdf>. Early, D., Barbarin, O., Bryant, D., Burchinal, M., Chang, F., Clifford, D., Crawford, G., Weaver, W., Howes, C., Ritchie, S., Kraft-Sayre, M., Pianta, R., & Barnett, W.S. (2005). *Pre-kindergarten in eleven states; Ncedl's multi-state study of pre-kindergarten & study of state-wide early education programs (SWEET) (NCEDL Working Paper)*. Chapel Hill, NC: University of North Carolina-Chapel Hill.
- 10 Barnett, W.S., Hustedt., J.T., Hawkinson, L.E. & Robin, K.B. (2006). *The state of preschool 2006*. The National Institute for Early Education Research.
- 11 Rhode Island KIDS COUNT. (2005). *Getting ready: Findings from the national school readiness indicators initiative, a 17 state partnership*. Providence, RI: Rhode Island KIDS COUNT.
- 12 Mitchell, A.W. (2005). *Stair steps to quality: A guide for states and communities developing quality rating systems for early care and education*. Alexandria, VA: United Way Success By 6. Retrieved from http://national.unitedway.org/files/pdf/sb6/StairStepstoQualityGuidebook_FINALforWEB.pdf. National Child Care Information Center. (2006) *Quality rating systems: definition and statewide systems*. Retrieved from <http://nccic.acf.hhs.gov/pubs/qrs-defsystems.html>.
- 13 Khasnabis, D. & Schultz, T. Unpublished paper prepared for the National Early Childhood Accountability Task Force.
- 14 Meisels, S.J. (2006). *Accountability in early childhood; No easy answers. Erikson Institute Occasional Paper #6, 7-16*. Retrieved from <http://www.erikson.edu/research.asp?file=publications8series#2006>. National Association for the Education of Young Children and the National Association of Early Childhood Specialists in State Departments of Education. (2003). *Early childhood curriculum, assessment and program evaluation: Building an effective, accountable system in programs for children birth through age 8 - Joint position statement of the national association for the education of young children and the national association of early childhood specialists in state departments of education*. Washington, D.C.: National Association for the Education of Young Children and the National Association of Early Childhood Specialists in State Departments of Education. Retrieved from http://naeyc.org/resources/position_statements/pscape.pdf.
- 15 A list of the commissioned papers and authors is included in Appendix A.
- 16 Information on state child care licensing standards and reviews from the National Association for Regulatory Administration and the National Child Care Information and Technical Assistance Center. (2006) *The 2005 child care licensing study: Final report*. Retrieved from http://nara.affiniscape.com/associations/4734/files/2005%20Licensing%20Study%20Final%20Report_Web.pdf. Information on QRIS standards and assessments from Mitchell, A.W. (2005). *Stair steps to quality: A guide for states and communities developing quality rating systems for early care and education*. Alexandria, VA: United Way Success By 6. Retrieved from http://national.unitedway.org/files/pdf/sb6/StairStepstoQualityGuidebook_FINALforWEB.pdf. and National Child Care Information Center. (2006). *Quality rating systems: definition and*

- statewide systems. Retrieved from <http://nccic.acf.hhs.gov/pubs/qrs-defsystems.html>.
- Information on state preK program standards and monitoring from Barnett, W.S., Hustedt, J.T., Hawkinson, L.E. & Robin, K.B. (2006). *The state of preschool 2006*. Rutgers, NJ: National Institute for Early Education Research, Rutgers University, 178-181, 198-205. Information on state early learning guidelines, preK and kindergarten assessments from Scott-Little, C. & Martella, J. (2006, April) *Standards-based education, child assessment and evaluation in early childhood programs: A national survey to document state-level policies and practices*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- 17 Many child care programs do not receive funding at levels sufficient to meet quality standards as delineated in state Quality Rating and Improvement Systems.
 - 18 As this report is being completed, committees in the U.S. House of Representatives and Senate have both enacted provisions in amendments to reauthorize the Head Start Act that would suspend the National Reporting System, pending the results of a National Academies study of early childhood assessment tools and appropriate uses of child assessment information.
 - 19 Macmillan, L. & Neugebauer, R. (2006). *Rush to assessment: The role private early childhood service providers play in assessment child and program outcomes*. Unpublished paper prepared for the National Early Childhood Accountability Task Force.
 - 20 Schumacher, R., Irish, K. & Lombardi, J. (2003). *Meeting great expectations: Integrating early education program standards in child care*. Washington, D.C.: Center for Law and Social Policy.
 - 21 Shore, R. (1998). *Ready schools: A report of the goal 1 ready school resource group*. Washington, D.C.: National Education Goals Panel. Pianta, R.C. & Kraft-Sayre, M. (2003) *Successful kindergarten transition: Your guide to connecting children, families & schools*. Baltimore, MD: National Center for Early Development and Learning, Paul Brookes Publishing Co. Bogard, K., & Takanishi, R. (2005). PK-3: An aligned and coordinated approach to education for children 3-8 years old. *Social Policy Report*, 19(3), 3-23. Reynolds, A.J., Ou, S. & Topitzes, J.D. (2004). Paths of effects of early childhood intervention on educational attainment and delinquency: A confirmatory analysis of the Chicago child-parent centers. *Child Development*, 75(5), 1299-1388.
 - 22 Wagner, S.L. (2003). Assessment in the early childhood classroom: Asking the right questions, acting on the answers. *Applied Research in Child Development*, 4 (fall, 2003), 1-16.
 - 23 Kagan, S.L., Moore, E. & Bredekamp, S. (1995). *Reconsidering children's early development and learning: Toward common views and vocabulary*. Washington, D.C.: National Education Goals Panel. Committee on Integrating the Science of Early Childhood Development & Commission on Behavioral and Social Sciences and Education, National Research Council Committee on Early Childhood Pedagogy. (2001). *Eager to learn: Educating our preschools*, B. T. Bowman et al. (Eds). Washington, D.C.: National Academy Press.
 - 24 Pianta, R.C. (2003). Standardized classroom observations from pre-k to third grade: A mechanism for improving quality classroom experiences during the P-3 years. Retrieved from <http://www.fcd-us.org/uplodocs/standardizedclrmobsfrompre-kto3rdfinal.doc>.
 - 25 Hamilton, B.E., Martin, J.S., Ventura, M.A., Sutton, P.D. & Menacker, F. (2005). Births: Preliminary data for 2004, *National Vital Statistics Reports*, 54 (8), 1-18.
 - 26 Capps, R., Fix, M., Ost, J., Reardon-Anderson, J. & Passel, J. (2004). *The health and well-being of young children of immigrants*. Washington, D.C.: The Urban Institute.
 - 27 National Task Force on Early Childhood Education for Hispanics. (2007) *Para nuestros niños: Expanding and improving early education for hispanics*. (Main Report). Tempe, AZ: Arizona State University. Laosa, L. & Ainsworth, P. (2007) *Is public pre-k preparing Hispanic children to succeed in school?* New Brunswick, N.J.: NIEER.
 - 28 Garcia, E., (2005). *Teaching and learning in two languages: Bilingualism and schooling in the United States*. New York, NY: Teacher's College Press. Tabor, P. (1997). *One child, two languages*. Baltimore, MD: Brookes Publishing Co.
 - 29 Espinosa, L. & Lopez, M. (2006). *Assessment considerations for young English language learners across different levels of accountability*. Unpublished paper prepared for The National Early Childhood Accountability Task Force and First 5 LA. Espinosa, L. (2005). Curriculum and assessment considerations for young children from culturally, linguistically and economically diverse backgrounds. *Psychology in the Schools*, 42(8), 1-16.

- 30 U.S. Department of Education. *26th Annual report to Congress on the implementation of the individuals with disabilities act*. Retrieved from <http://www.edgov/about/reports/annual/osep/2004/26th-vol-1-sec-1.doc>.
- 31 Information on the Office of Special Education Program's child assessment reporting effort is available at <http://www.fpg.unc.edu/~ECO/index.cfm>
- 32 Kagan, S.L., Tarrant, K. & Berliner, A. (2005). *Building a professional development system in South Carolina: Review and analysis of other states' experiences*. New York: Columbia University: National Center for Children and Families.
- 33 Connecticut's Early Childhood DataCONNECTIONS Project reviewed all early childhood data systems in six state agencies, three federal offices and two private non-profit organizations. The review summarized current data in the areas of supply and demand for early childhood services, workforce characteristics, program quality, costs, funding and child and family outcomes. The study describes the capacity of current systems to answer key policy questions, as well as identifying shortcomings of existing systems in terms of both quality and content. An action plan for future data efforts is framed in terms of ten recommendations to improve linkages across existing data bases and to guide new data collection efforts. Priorities for new data include workforce data, a quality rating system, a household survey to document child care needs and family preferences, and a child data tracking system to link early childhood systems with the Department of Education's student tracking system. Wilson, S. (2006). *A development plan for early care and education data and research systems*. Farmington, CT: Child Health and Development Institute of Connecticut.
- 34 Publications of the National Center for Educational Accountability's Data Quality Campaign include similar recommendations for developing unique student identifier numbers and a quality assurance effort for state elementary and secondary education data systems. National Center for Educational Accountability. (2006). *Creating a longitudinal data system: Using data to improve student achievement*. Retrieved from http://www.dataqualitycampaign.org/files/Publications_Creating_Longitudinal_Data_Systems.pdf.
- 35 Lewis, A. (2003). Hi ho, hi ho, it's off to tests we go. *Phi Delta Kappan*, 84(7), 483-484. Meisels, S.J. & Atkins-Burnett, S. (2004). The Head Start National Reporting System: A critique. *Young Children*, 59(1), 64-66. Raver, C.C. & Zigler, E. (2004). Another Step Back? Assessing Readiness in Head Start. *Young Children*, 59(1), 58-63. Moore, E. & Yzaguirre, R. (2004). Head Start's National Reporting System fails our children. Here's why. *Education Week*, 9 June. Rothstein, R. (2004) Too young to test. *The American Prospect Online*, Nov 1. GAO, *Head Start: Further development could allow results of new test to be used for decision making*, GAO-05-343. Washington, D.C.: May, 2005. Meisels, S. J. (2006) Accountability in early childhood: No easy answers. *Erikson Institute Occasional Paper #6*, 7-16. Retrieved from <http://www.erikson.edu/research.asp?file=publications8series#2006>. For another view, see Zill, N. (2007) *The Head Start National Reporting System as a model for systems aimed at assessing and monitoring the performance of preschool programs*. Unpublished paper for the National Early Childhood Accountability Task Force.
- 36 Harms, T., Clifford, R.M., & Cryer, D. (1998). *Early childhood environment rating scale* (Rev.ed.). New York: Teachers College Press.
- 37 Early, D. et al, 2005. Sylva, K., Siraj-Blatchford, I., Taggart, B., Samons, P., Melhuish, E., Elliot. K. & Totsika, V. (2006). Capturing quality in early childhood through environmental rating scales. *Early Childhood Research Quarterly*, 21(1), 76-92. LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D. & Barbarin, O. (2007). Observed classroom quality profiles in state-funded prekindergarten programs and associations with teacher, program and classroom characteristics. *Early Childhood Research Quarterly*, 22(1), 3-10.
- 38 La Paro, K., Pianta, R., Hamre, B. & Stuhlman, M. (2002). *CLASS: Pre-K version*. Unpublished measure, University of Virginia.
- 39 Grafwallner, R. (2006). *Maryland model for school readiness (MMSR) kindergarten assessment: A large-scale early childhood assessment project to establish a statewide instructional accountability system*. Unpublished paper for the National Early Childhood Accountability Task Force.

Acknowledgements

This report was jointly written by Thomas Schultz and Sharon Lynn Kagan and edited by Rima Shore. It reflects the contributions of a host of people and organizations.

The Task Force was conceived by Susan Urahn of The Pew Charitable Trusts, and created and funded by The Pew Charitable Trusts, the Foundation for Child Development and the Joyce Foundation. Susan Urahn and Sara Watson of The Pew Charitable Trusts, Fasaha Traylor of the Foundation for Child Development and Jana Fleming of the Joyce Foundation served as ex officio members of the Task Force and contributed within and outside the Task Force meetings to the substance, process and strategy of the project.

Presentations from the following individuals, as well as the commissioned papers cited in Appendix A, enriched and informed the Task Force's deliberations:

- Lou Danielson and Jennifer Tschantz, Office of Special Education Programs
U.S. Department of Education
- David Dickinson, Vanderbilt University
- Kathy Hirsch-Pasik, Temple University
- Marylou Hyson, The National Association for the Education of Young Children
- Jack Jennings, Center for Education Policy
- Michael Lopez, National Center for Latino Child and Family Research
- John Love, Mathematica Policy Research, Inc.
- Anne Mitchell, Early Childhood Policy Research
- Ruby Takanishi, Foundation for Child Development
- Grover Whitehurst, Institute for Education Sciences, U.S. Department of Education
- Martha Zaslow, Child Trends, Inc.

Thomas Schultz, project director at The Pew Charitable Trusts, was responsible for overall management of the project. Debi Khasnabis, research assistant, provided valuable background research and served as recorder for the majority of Task Force meetings. Jennifer Rosenbaum of Pre-K Now, Karen Davison, Head Start Fellow and Jeannette Lam of The Pew Charitable Trusts also served as recorders. Judy Carn, Monique Pettway, Jeannette Lam, Emily Hanson, Nancy Clark and Karen Orth provided logistical support and fiscal management expertise. Lorie Slass, Jeremy Ratner and Jessica Riordan provided assistance in media relations and production of the final report. Ellen Wert proofread the document. 202design was responsible for the design of the report.

Task Force members Eugene Garcia, Maryann Santos de Barona, Mary Beth Bruder and Don Rock contributed valuable issue briefs and memos to inform the report.

