Issues in Designing State Accountability Systems

August 2004

Issues in Designing State Accountability Systems

Gloria Harbin, University of North Carolina at Chapel Hill
Beth Rous, University of Kentucky
Mary McLean, University of Wisconsin at Milwaukee

This document is the result of initial collaborative work between two research projects that were funded by the Office of Special Education Programs (OSEP). One of these projects was a directed research project that examined state preschool accountability standards. The other project was a field-initiated study, designed to identify, measure, and predict outcomes of coordinated early intervention service delivery. In addition, in-depth work by the authors in several states also greatly contributed to understanding the issues presented.

Table of Contents

Johnson Elementary serves as a local Head Start grantee and receives funding from the state pre-kindergarten program. Joey attends Johnson Elementary as a preschool student and participates in the 619 program, Head Start and goes to the before and after school child care program offered at the school. The local school administrator complies with the APR requirements for the 619 program and the state pre-kindergarten program, participates in the Head Start monitoring system, and must comply with the regulatory requirements for child care licensing and monitoring. This fall, Joey was tested for the 619 program, participated in the National Reporting System for Head Start, and the state pre-kindergarten screening and assessment process. This spring, Joey will participate in the same assessment processes as in the fall, but also will be assessed in the child care program to meet new requirements for standards and accountability developed by the state in response to the Good Start Grow Smart Initiative.

P rograms for young children with disabilities are facing a significant challenge. Federal agencies, such as the Office of Special Education Programs (OSEP), in the U.S. Department of Education, are being asked to provide effectiveness data for the programs they fund. Despite an array of challenges, OSEP must provide data on the effectiveness of the services provided under Part C and Section 619 of Part B of the Individuals with Disabilities Education Act (IDEA). State legislatures also are demanding accountability data about the effects of the programs they fund. These data are, in fact, long overdue. As states have initiated efforts to develop accountability systems, they have been hampered by the lack of rigorous models, the fragmentation of accountability initiatives, and the lack of consensus among stakeholders (including researchers) about which outcomes of service delivery should be measured.

At the Conference on Research Innovations in Early Intervention (CRIEI) - February 2004 in San Diego, CA, one of the working sessions was designed to provide a forum for a national discussion among researchers who are assisting states in various aspects of designing state accountability initiatives. As a result of this meeting, the group asked the presenters to transform the content of the presentation into a document that could guide researchers in their work with state and local policy makers. This document is the result of that request and will hopefully facilitate and continue important dialogue in this area.

Purpose

The purpose of this working document is to provide a resource for state policy makers and researchers who are interested in collaboratively designing and implementing rigorous and systematic accountability models that yield accurate data.

Format

This document is divided into four sections: 1) federal goals and indicators; 2) multiple levels of an accountability system (e.g., federal, state and local); 3) identifying outcomes or standards and; 4) measuring outcomes or standards. The narrative in each section provides a background and an explanation of the issues to be addressed by those individuals involved in developing accountability systems. The authors hope that this information will aid accountability system developers to use the checklists at the end of each section. These three checklists contain a series of questions to identify issues and guide decisions that need to be addressed in the development of a state accountability system. The three checklists are interrelated. Decisions made about questions in each section affect decisions in the others.

Federal Goals and Indicators

Historically, federal laws such as IDEA often include the goals of the legislation, thus indicating the desired outcomes intended by Congress. Previously, it was assumed that the programs funded by OSEP were designed to accomplish the legislative goals of IDEA. The Annual Report To Congress provided limited data on whether the goals of the legislation were met - whether children had progressed and whether families were more capable. More recently, Congress has acted to request more systematic data about the effects of the programs on children and their families.

In 1993, Congress enacted the Government Performance and Results Act (GPRA), which mandated each federal agency to identify goals and indicators for all of its programs and to report progress toward those goals to Congress on an annual basis. Thus, GPRA created a government-wide focus on results designed to enhance decision making by Congress. GPRA has generated the need for data concerning aspects of the service system (e.g., number of children served in inclusive settings) as well as child assessment data that can be aggregated for reporting at the state and federal levels. Currently, there are four GPRA indicators for Part C of IDEA (Infant/Toddler Programs) and two indicators for Section 619 of Part B (Preschool Special Education Programs) of IDEA. It is possible that these indicators may be revised in the near future. However, the current, officially adopted indicators are as follows:

Part C GPRA Indicators:

- Increase in the # of states that serve more than 2% of the general population in 0-3 and more than 1% of the general population in 0-1 range
- The percentage of children receiving services in programs designed for their typically developing peers will increase
- The percentage of children participating in Part C that demonstrate improved and sustained functional abilities will increase
- The percentage of families that report that early intervention services have increased their capacity to enhance their child's development will increase

Preschool GPRA Indicators:

- Percentage of preschool children who receive special education in inclusive settings will increase
- Percentage of preschool children with disabilities receiving special education and related services who improve their early language/communication, pre-reading and social-emotional skills will increase
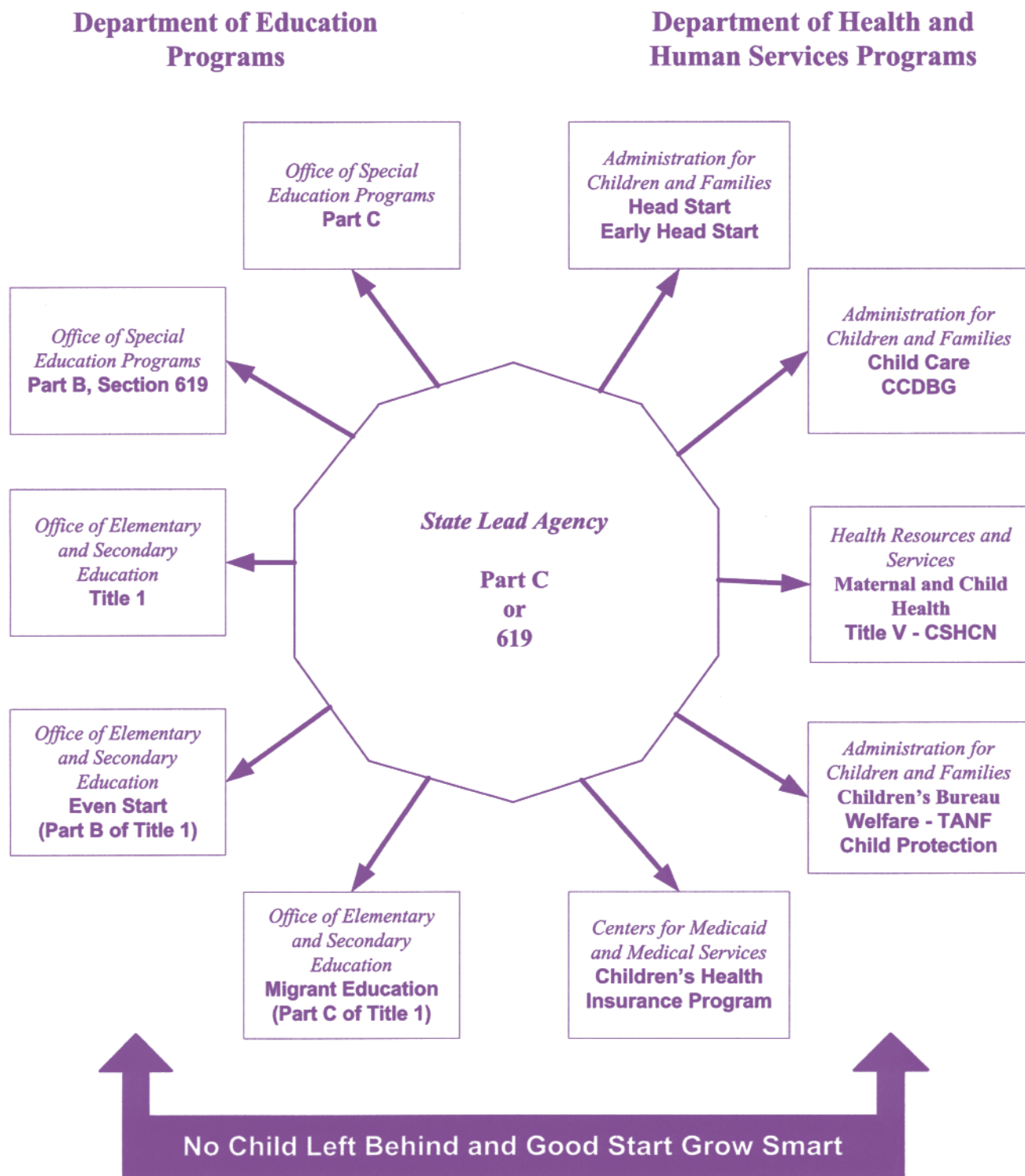
In turn, OSEP is including these indicators and others in the Continuous Improvement Focused Monitoring System (CIFMS) which is used to monitor states on the implementation of IDEA. An even more detailed set of the indicators for state self-assessment is being developed by the federally-funded National Center on Special Education Accountability and Monitoring (NCSEAM) under the direction of Dr. Alan Coulter at the Louisiana State University Health Sciences Center in New Orleans.

In an accountability initiative spear-headed by another governmental agency, the Office of Management and Budget in 2002 developed the Program Assessment Rating Tool (PART) to provide a link between GPRA and the budget development process. The first PART assessments of the Infant/Toddler and Preschool programs of IDEA have been completed at this point in time. Unfortunately, the rating of the assessments was "Results Not Demonstrated" for both programs, adding a sense of urgency to the need for large scale assessment information from the states on the progress of young children with disabilities and their families (www.whitehouse.gov/omb/budget/fy2004/pma/html). In addition, the PART Assessment directed the Department of Education to develop performance goals and measures for both programs. OSEP has recently funded the Early Childhood Outcomes (ECO) Center, directed by Dr. Kathy Hebbeler of SRI International, Menlo Park, California to lead that effort.

Just as there are GPRA Indicators and PART assessments for the Infant and Toddler (PART C) and Preschool Special Education Programs (Section 619 of Part B) which are within IDEA, other federal agencies also are required to delineate GPRA Indicators for success and to be assessed through PART as well. Consequently, Part C (Infant & Toddler) and Section 619 (Preschool Special Education) programs find themselves embedded in a complex context, in which there are multiple federal accountability initiatives going on simultaneously. Other federal programs that have linkages to the Infant-Toddler and Preschool Special Education Programs are immersed in their own efforts to keep pace with federal outcome requirements. For example, Head Start, Children with Special Heath Care Needs (CSHCN), Temporary Aid to Needy Families (TANF), and the Developmental Disabilities Programs also are identifying accountability indicators. Even within the same federal agency (e.g., Department of Education), separate programs are delineating their own GPRA indicators (e.g., Even Start, Title I, Bilingual Education, Programs for children with disabilities).

In addition to GPRA and PART, most federal programs (e.g., CSHCN, TANF) are required to develop and implement monitoring activities, which have their own set of indicators (e.g., similar to the CIFMS developed by OSEP and described earlier). Programs such as Part C and Section 619 of Part B of IDEA also have data reporting requirements that are separate, and sometimes different from, the monitoring requirements. Finally, some federally funded programs must participate in some form of national assessment of their service recipients (e.g., National Education Assessment and Head Start National Assessment). Figure 1 depicts an array of relevant federal programs that provide services to some of the children and families served by the Infant-Toddler and Preschool Special Education Programs. This figure indicates that Part C and Section 619 programs are not the only federal programs involved in federal accountability initiatives. Unfortunately there is currently no mechanism or group that is trying to coordinate these various agencies' accountability efforts. Consequently, several of the agencies are likely to have identified similar, but not identical indicators.

Figure 1

Department of Education Programs

Department of Health and Human Services Programs

Office of Special Education Programs
Part C

Administration for Children and Families
Head Start
Early Head Start

Office of Special Education Programs
Part B, Section 619

Administration for Children and Families
Child Care
CCDBG

State Lead Agency

Part C
or
619

Office of Elementary and Secondary Education
Title 1

Health Resources and Services
Maternal and Child Health
Title V - CSHCN

Office of Elementary and Secondary Education
Even Start
(Part B of Title 1)

Administration for Children and Families
Children's Bureau
Welfare - TANF
Child Protection

Office of Elementary and Secondary Education
Migrant Education
(Part C of Title 1)

Centers for Medicaid and Medical Services
Children's Health Insurance Program

No Child Left Behind and Good Start Grow Smart

The current "state-of-the-development" of federal accountability initiatives, described in the previous section, presents challenges to state policy makers, who are attempting to develop state accountability systems for young children with disabilities. The current federal efforts are not coordinated, nor are the goals and indicators selected by these various initiatives always the most important. Many have been selected for expediency – they are the goals for which there are national data available. It is very likely that the federal goals for the various agencies and their programs will continue to evolve. Hence, it would be unwise to build state accountability systems entirely on the shifting sands of the current federal goals and indicators. State policy makers need to keep these requirements in their sights, but should have a broader and more futuristic vision and array of desired outcomes, if they are going to design accountability systems that provide the type of child and family outcome data sought by state legislatures, as well by as the U.S. Congress.

The identification of standards, outcomes, and their accompanying measurement processes is critical to the ability of state agencies to provide effectiveness data for the programs serving young children with disabilities and their families. However, state policy makers and administrators must take into consideration the multiple and complex systems in which young children receive early intervention and early childhood services, as well as the multiple levels of accountability currently in place, or in development across these systems. This requires that state program administrators interact with one another in order to be knowledgeable about the accountability initiatives, requirements, and measures of other programs.

The state-of-the-art for serving children with disabilities continues to be framed within the context of inclusive services. For infants and toddlers, this framework includes services within "natural environments"; for preschoolers, "least restrictive environments." Because the early childhood system of services includes multiple agencies and programs that differ from state to state, addressing accountability and measuring outcomes and impacts within a single program can be misleading. For example, a 2-year-old with disabilities may receive early intervention services in the home and attend a childcare program during the day. This same child also may receive services through Children with Special Health Care Needs (CSHCN) and the family receives Temporary Aid to Need Families (TANF). If the child's interventions are embedded in routines and provided by the child's family and caregivers in the child care program, who or which program should claim responsibility for the child's gains? If the child's health was not stabilized, he would not be well enough to attend child care or participate in intervention. Without sufficient nutrition to stabilize health and provide sufficient energy, once again, the child would not be able to optimally take advantage of intervention opportunities. So who is responsible for the child's gains? Everyone contributed: family, child care givers, early interventionists, CSHCN staff, and TANF staff. Similarly, a 3-year-old child with disabilities may attend a blended public school (619 and Title 1) and Head Start program in the morning and a childcare program in the afternoon. Once again, who is responsible for his progress?

Local programs are being asked to gather more discrete data on how the program has directly benefited the child and/or family. As a result, agencies at the local level that serve children through multiple education and human service programs (e.g., Head Start, child care, public schools, Early Intervention) are expected to provide data to multiple state and federal agencies. All too frequently these local programs are burdened by being requested to provide different data to different agencies at different intervals.

Multiple Accountability Initiatives

As previously indicated, there are multiple accountability initiatives at the federal and state level. These initiatives have been created separately and are evolving on separate, yet often parallel tracks.

Unfortunately, in some instances, there is the prospect of tracks intersecting and trains wrecking. Two of these broad initiatives focus on children below compulsory school age, while the other two broad initiatives focus on school-age children. The four parallel accountability initiatives are described below. Each of these initiatives has taken place within a particular organizational culture that has influenced the process, target, and terminology used.

First, over the last 20 years, the educational system for school aged children has seen several cycles of reform. The objective of these various reform movements has been to demonstrate increased student achievement (McLaughlin, Henderson, and Rhim, 1998). Consequently, the public schools in each state have developed educational standards for school-aged children, usually K-12. The focus of the current cycle of education reform is based on requirements within No Child Left Behind (Elementary and Secondary Education Act, 2001) and reflects the demands for higher student achievement and increased accountability. Demands have resulted in the development of "high stakes" accountability systems for both the K-12 educational system and for the students as well. Years ago, children with disabilities often were excluded from large scale achievement tests. However, current legislation requires their inclusion.

On the second track, the Individuals with Disabilities Education Act (IDEA), re-authorized in 1997, addresses the assessment of pre-kindergarten and school-aged children with disabilities for accountability purposes. According to IDEA '97, states must establish performance goals and indicators for children with disabilities that are consistent with goals set for typically developing children. Furthermore, children with disabilities must be included in any state or district-wide assessments that are established for typically developing children. The inclusion of children with disabilities in statewide assessment systems is not only supported by IDEA, but also by Section 504 of the Rehabilitation Act of 1973 and by the Americans with Disabilities Act of 1990 (ADA). Since assessment systems are assumed to result in direct benefits to children (Thurlow, Elliott & Ysseldyke, 2003), exclusion of children with disabilities from such assessment would be a violation of Section 504 and the ADA. More recent support for inclusion in state assessment systems has come from the No Child Left Behind Act (the Elementary and Secondary Education Act) discussed above, which requires that states report the results of yearly assessments in grades 3 through 8 and one assessment in grades 10 through 12. The scores from these assessments must be disaggregated by gender, disability, ethnicity, English proficiency and migrant status (U.S. Department of Education, 2002).

On the third track, the federal and state governments have continued to increase their investment in early care and education programs for preschool children. These programs include programs for preschool-aged children with disabilities and their families under IDEA, Title 1, Head Start, Early Head Start, Early Reading First, Even Start and state funded pre-kindergarten programs. As the investment in early childhood has grown, so too has the expectation of documented results. State Pre-k programs are often influenced by the school accountability initiatives, attempting to fit into the already developed academic standards. Most recently, the Good Start Grow Smart Initiative by President Bush identified three major reasons why the quality of early care and education systems remained low: 1) lack of alignment between what happens in early childhood settings and expectations once children are in school; 2) lack of program evaluation data that includes how children are being prepared to succeed in school; and 3) little dissemination of information about how to prepare children to be successful in school. The Good Start Grow Smart legislation addresses accountability of early childhood programs by asking states to voluntarily develop guidelines

on literacy, language, and pre-reading skills for children ages 3 to 5 that align with State K-12 standards. It is expected that these guidelines could be adapted for use across the various early care and education settings.

The final track is the accountability initiatives engaged in by some state Part C programs for infants and toddlers with disabilities. These states sought to design an "evaluation" of their Part C programs. Instead of using the term "standards," most of the states used the term "outcomes." The outcomes identified fell into one of three categories: system or program, family, child (Roberts, Innocenti, & Goetz, 1999). However, the most frequently identified outcomes were system outcomes. Those individuals who designed the early Part C evaluations were confronted with many challenges and consequently formed a Consortium to increase their effectiveness in their individual states. Unlike many of the school-age accountability initiatives, most of the evaluators were University researchers who received a sub-contract from the state Part C lead agency, many of which were not education lead agencies. In addition, unlike the school-age accountability systems, the Part C evaluations looked beyond the child to the family and beyond the child's education (cognitive, language, etc.) to the child's health and well-being.

Confusion About Terms

A Nation At Risk (1983) and several reports that followed helped to trigger a major effort to address the quality of American schools at all levels through higher standards (Kelly, 1993). The concept of standards is not unique to the field of education. Standards are used throughout many industries (technology, medicine, etc.) to help ensure and monitor quality. Therefore, during the accountability movement in education, researchers and administrators drew from a number of sources outside education to design their standards and accountability systems. Likewise, as the accountability movement has moved into the early childhood arena, researchers and administrators have looked to the K-12 accountability system as a source of information.

While the concept of borrowing and adapting from others is reasonable and expected, and the early childhood field can benefit from the research and experiences of the K-12 system, the wholesale use of K-12 terminology in the design of early childhood accountability systems can cause confusion for early childhood professionals. This is especially true given the number of diverse programs and theoretical perspectives that make up the early childhood community. Some of the most commonly confused and debated terms in use today are "standards" and "outcomes." Basically, there is some debate as to whether the terms standards and outcomes can be used synonymously, or if there are differences between the two. In partial response to this debate, it should be noted that historically, the use of the term standards is more common for programs (i.e., public schools) with an overall focus on the provision of services for typically developing children; while the use of the term outcomes is more common for programs with an overall focus on specialized populations (e.g., children in poverty, children with disabilities).

In keeping with this concept, to date over 40 states have developed or are in the process of developing state Early Childhood standards, and many of these states use the term "standards" as a way to describe the knowledge and skills children need or how we can measure whether children are making progress (Rous, Kohner-Coogle & Stewart, 2004). The programs involved in this effort at the state level are varied, and include state funded preschool programs and preschool programs for children with disabilities. Essentially these state documents describe two concepts: content and performance standards.

Traditionally the term "standards" has been used to articulate the level of quality or the distinctive skills or characteristics to be: 1) attained by children; 2) accepted as the norm; and 3) measured to judge progress. For example at the child level a standard could be, "the child will understand the effect of his or her behavior on others in the environment."

Along with the development of state standards for young children, there are some federal programs that are in the process of developing outcomes. Head Start is using an Outcomes Framework to describe how children are benefiting from the Head Start program. The Office of Special Education Programs (OSEP) is in the process of identifying outcomes as a way to measure the impact of Part C and 619 programs on children and families. Outcomes have traditionally related to the result or the effect of services, instruction, and/or programs on the attainment of skills or characteristics. An example of a child level outcome is, "the child uses language to convey and comprehend communicative and social intent."

Multiple Levels of Standards and Outcomes

While the most recent emphasis in early childhood has been on the development of standards and/or outcomes at the child level, attention also should be given to the interconnectedness of child level standards and outcomes and other levels of standards and outcomes at the state or national level. Standards and outcomes that are developed for the program and personnel levels can have an impact on the ability of program staff to fully understand, implement and track progress of child outcomes and standards. Each of these is discussed more fully below.

Personnel standards and outcomes are most typically designed at the state program level. These standards can take several forms. Many states have developed specific certification programs for early childhood and early childhood special education personnel, sometimes through blended programs. These certification programs generally develop teacher or provider standards that provide the framework for university course construction and delivery. Some states have expanded their personnel level standards by developing professional core content and articulation agreements. Core content provides information on what professionals need to know and be able to do at various levels (entry through masters) to work effectively with young children and their families. Other states have developed a set of personnel competencies which must be documented.

There are three types of program standards or outcomes. In the first type, there are broad indicators used to determine overall effectiveness of programs. For example, Maternal and Child Health programs may collect data related to the birth and death rates of young children to determine if the state program is having an impact in reducing the child death rates. If Part C and 619 programs wanted to use this type of indicator to judge the effectiveness of their program or system, examples might include: earlier identification (decrease in mean age at entry), increase in number of informed consumers, increase in smooth transitions. State agencies also may develop cross-program outcomes and indicators and then identify existing data sources that can be used to track change over time. An example of this type of initiative is the School Readiness Indicator Initiative - a 17 state initiative designed to help states develop school readiness indicators and to develop a comprehensive national set of school readiness indicators to inform public policy for young children and their families (Rhode Island Kids Count, 2004).

The second type of program standard is linked to monitoring the quality of services. These standards often are based on federal monitoring requirements. Local programs must comply with these standards in order to continue to serve children. Infant-Toddler and 619 program standards are an example of this type.

Finally, most states have in place state specific program or environmental standards. Most typically, program standards are related to regulation and state licensing requirements for the specific programs (e.g., child care). More states have added an additional level of program standards that are designed to enhance the quality of child care programs. In some states, these quality enhancement systems are designed around a rating system for early care and education programs across agency types that often include a self-study component and tiered rating process. As more state programs are designing service delivery systems for young children in inclusive settings or natural environments, these rating systems have increased the number of items that address the issues of young children with disabilities. This is also true at the national level for the new National Association for the Education of Young Children (NAEYC) accreditation process, which has increased the number of items across all areas that specifically address the special needs of children in early care and education settings.

Alignment of Initiatives

Regardless of which term is used, there are commonalities between standards and outcomes at all levels to intent and purpose (Shore, Bodrona, Leavy, 2004). Both standards and outcomes can be used to provide a framework through which administrators, teachers, providers and families can articulate expectations. Once expectations are clear, then the standards and outcomes provide a mechanism by which changes in behavior can be observed and recorded.

Given the two approaches in the frameworks used to help programs measure impact and change in behavior (i.e., standards approach and outcome approach), early intervention and preschool programs within states must determine the path they are going to take in designing their own accountability processes. There are three choices: 1) design a specialized set of expectations for early intervention and preschool special education that stand alone, 2) develop a set of expectations that are clearly aligned with expectations from across the diverse and inter-related educational programs within the state and support the inclusion of young children as part of the larger accountability process, or 3) develop a set of expectations that are not only aligned with educational programs, but with health and other human services as well. For states that opt to design systems that are congruent and provide alignment among programs at the state level, consideration should be given to the design of the state standards and outcomes in terms of key features (e.g., ages and populations), areas addressed within the standards and outcomes, and measurement and reporting processes. Table 1 provides an example of the alignment across the various early childhood accountability initiatives in Kentucky.

Table 1   Alignment of Accountability Initiatives in Kentucky

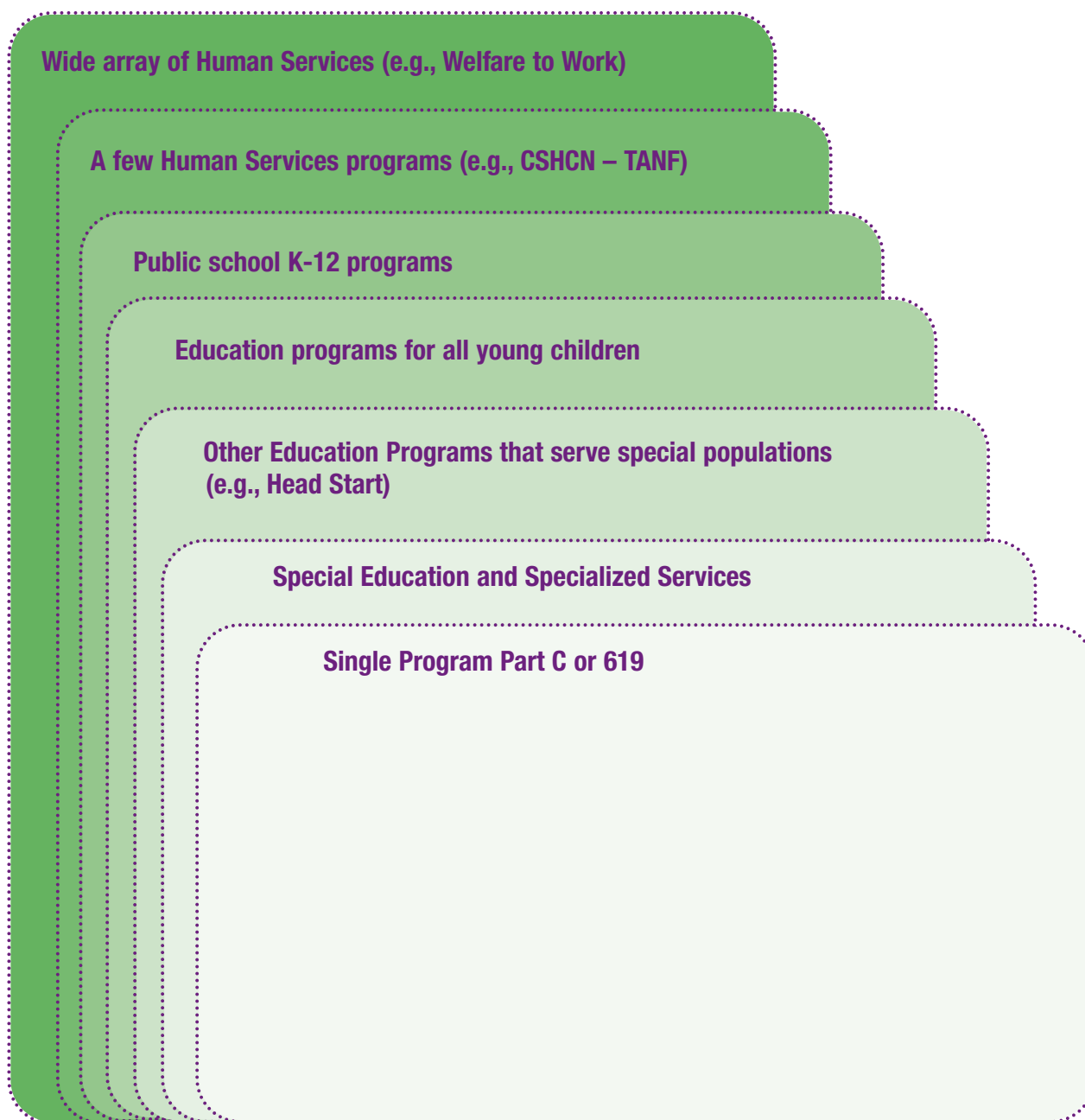| Key Features	Birth to 3 (KY)	3 to 5 (KY)	Head Start (National)	K – 12 (KY)
Population	Universal	Universal	Low income children	Universal
Type	Standards	Standards	Outcomes	Standards; Some Expansion Documents to Address Specific Areas for Children with Special Needs
Alignment	Aligned with 3-4 Standards	Aligned with K-12 Standards and Head Start Outcomes		
Areas Addressed	Communication	Language Arts	Language Development Literacy	Language Arts
	Cognitive	Mathematics Science	Mathematics Science	Mathematics Science
	Social Emotional	Social Studies (Relationships)	Social Emotional	Social Studies
	Motor	Health Education Physical Education	Physical Health/Development Approaches to Learning	Physical Education Health Education
	Creative Expression	Arts and Humanities	Creative Arts	Arts and Humanities
Measurement and Reporting	Continuous Assessment System with Recommended Instruments for Screening, Diagnostic, Classroom/ Instructional	Continuous Assessment System with Recommended Instruments for Screening, Diagnostic, Classroom/ Instructional	National Assessment System	Commonwealth Accountability and Testing System
	Aligned with Child Standards	Aligned with Child Standards	Aligned with Child Outcomes	Aligned with K-12 Program of Studies

There are a number of critical questions that states will need to address when designing state level accountability systems. The first set of questions relates to the level and alignment of standards and outcomes. Specifically, at what level and in what order should standards and outcomes be aligned (Figure 2)? Is it critical to only align standards and outcomes that are developed across programs within a state or should states align state standards and outcomes to national outcomes, specifically those that have been developed or are in the process of being developed for specific programs and populations (Head Start, OSEP)? If states develop program specific outcomes, how do these outcomes link with outcomes developed at the national level (i.e., Part C state outcomes and OSEP outcomes)?

Figure 2
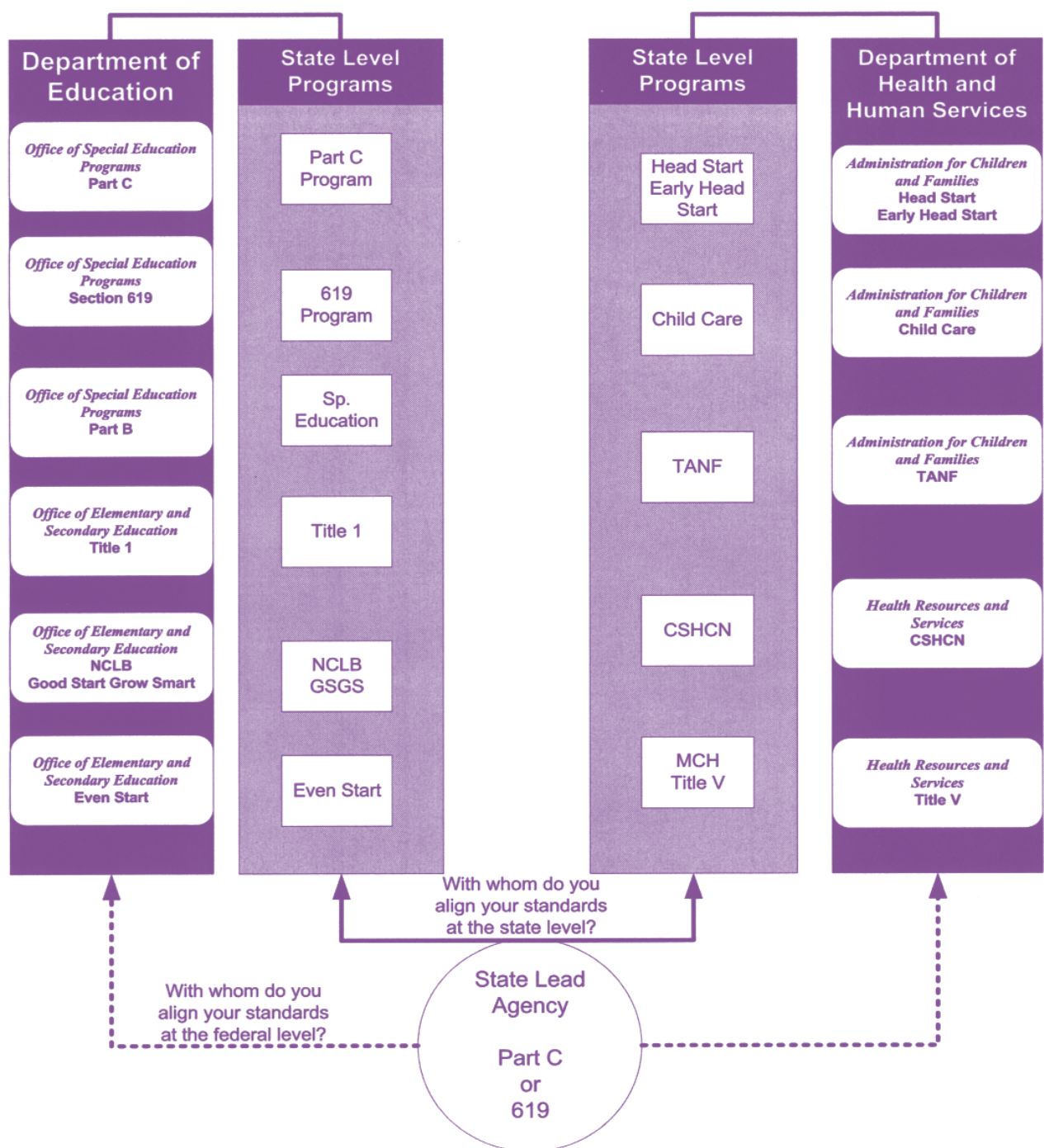
At another level, state agencies must consider the multiple layers of accountability (Figure 3) given that specific programs generally operate within the context of a broader system of services.

Figure 3   Layers of Accountability

Wide array of Human Services (e.g., Welfare to Work)

A few Human Services programs (e.g., CSHCN – TANF)

Public school K-12 programs

Education programs for all young children

Other Education Programs that serve special populations (e.g., Head Start)

Special Education and Specialized Services

Single Program Part C or 619

For example, Part C and 619 programs operate within the context of other programs for specialized populations; early childhood programs operate within the context of the broader educational system. Therefore, should state policy makers develop state specific outcomes for specialized programs (e.g., Part C, 619) operating within the state and align those within a comprehensive set of state early childhood standards that are more universally designed? Should state policy makers attempt to align standards and outcomes across all early childhood programs and then align with the public school standards? Should Part C and 619 state policy makers gather all existing standards and outcomes and take these into account when developing Part C and 619 standards or outcomes? Finally, should state policy makers from all programs come together to collaboratively develop a single and cohesive set of standards or outcomes?

The second set of questions relates to the implications for local programs. How will multiple levels of standards and outcomes affect local programs? If local programs are expected to provide multiple types of data for multiple types and levels of standards and outcomes, what is the impact on service delivery, staff development, staff retention and ultimately child and family outcomes? For example, suppose that a school district is a Head Start grantee and the district also receives funds for preschool children with disabilities. This preschool program must comply with OSEP monitoring and data provision requirements, the Head Start monitoring system, and the state's preschool or pre-kindergarten accountability requirements. If the school provides pre and post program child care, it also must meet the child care licensing requirements. The lack of similarity in these is burdensome to local school personnel.

The third set of questions relates to measurement of child and family outcomes. To what degree will multiple levels of standards and outcomes affect the ability to collect valid and reliable data on child and family outcomes within a specific program? If children and families are served in multiple programs over the course of a day or week, can positive or negative outcomes be solely attributed to that specific program? Can a single assessment process be used across multiple program standards and outcomes? A child in the program described above would receive testing as part of the 619 program, have to participate in the Head Start National Assessment, and possibly would be required to participate in the state's school accountability assessment.

The fourth and final set of questions relates to the use and implications of universally designed standards and assessment systems. Can universally designed state standards and outcomes be used as a mechanism to measure multiple levels of outcomes across programs? Can universally designed early childhood standards be appropriately linked with individual sets of outcomes identified by multiple programs, in a way that provides continuity across early childhood programs? Can assessment systems that are linked with universally designed state standards, in turn be linked with specific program outcomes from multiple programs in a way that provides valid and reliable data to measure child and family outcomes?

The next page contains a set of questions (Questionnaire A) to guide the development of state accountability systems.
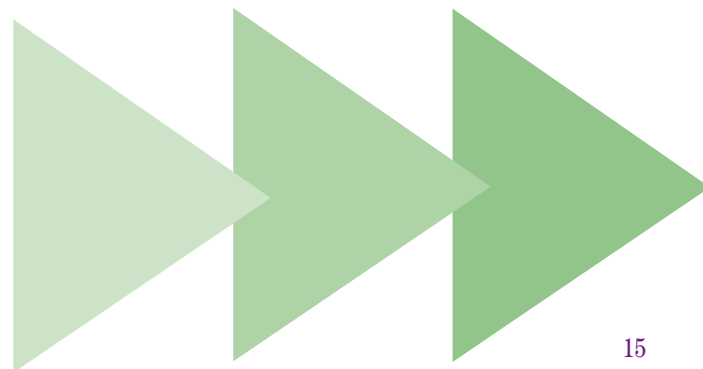
Questions to Guide the Development of Accountability Systems

1. At what level will early childhood state standards or outcomes be developed?
   [Check all that apply].

   ☐ Child level
   ☐ Program level
   ☐ Personnel level
   ☐ State level

2. What is the overall frame within which the standards or outcomes will be developed?
   [Check only one option].

   ☐ Individually by program
   ☐ Jointly between programs that serve like populations (e.g., special needs, at-risk)
   ☐ Collaboratively designed to be used universally with all programs across the state

3. What is the overall frame within which different levels of standards will be aligned?
   [Check only one option].

   ☐ Standards and outcomes for each level will be developed and implemented
     independently
   ☐ Standards and outcomes across all levels will be considered during the development
     and implementation process
   ☐ Standards and outcomes across all levels will be linked and referenced
     (e.g., child standards will be tied to personnel competencies)

4. What process will be used for the identification and development of standards or outcomes
   across all levels? [Check all that apply].

   ☐ Lead agency selects
   ☐ Revision and extension of existing standards/outcomes
   ☐ Adoption and use of standards or outcomes identified by another state
   ☐ Revision of standards or outcomes identified by another state
   ☐ Use of work group with Lead agency representatives
   ☐ Use of workgroup with multiple agency representatives
   ☐ Use of university faculty across the state
   ☐ Literature review

5. What formats will be available? [Check only one option].

       ☐ One version for all programs, professionals and families
       ☐ Professional version only
       ☐ Single document that includes professional and family version
       ☐ Separate documents for professionals and families

6. What process will be used for validation of standards and outcomes at all levels?
    [Check all that apply].

       ☐ Internal State Review by Experts
       ☐ External Review of National Experts
       ☐ Field Review and Input by State Providers
       ☐ Field Review and Input of Nationally Representatives Group

7. What process will align data collection and measurement processes across all levels of
    standards and outcomes? [Check only one option].

       ☐ Agencies collect and report data independently
       ☐ Agencies collect data independently and share with other agencies
       ☐ Agencies collect data, share with single source who reports data by area and program
       ☐ Agencies collect data and enter into single management systems to generate report
       ☐ Third party collects and reports data

8. How will the differences in terms be addressed? [Check only one option].

       ☐ No one seems to realize or recognize that different programs or initiatives are using
         different terms (e.g. outcomes, standards, indicators, benchmarks)
       ☐ Our agency will use the terms used by our federal funding agent.
       ☐ Our agency will use the terms used by another initiative, because we feel it makes
         sense to do so.
       ☐ We will engage multiple agencies and initiatives in a conversation to see if we can
         find some commonalities.
       ☐ All agencies have agreed to use the same terms within our state.

Although all states have been providing services statewide to eligible infants, toddlers, and preschool children with delays and disabilities for over a decade, little is known about the quality and effectiveness of the diverse services provided in each state. Although research indicates that exemplary programs yield positive outcomes (Faran, 1990; Guralnick, 1997), little is known about whether the set of hypotheses undergirding Parts B & C of IDEA are accurate and whether they have resulted in the intended outcomes in typical programs (Harbin, 2000).

As mentioned earlier, recent interest in accountability at both the federal and state level has increased the states' interest in determining the results of service delivery (Patton 1997; Roberts et al., 1999). Members of Congress, as well as members of state legislatures across the country, are eager to know if the money invested in these programs for infants, toddlers, and preschool children with disabilities and their families have resulted in positive benefits for the targeted children, and their families. The enactment of the Government Performance Results Act (GPRA) by the U.S. Congress demonstrates the increased desire and demand for accountability.

State legislatures also provide substantial funding for services to infants, toddlers and preschool children with disabilities, and just like the U.S. Congress, are increasing their demands for accountability. No longer is it sufficient to report merely the number of children served or the number of services provided (Kibel, 1999; Patton 1997; Roberts et al., 1999). Legislatures across the country are beginning to ask for, and in some states demand, an account of more specific results and benefits obtained from precious fiscal resources. The message is becoming clear: if you can't tell us just how this funding is making a difference in the lives of children and their families, agencies will no longer continue to receive funds. Therefore, some state agency administrators of Infant and Toddler or Preschool (Part C of IDEA) programs for children with disabilities (Section 619, Part B of IDEA) are eager to identify outcomes or standards, as well as accompanying measures to justify the adequacy of these service programs.
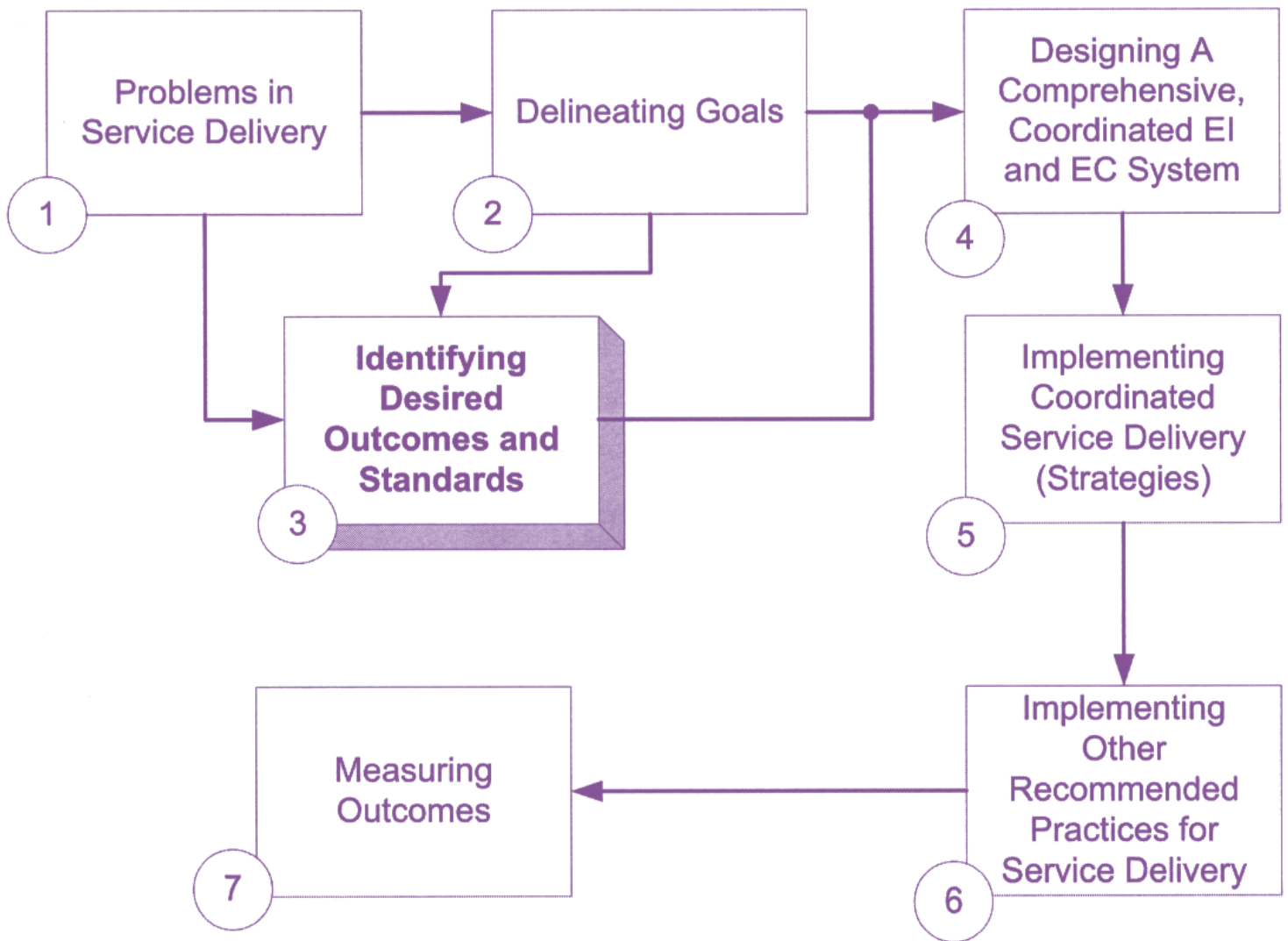
Since the legislation was enacted in 1986, why is it, then, that we are not further along in the process of evaluating the results of Early Intervention and Preschool programs? The answer lies in the complexity of identifying, defining, and measuring outcomes, or results, of this complex service system. As is frequently the case with public policy, legislation was enacted and the field struggles to implement it (Bailey, McWilliam, et al. 1998; Harbin, 1993; Knapp, 1995; Roberts et al., 1999; Weatherly, 1979).

The Logic Model depicted in Figure 4 is a commonly used program planning and evaluation framework (Gallagher, Surles, & Hayes, 1973; Patton, 1997; Stufflebeam, 1985). This model for evaluating services indicates that desired outcomes should be delineated early in the process (see box 3) in order to guide evaluation. However, due to the many complexities of serving infants and toddlers, as well as the demands for immediate implementation requiring many changes in the way services are delivered, professionals, policy makers, and parents have focused their energy in boxes 4, 5, and 6 (designing and implementing this multi-agency service system, and implementing multiple recommended practices.) Before we can measure the outcomes (box 7) and answer the accountability questions posed by Congress and state legislatures, we must first delineate the desired outcomes.

However, to merely expediently select outcomes or standards without knowing or understanding which outcomes or standards are considered most important is not only illogical, but ineffective.

Figure 4 – Logic Model Evaluating Coordinated Services

Problems in Service Delivery **1**

Delineating Goals **2**

Designing A Comprehensive, Coordinated EI and EC System **4**

Identifying Desired Outcomes and Standards **3**

Implementing Coordinated Service Delivery (Strategies) **5**

Measuring Outcomes **7**

Implementing Other Recommended Practices for Service Delivery **6**

Although it may seem easy and straightforward, identifying child and family outcomes and standards is a complex process. Agreement regarding the selection of child and family outcomes of comprehensive, coordinated, inclusive services has not yet been achieved (Spiker et al., 1996; Guralnick, 1997; Knapp, 1995). Several factors have made the identification of service outcomes difficult.

Foremost among these is the breadth of possible outcomes and standards, given the legislative and programmatic goals of the service system, as well as the diverse population served. In addition to the legislative goals, Bailey and Wolery (1992) drew on the professional literature and suggested several specific goals. Harbin and her colleagues (2004) conducted a series of structured focus groups with seven diverse stakeholder groups (families, service providers, local program administrators, state policy makers, state technical assistance staff, and national experts). Harbin used the child and family outcomes identified through these focus groups to construct an instrument that was subsequently used in a Delphi study containing the same broad stakeholder groups. Participants achieved consensus on 21 broad child outcomes; 19 broad family outcomes, and 52 system outcomes or program standards, which are grouped into eight broad categories, such as system entry, service planning, and consumer knowledge. The outcomes identified by these diverse stakeholder groups are contained in Questionnaire B following the guiding questions.

Many states have undertaken their own processes to identify outcomes and standards to be addressed by the state and its communities (Lund, Rous, Moffett, Wood & O-Keefe, 2002; Roberts et al., 1999). However, there is still no national consensus about the outcomes or standards to be addressed. Therefore, there is no way to compare the progress of children and families in one state with those in another. Nor is there any way to aggregate data across states to provide a national picture of the benefits of service provision. Consequently, as mentioned earlier, the Office of Special Education Programs (OSEP) in the U.S. Department of Education (U.S. DOE) recently funded the Early Childhood Outcomes Center (Hebbeler et al., 2003) to try to achieve consensus and facilitate more consistency across the country.

A second but related issue is the question: who is the target of EI or EC programs for children with disabilities? In spite of the fact that the law requires the development of the Individualized Family Service Plan (IFSP), which defines the family as a legitimate target of intervention and support (Harbin, 1993), most evaluations generally have focused on progress made by the child. Only recently have family outcomes begun to be systematically identified and measured (Harbin & Neal, 2004; Harbin, McWilliam & Gallagher, 2000; Kahn, 2000; Roberts et al., 1999).

The third difficulty in defining outcomes is the diversity of the population that is being served (i.e., differences in degree of severity of the disability, type of disability, and cultural differences). Individual needs means individualized services; individualized services results in individual outcomes. Is it possible to have a single set of outcomes or standards for such a diverse group of abilities and interventions? Differentiation among goals, outcomes or standards, indicators or benchmarks, and strategies used to achieve outcomes and meet standards is the fourth factor that hinders consensus.

Even after individuals gain a better understanding of the definition of "outcome" or "standard" they often experience a fifth barrier, and constrain their identification of outcomes or standards because they fear that their agency would be held responsible for achieving difficult outcomes and standards. Such "fear among providers . . . that they're being asked to take responsibility for, and be judged on, something over which they have little control" (Patton, p. 158) is a common one, but needs to be overridden if we are to truly answer the questions "What is the Early Intervention and Early Childhood system trying to achieve with children and their families? How will children and families be different as a result of participation in Early Intervention and Early Childhood programs?" (Paraphrase of Patton, p. 154). Questionnaire B contains the questions to guide the selection of outcomes and standards.
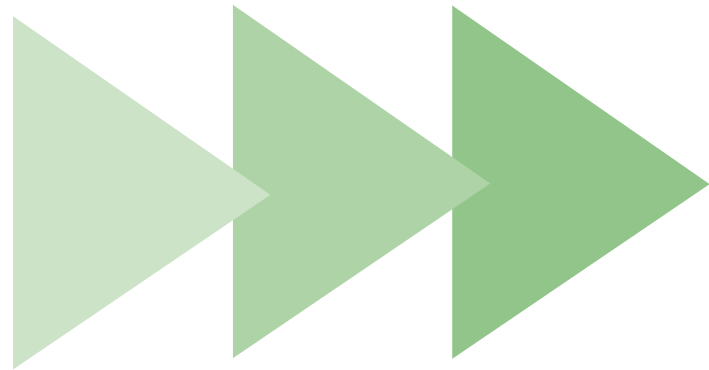
Questions to Guide Selection of Standards and Outcomes

1. At what level do you wish to measure outcomes and standards? [Check all that apply].

    ☐ Child
    ☐ Family
    ☐ Program
    ☐ Multi-agency service system
    ☐ Personnel

2. What is the age range of the children targeted for this evaluation? [Check only one option].

    ☐ Birth-to-3
    ☐ 3-to-5
    ☐ Birth-to-5
    ☐ Birth-to-8
    ☐ Other

3. Who are the children targeted for this evaluation? [Check all that apply].

    ☐ Children with delays and disabilities
    ☐ Children with multiple risk conditions
    ☐ Children with limited types of specified risks
    ☐ Children with any type of risk condition, biological or environmental
    ☐ All young children

4. What is the breadth of the array of services to be evaluated? [Check the primary services to be evaluated with a "P", and the secondary services to be evaluated with an "S"].

    ___Child education & therapeutic services   ___Family basic needs (food, clothing, housing)

    ___Childcare   ___Adult education and information

    ___Child health and medical   ___Adult training

    ___Child protection   ___Family support

    ___Food   ___Mental health

    ___Clothing   ___Economic

    ___Housing   ___Transportation

    ___Dental   ___Cultural/social

    ___Mental Health   ___Recreation

    ___Family medical   ___Legal

    ___Family dental

5. Which agencies are the target of evaluation? [Check all that apply].

☐ Lead agency
☐ Other agencies that provide educational and therapeutic services
  (e.g., Head Start, Early Head Start, Even Start, Public preschool)
☐ Other human service providers of child services (e.g. health, child protection,
  food, housing)
☐ Agencies that provide family services
☐ Agencies that provide services to children at risk
☐ Agencies that provide services to typically developing children

6. What is the process of identification of outcomes or standards? [Check all that apply].

☐ Lead agency selects
☐ Use of outcomes or standards identified by another state or community
☐ Use of outcomes or standards identified by other accountability initiatives within
  the state
☐ Use of families, PTI, or other family support group
☐ Use of focus groups, but not rigorous
☐ Use of rigorous focus groups, with process for obtaining widespread stakeholder
  input into and validation of, outcomes obtained by Focus Group
☐ Use of agency representatives beyond the lead agency
☐ Literature review
☐ Delphi study

7. Will there be indicators selected? [Check one].

☐ Only outcomes or standards - no indicators
☐ Each outcome or standard will have one or more indicators
☐ Scores on assessments will serve as the indicators
☐ Benchmarks across a developmental continuum

8. What is the relationship with outcomes or standards selected by other Early Childhood and Early
    Intervention initiatives? [Check only one option].

☐ No relationship
☐ Have identified the outcomes or standards identified by other initiatives in the state
  and tried to be as similar as possible
☐ Asked other initiatives to provide feedback on the outcomes or standards selected
  by lead agency
☐ Representatives of other initiatives actively participated in the outcome or standard
  selection process
☐ There is a collaborative group that is selecting a set of outcomes or standards to be
  used across all Early Childhood and Early Intervention initiatives

Child Outcomes	Family Outcomes
• Child is accepted by the community.	• Families participate more in their community and make use of natural community resources.
• Child is able to function appropriately in settings where children without disabilities are served.	• Families demonstrate an increased capacity to plan for and meet their child's basic needs (e.g., food, shelter, clothing, and health).
• Child's development is improved in the individual areas targeted for intervention.	• Families possess realistic expectations for their child.
• Child communicates effectively.	• Families demonstrate an increased capacity to enhance their child's development (e.g., social, emotional, cognitive, language, motor, and self-help).
• Child's potential is maximized.	• Families' relationship with their child is improved.
• Child's ability to be appropriately engaged in the environment is increased.	• Families experience increased meaningful participation in decisions regarding their child's services and placements.
• Child is able to use skills and abilities in different settings (generalization).	• Families have increased knowledge of the available specialized and non-specialized community and state resources.
• Child adapts to change.	• Families possess increased understanding of how to navigate the service system.
• Child is self-directed and experiences increased independence.	• Families are able to comfortably incorporate professional assistance in family life.
• Child experiences feelings of success.	• Families possess increased knowledge of services and options.
• Child is able to remain living with his/her family instead of being placed in an institution.	• Families know and exercise their rights.
• Child's level of disability is stabilized or reduced.	• Families are effective advocates.
• Child's need for specialized services is reduced.	• Families possess a sense of empowerment.
• Child has more friends.	• Families possess increased knowledge of financial resources.
• Child experiences improved relationships with significant adults.	• Families experience increased financial stability.
• Child's health is improved and does not negatively affect child's development.	• Families experience increased stability and cohesion.
• Child's basic needs are met.	• Families experience increased quality of life.
• Child is able to "be a kid," participating in and enjoying activities similar to those of typically developing children.	• Families have increased confidence to address life's challenges.
• Child is happy, feels loved, and feels good about self.	• Families have satisfying relationships with friends, family, neighbors, and professionals.
• Child is in a stable, permanent, and safe living environment.	• All family members are healthy.
• There is a decrease in the number of injuries sustained by child.	

System Entry
- all children are found and served
- equal access
- early identification and entry
- timely entry
- easy entry

Service Planning
- overlaps and gaps in services identified and addressed
- IFSPs/IEPs contain all needed services
- planning (IFSP/IEP) for individuals reflect coordination and integration across providers and agencies
- plans reflect strengths and assets of families and children

Service Provision
- child's needs are met
- family's needs are met
- services are accessible
- broad array of services and resources received
- sufficient services and resources provided
- coordinated infrastructure
- natural settings and supports used
- quality of services high
- services integrated into broader array of community resources
- families' culture and values are respected
- services reflect strengths and assets of families and children

Coordinated Infrastructure
- services coordinated across providers and agencies
- policies and procedures coordinated across agencies
- system is flexible
- there is a structure to facilitate coordinated service delivery
- adequately trained staff facilitate coordinated service delivery
- effective interagency communication facilitates coordination

Transition
- children and families experience smooth transitions and continuity of services
- children and families experience support during transition
- unnecessary transitions are eliminated

Finance
- coordinated funding streams
- increased public support for funding
- sufficient funds available to meet needs
- equal access to services regardless of family finances
- all children and families are insured
- methods for making funding decisions are fair

Accountability
- parents' complaints are addressed quickly
- families have easy access to due process and mediation
- families are satisfied consumers
- community leaders satisfied
- service providers satisfied with job effectiveness
- all child and family outcomes are achieved in each community
- staff feels comfortable and safe in advocating for families
- parents employed throughout system to ensure that their needs are formally represented
- families and providers perceive increase in community acceptance
- mechanism exists for multi-agency feedback, planning, and continual improvement of interagency interactions
- compliance monitoring is collaborative and proactive

Consumer Knowledge
- families informed consumers about identifying services and developing child's plan
- families informed consumers about services and system
- families informed consumers about financial resources
- families informed consumers about transition
- families know they have an array of options for services

As indicated earlier, the FY 2002 PART Assessment of the Infant/Toddler and Preschool programs of IDEA resulted in the designation "Results Not Demonstrated" for both programs. In addition, the PART assessment found that new measures were needed for both programs. Specifically, the absence of child performance data is indicated for both infant/toddler and preschool programs. As a result, the Department of Education, and specifically the Office of Special Education Programs, finds itself in need of developing a system for measuring the impact of these important federal programs on the children and families served. Unfortunately, there is a growing realization among administrators and researchers alike that considerable work will be needed to develop the measures needed (Carta, 2002; Shonkoff, 2002; Wolery & Bailey, 2002).

While accountability for school-aged children with disabilities has received considerable attention in recent years (Thompson & Thurlow, 2003), the Early Intervention and Early Childhood Special Education communities may perhaps be best characterized as approaching the awareness level relative to systems for measuring progress towards standards or outcomes. Existing instruments in the field, developed primarily for the purpose of determining eligibility (norm-referenced) or for program planning (criterion-referenced), may not be appropriate for the task of annual performance reporting of child progress. The inadequacy of classic psychometric theory and the advantages of other measurement models for early childhood assessment have been discussed for more than a decade (McConnell, Priest, Davis & McEvoy, 2002; Neisworth & Bagnato, 2004; Snyder & Sheehan, 1992). However, other assessment models have been applied to the assessment of young children only infrequently. Two examples will be provided here.

While virtually all school-aged accountability assessments follow the Item Response Theory (IRT) measurement model, there has been little application of this model to the assessment of young children. One notable exception is the work going on in the state of California, funded by the California Department of Education's Child Development Division and Special Education Division (McLean, Kuschner, Heal & Drouin, 2004). California has developed an instrument called the Desired Results Developmental Profile (DRDP), which is an observational, curriculum-embedded instrument for children from birth-through-age-four covering the major domains of early child development. The DRDP employs Item Response Theory, specifically the Multidimensional Random Coefficients Multinomial Logic (MRCML) model, to yield scale scores. In addition, adaptations have been developed (the DRDP access), so that the DRDP will be appropriate for most children with disabilities in this age range as well. An example of the application of Item Response Theory (Rasch scaling) to family assessment is currently being developed by the National Center for Special Education Accountability and Monitoring (NCSEAM) mentioned earlier.

Another approach to outcome measurement was proposed by the Early Childhood Research Institute on Measuring Growth and Development (McConnell, McEvoy & Priest, 2002). The ECRI on Measuring Growth and Development developed Individual Growth and Development Indicators (IGDIs). IGDIs are quickly administered and repeatable indicators of child status in the five developmental domains. These domains include: language, social, cognitive, motor, and adaptive. IGDIs are not designed to be comprehensive samples of child performance in a particular domain but rather components or correlates of more comprehensive measures. The IGDIs have been developed following a general outcomes measurement approach (McConnell, Priest, Davis and McEvoy, 2002). General outcomes measurement uses standardized, prescriptive procedures with repeated administration of the same procedures over time. The resulting slope of progress for repeated administrations is used as an indicator of rate of development in the domain. So for example, in the language domain, for preschool children, a picture naming task has been developed.

The number of pictures that a child can name in a designated period of time is measured. This task remains the same across the preschool years. The resulting slope from repeated measurement of the same task over time is interpreted as an indicator of general language development.

Federal law clearly sees the family as a legitimate focus of EI. However, there is a lack of tools to capture the wide array of possible family outcomes. There are a variety of instruments that measure particular areas that might be targeted for intervention. For example, Dunst and his colleagues (1984) developed a tool to measure the amount of social support available to, and used by families. A research team from the Beach Center on Disabilities at the University of Kansas has identified five broad areas, or domains, of family outcomes, and has developed a measure to assess each (Park et al., 2003), with particular focus on the dimension of addressing the family's quality of life (Turnbull et al., 2004). Harbin and Neal (2004), drawing upon the results of the Focus Groups and Delphi study described earlier, developed an instrument to measure a wide array of family outcomes.

The National Early Intervention Longitudinal Study (NEILS) included items to obtain families' perceptions on the impact of EI services on their family, as well as on their child (Bailey et al., 1998). This study used a broader survey approach to assessing outcomes most often used by public health, instead of the types of instruments that arise out of the traditions of psychology. Another national initiative funded by OSEP to identify and measure family outcomes was the ECRI on Measuring Growth and Development (1998) mentioned above. Through a subcontract with PACER, Inc., families identified seven outcomes that were reduced to four by the ECRI research team and advisory board. In January of 2003, the OSEP convened a group of stakeholders to discuss and identify family outcomes and possible measures. This group identified five outcomes and a series of indicators for each. The group discussed measurement issues, but did not reach consensus on measures.

As part of a major longitudinal study, Head Start has been assessing four family outcomes of services as part of the Family and Child Experiences Survey (FACES), and funded by the Administration on Children, Youth, and Families (1998). More recently, the Council on Quality and Leadership (2004) identified 14 family outcomes and suggested an individualized approach to measuring these outcomes. Finally, through a series of multi-stakeholder focus groups and Delphi studies, the Service Coordination Research and Training Center identified both immediate and long-term outcomes of effective service coordination (Bruder, in press). This project is in the process of exploring instruments to measure these outcomes. The previous discussion indicates that more progress has been made to identify family outcomes than in developing tools to measure them. Questionnaire C contains the questions to be used to guide decisions about measurements used in accountability systems.
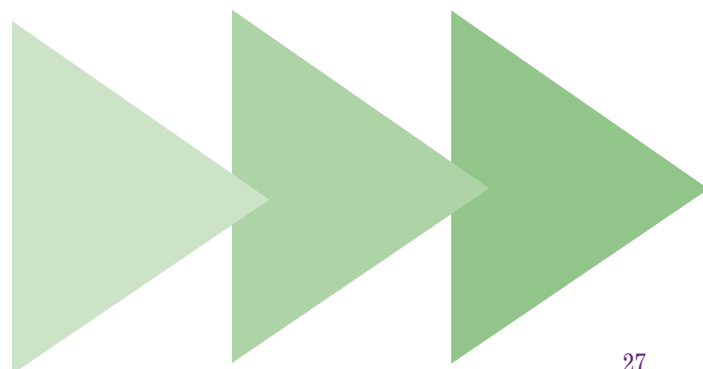
Questions to Guide the Development of Measurement Processes

1. Who will be involved in the assessment?

a. Children
- ☐ All children in EI and EC SpEd programs
- ☐ A sample of children in EI and EC SpEd programs
- ☐ Only children who have been in the program for at least 1 year
- ☐ Only children who are exiting EI or EC SpEd programs
- ☐ All children (typical & atypical) in all EI and EC programs
- ☐ Sample of all children (typical & atypical) in all EI and EC programs

b. Families
- ☐ All families of children with disabilities
- ☐ All families of children with risks and disabilities
- ☐ A sample of families of children with risks and disabilities
- ☐ All families whose children (typical and atypical) participate in any EI or EC program in the community
- ☐ A sample of families whose children (typical and atypical) participate in any EI or EC program in the community

2. When should children and/or families be assessed?

- ☐ At entry and exit from Part C and 619
- ☐ At entry and once a year and then at exit
- ☐ Twice a year
- ☐ Beginning and end
- ☐ Every 6 months
- ☐ If upon entry, should it be immediately or after some amount of time has passed (e.g., 1 month)
- ☐ Must all states assess according to the same timeline?

3. How should English language learners be assessed?

- ☐ All children's language will be assessed in English
- ☐ All children's language will be assessed in the dominant language
- ☐ Both English and the dominant language (if different) will be assessed
- ☐ Translation of instruments/use of interpreters:
    - How will appropriateness of translation be determined?
    - What guidelines for selecting/using interpreters need to be followed?
    - Must norms for translated instruments be obtained?
    - Should any particular children be excluded because the language they speak is not very prevalent in a community and therefore obtaining a translation or interpreter would be too difficult?

4. How should families who are English language learners be assessed?

- ☐ Translation of instruments/use of interpreters:
  - • How will appropriateness of translation be determined?
  - • What guidelines for selecting/using interpreters need to be followed?
  - • Should any families be excluded because the language they speak is not very prevalent in a community and therefore, obtaining a translation or interpreter would be too difficult?

5. Will there be one method of child assessment or multiple options?

- ☐ Direct assessment
- ☐ Observation in natural environments
  - • Which natural environments
- ☐ Parent report
- ☐ Service provider report or rating
- ☐ IFSP or IEP review
- ☐ Others

6. Will there be one method of family assessment or multiple options?

- ☐ Family satisfaction with services
- ☐ Family perception of current status
- ☐ Family perception of progress
- ☐ Family report of capacity to enhance child's development
- ☐ Service provider's perception of family's progress in enhancing child's development
- ☐ Family's report/rating of experiences in direct service delivery
- ☐ Family's report of experiences with every stage of service delivery (child referral, intake, assessment, IFSP/IEP, services)
- ☐ Service provider report of services provided to families

7. How will system outcomes be assessed?

- ☐ Use of state database
- ☐ Perceptions of families
- ☐ Perceptions of service providers from lead agency only
- ☐ Perceptions of service providers from multiple agencies
- ☐ Outside evaluation (e.g., University)
- ☐ Monitoring

8. How will appropriate instruments be selected?

  ☐ Should all states use the same instrument?
  ☐ Should there be a list of approved instruments from which states can choose?
  ☐ Should any instrument currently being used by programs or states be eligible?

9. Should the measurement report "status" or "change"?

  ☐ Change will require more assessments.
  ☐ How much change is enough?
  ☐ Should change be reported in rate of acquisition of skills?
  ☐ What benchmarks should be set for change?
  ☐ What benchmarks should be set for measuring status?

10. Do the instruments selected have the following psychometric properties?

  ☐ Reliability
  ☐ Validity
  ☐ Utility
  ☐ Sensitivity
  ☐ Cultural bias

11. Who will perform the assessments?

  ☐ Qualifications?
  ☐ Protected tests?
  ☐ Specially trained assessors?
  ☐ Familiar adult?
  ☐ Impartial assessors or service providers?
  ☐ Cost of training large numbers of assessors and getting them reliable?

12. What are the issues around accommodations and alternate assessments?

  ☐ Should allowable accommodations be specified (standardized)?
  ☐ Should there be guidelines for who needs accommodations?
  ☐ If observing in natural environments, should we assume that accommodations are already being made?
  ☐ Is an alternate assessment necessary or can every child be assessed on the same instrument if we allow "out of level" assessment?

Reference List

Administration on Children, Youth, and Families, (1998). Head start program performance measures, second progress report. Washington, D.C.: Head Start Bureau, ACYF, U.S. Department of Health and Human Services.

Bailey, D., McWilliam, R., Darkes, L., Hebbeler, K., Simeonsson, R., Spiker, D., & Wagner, M. (1998). Family outcomes in early intervention: A framework for program evaluation and efficacy research. Exceptional Children, 64, 313-328.

Bailey, D. B., & Wolery, M. (1992). Teaching infants and preschoolers with disabilities. Columbus, OH: MacMillan.

Bruder, M.B. (In press). Service coordination and integration in a developmental systems approach to early intervention. In M. J. Guralnick (Ed.), A developmental systems approach to early intervention: National and international perspectives. Baltimore: Brookes.

Carta, J. J. (2002). An early childhood special education research agenda in a culture of accountability for results. Journal of Early Intervention, 25(2), 102-104.

Council on Quality Leadership, (2004). Personal outcome measures for families with young children. Towson, M.D.: The Council on Quality and Leadership.

Dunst, C. J., Jenkins, V., & Trivette, C. (1984). Family Support Scale: Reliability and validity. Journal of Individual, Family, and Community Wellness, 1(4), 45-52.

Dunst, C., Vance, S., & Cooper, C. (1986). A social systems perspective of adolescent pregnancy: Determinants of parent and parent-child behavior. Infant Mental Health Journal, 7, 33-48.

Early Childhood Research Institute on Measuring Growth and Development. <01 Author, Analytic>,(1998). Family outcomes in a growth and development model (Technical report number 7). Minneapolis: Center for Early Education and Development, The University of Minnesota.

Gallagher, J., Surles, R., & Hayes, A. (1973). Program planning and evaluation. Chapel Hill, NC: TADS.

Guralnick, M. J. (1997). The effectiveness of early intervention: Directions for second generation research. In M.J. Guralnick (Ed.), The effectiveness of early intervention (pp. 3-20). Baltimore: Paul H. Brookes.

Harbin, G. L. (1993). Family issues of children with disabilities: How research and theory have modified practices in intervention. In N. Anastasiow, & S. Harel (Eds.), At-risk infants: Interventions, families, and research (pp. 101-111). Baltimore: Paul H. Brooks.

Harbin, G.L. (2000). Implementing early intervention policy: Are we making the grade? Journal of Early Intervention, 24(2), 103-105.

Harbin, G.L., McWilliam, R. A., & Gallagher, J. J. (2000). Services for young children with disabilities and their families. In J.P. Shonkoff & J.P Meisels (Eds.), Handbook of early childhood intervention (2nd ed.), (pp. 387-415). New York: Cambridge University Press.

Harbin, G. L., & Neal, P. (2004). Family benefits inventory. Chapel Hill, NC: FPG Child Development Institute, The University of North Carolina at Chapel Hill.

Harbin, G.L., Pelosi, J., Kameny, R., McWilliam, Kitsul, Y., Fox. E., & Rodriguez, I. (2004). Identifying and predicting successful outcomes of coordinated service delivery. Chapel Hill, NC: FPG Child Development Institute, The University of North Carolina at Chapel Hill.

Hebbeler, K., Carta, J., & Baily, D. (2003). Center on Outcomes for Infants, Toddlers, and Preschoolers with Disabilities. Research, Innovation, and Evaluation: Discretionary Projects Supported by the Office of Special Education Programs under the Individuals with Disabilities Education Act.

Kahn, L. (2000). Personal communication.

Kibel, B. M. (1999). Success stories as hard data: An introduction to results mapping. New York: Kluwer Academic/Plenum Publishers.

Knapp, M. S. (1995). How shall we study comprehensive, collaborative services for children and families? Educational Researcher, 24(4), 5-16.

Lund, I.A., Rous, B.S., Moffett, C.L., Wood, K., & O'Keefe, H. (2002). State profiles on pre-kindergarten standards and accountability systems. Lexington, KY: State Preschool Accountability Research Collaborative, University of Kentucky Interdisciplinary Human Development Institute.

McConnell, S. R., McEvoy, M. A., & Priest, J. S. (2002). "Growing" measures for monitoring progress in early childhood education: A research and development process for individual growth and development factors. Assessment for Effective Intervention, 27(4), 3-14.

McConnell, S. R., Priest, J. S., Davis, S. D., & McEvoy, M. A. (2002). Best practices in measuring growth and development for preschool children. In A. Thomas & J. Grimes (Eds.), Best practices in school psychology (4th ed.), (pp. 1231-1246). Washington, D.C.: National Association of School Psychologists.

McLaughlin, M. J., Henderson, K., & Rhim, L. M. (1998). Snapshots of reform: How five local districts are interpreting standards-based reform for students with disabilities. Alexandria: VA: Center for Policy Research on the Impact of General and Special Education Reform, National Association of State Boards of Education.

McLean, M., Kuschner, A., Heal, H.B., & Drouin, C (2004, April). Inclusive assessment: Valid and feasible assessment of development for children with disabilities. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

National Commission on Excellence in Education. (1983). A nation at risk: the imperative for educational reform. Washington, D.C: United States Department of Education.

Neisworth, J.T., & Bagnayo, S.J. (2004). The mismeasure of young children: The authentic alternative. Infants and Young Children, 23(46, 198-212.

Park, J., Hoffman, L., Marquis, J., Turnbull, A., Poston, D., Mannan, H., Wang, M., & Nelson, L. (2003). Toward assessing family outcomes of service delivery: Validation of a family quality of life survey. Journal of Intellectual Disability Research, 47(4), 367-384.

Patton, M. Q. (1997). Utilization-focused evaluation (3rd ed.). Thousand Oaks: Sage.

Rhode Island Kids Count (2004). Getting ready: The school readiness indicators initiative. Available: http://www.gettingready.org/matriarch/default.asp

Rous, B., Kohner-Coogle, M., & Stewart, J. (2004). Links to early childhood standards. Lexington, KY: University of Kentucky State Preschool Accountability Research Collaborative

Roberts, R., Innocenti, M., & Goetze, L. (1999). Emerging issues from state level evaluations of early intervention programs. Journal of Early Intervention, 22(2), 152-63.

Shonkoff, J. P. (2002). A call to pour new wine into old bottles. Journal of Early Intervention, 25(2), 105-107.

Snyder, S., & Sheehan, R. (1992). The Rasch measurement model: An introduction. Journal of Early Intervention, 16(1), 87-95.

Stufflebeam, D. (1985). Systematic evaluation: A Self-instructional guide to theory and practice. Boston, MA: Kluwer/Nijhoff.

Thompson, S., & Thurlow, M. (2003). 2003 State special education outcomes: Marching on. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved 8/14/04, from the World Wide Web: http://education.umn.edu/NCEO/OnlinePubs/2003StateReport.htm./

Thurlow, M. L., Elliott, J. L., & Ysseldyke, J. E. (2003). <u>Testing students with disabilities: Practical strategies for complying with district and state requirements</u> (2nd Ed.). Thousand Oaks, CA: Corwin Press.

Turnbull, A., Marquis, J., Hoffman, L., Poston, D., Mannan, H., Wang, M., & Summers, J., (In press). <u>A new tool for assessing family outcomes: Psychometric evaluation of the Family Quality of Life Scale</u>.

U.S. Department of Education. (2001). <u>Elementary and Secondary Education Act of 2001 (No Child Left Behind)</u>. Washington, D.C.: U.S. Department of Education.

U.S. Department of Education. (2002). <u>No Child Left Behind: A desktop reference</u>. Washington, D.C.: U.S. Department of Education.

Weatherly, R. A. (1979). <u>Reforming special education: Policy implementation from state level to street level</u>. Cambridge, MA: MIT Press.

Wolery, M., & Bailey, D. B. (2002). Early childhood special education research. <u>Journal of Early Intervention, 25</u>(2), 88-99.

Acknowledgements