

Issue Brief



Education Policy Studies Division
Contact: Anna Lovejoy, 202/624-5331 or
alovejoy@nga.org
May 19, 1995

By the Bucket: Achieving Results for Young Children*

Summary

This *Issue Brief*, one in a series of National Governors' Association (NGA) publications emphasizing the needs of young children and what states can do to meet those needs, offers a conceptual framework for thinking about how states can begin the process of improving results for young children and their families by focusing on enhanced accountability. The series is a part of the Governors' Campaign for Children initiative of NGA Chair Howard Dean, M.D., of Vermont, a year-long effort to help policymakers better understand the needs of young children and develop strategies to address those needs.

The *Issue Brief* describes four categories, or "buckets," of data that can be used to measure results for children, families, services, and systems. The buckets, though independent and distinct, form a continuum. One end of the continuum represents the data items most directly related to children's performance and the other end represents data items more remote from children's performance but related to systemic performance. Together, the buckets represent an organized way to consider whether the nation is meeting the first National Education Goal: By the year 2000, all children will start school ready to learn.

Introduction

Profoundly concerned by the ineffectiveness of America's social institutions and motivated by a desire to improve their efficiency and cost-effectiveness, policymakers and citizens are calling for more stringent accountability. Although no institution is immune from such vociferous pleas, schools and agencies serving children and families seem to be at the forefront of the debate. In education, for example, calls for reform are rampant, with standards specification, outcomes-based accountability, and results-driven evaluation marshaling prominence in discussions of school-based management, charter schools, and comprehensive services. Simultaneously, concerns about the nation's younger children and their readiness for school have also generated heated debate: What is meant by readiness? Who is to be ready (children, schools, and/or communities)? How is readiness to be assessed? The subtext of many of the readiness conversations is fueled by concerns about performance and effectiveness, and it is logically

accompanied by calls for a sharper delineation of outcomes, specified accountability standards, and a results-driven orientation to services for young children.

Hardly harmonious, this press for an outcomes orientation for preschool-aged children has met with staunch vocal resistance among many early educators who argue that readiness is an ill-defined construct and that readiness and young children are not amenable to formal assessment. Moreover, early educators point to data that cogently and correctly depict misuses and abuses in the testing of preschoolers. Within the context of these legitimate concerns regarding testing, however, efforts to invoke various forms of child readiness assessment are rapidly mounting. Sometimes assessment efforts are motivated by well-meaning intentions to discern levels of child performance for placement purposes. Other times they are launched to improve pedagogical or service quality. Sometimes they have been inspired by community or state planning efforts to monitor service availability. Still other efforts take hold in response to accountability mandates.

Whatever their impetus, contemporary efforts to assess children's readiness are emerging amidst attitudinal skepticism, definitional ambiguity, and operational inconsistency—conditions that muddy intent and impede utility. On the definitional side, for example, “results” are often confused with “outcomes” and “interim indicators” are often confused with “benchmarks.” On the operational side, concurrent but distinct child, school, and community readiness assessments are being launched, sometimes with little coordination and sometimes with even less understanding of how such efforts might be synergistic. Such heightened activity, confused all the more by inconsistent definitions and purposes, demands that we take stock and closely examine what is meant by children's readiness for school and what data are needed to determine readiness. This *Issue Brief* attempts to do so by distinguishing among discrete categories of data that can be used to assess children's readiness (data buckets) and by suggesting how such buckets might be most effectively used, given the differing purposes of assessment.

Defining Results for Young Children

The terms “results” and “outcomes” often conjure up multiple images. “Results” is sometimes used as an umbrella term, signaling a movement from accountability based on *inputs* to accountability based on *accomplishments*. In this sense, “results” refer to a clear paradigm shift that substitutes an emphasis on process with an emphasis on effects. Sometimes this shift is discussed as a movement from a focus on means to a focus on ends, and from a focus on inputs to a focus on outcomes, where inputs are interim or enabling conditions that evoke outcomes. Here is where the discussion becomes slippery. First, though “outcomes” is sometimes used as a synonym for “results,” “outcomes” also conjures up a quite distinct meaning. The term “outcomes” presumes that a standard has been specified and that activity—teaching, service provision, and measurement—is directed, or redirected, toward that standard, sometimes at the expense of a concern with the inputs or the process. The movement toward outcomes-based education, though designed to promote academic excellence and equity, has been challenged on many counts—nationalization of standards and curriculum, excessive intrusion by the federal government into local affairs, and paradoxically, an emphasis on outcomes at the expense of process. To date, “results” is a more generic, if not accepted, term.

Beyond the semantic debate on results and outcomes, a second set of issues arises regarding the relative value placed on them. Some contend that inputs, rather than simply being a “means” to an “end,” are actually “ends” in themselves. They may actually constitute results. As such, they have sometimes been labeled “interim outcomes” and are regarded as quite suitable indicators of accomplishments. Although adopting a similar view, others avoid the term “interim outcomes,” substituting “benchmarks” as a more appropriate construct. A benchmark connotes a suitable reference point en route to a known position—a point on a continuum and an accomplishment.

Although others may elect to quibble with or to add to these distinctions, few would argue that for a long time, those concerned about children’s readiness have been caught in a seemingly endless nomenclature debate. This *Issue Brief* suggests that, though important, discerning between the nuanced labels is perhaps less significant than achieving clarity about what it is we wish to know. It simply uses the term “results” and suggests that there are four distinct buckets of results that are germane to children’s readiness. Each bucket is discernible and knowable, each demands its own data elements and approaches to data collection, and each evokes its own assessment processes and considerations. The buckets, though independent and distinct, can and should be used in concert with one another, depending on the defined purposes of the data collection. Together, the buckets form a continuum, with one end representing the data items most directly related to children’s performance and the other end representing the data items more remote from child performance but related to systemic performance. In concert, the buckets represent an organized way of approaching what should be assessed.

Bucket One: What Children Know and Can Do. This bucket contains information on what children know and can do. Such information must be gathered by observing children directly so that the data represent a precise reflection of children’s behavior and performance. Behaviors in this bucket include dimensions related to children’s motor development, their social and emotional development, their use of language, their cognition and general knowledge, and the way in which they approach learning.¹ The bucket focuses on all of these dimensions, with child behavior typically recorded intermittently over time from more than one data source. Examples of bucket one indicators by dimension include the following.

- **Motor development**—Jumps, walks a six-foot balance beam, cuts, and does an “x”-piece puzzle.
- **Social and emotional development**—Accepts responsibility for own actions, takes turns, and forms and maintains friendships with others.
- **Language usage**—Initiates and sustains conversations, listens to others, recites poems, repeats a sentence in correct word order, follows a verbal direction containing three steps, and tells about a picture when looking at it.
- **Cognition and general knowledge**—Matches, sorts multiple shapes and colors, and seriates by size.
- **Approaches toward learning**—Takes risks, is persistent, demonstrates curiosity, and uses materials in inventive ways.

Bucket Two: Child and Family Conditions. This bucket contains information regarding the conditions that surround and encase what children know and can do. Such information may be gathered from reviews of documents, including health records; interviews with family members and service providers; and direct observations and conversations with children and their families. The premise undergirding this bucket is that children’s knowledge and skills are directly related to their health status and to the conditions in which they live. Using information from individual children, this bucket generally reports data based on aggregated prevalences and percentages. Child and family conditions may be grouped into categories, for example, child health conditions or family income conditions, with positive and negative indicators in each category.

- **Child health conditions**—Prevalence of children born with low birthweights; prevalence of children who are fully immunized; prevalence of children who have functional limitations because of health conditions; prevalence of children who have age-appropriate heights and weights; and prevalence of children who are in good physical health, with no vision or hearing impairments.
- **Family income conditions**—Prevalence of children who live in poverty; prevalence of children who live above the poverty level; and prevalence of children with two parents, or one parent, employed.
- **Family life conditions**—Prevalence of children who are born to teen mothers or substance-using parents; prevalence of children who are abused; prevalence of children who live in foster care; prevalence of children whose television viewing is monitored; prevalence of children who live in two-parent families; and prevalence of children who live in comparatively crime-free neighborhoods.

Bucket Three: Service Provision and Access. A third bucket of readiness information characterizes the services that exist and those to which children and families have access. It is distinct from the behaviors (bucket one) or conditions (bucket two) associated with children’s readiness. More than a tally of raw services, this bucket focuses on actual access to services, with items typically reported in prevalences or percentages. Indicators often include access to services among specific populations or individuals with particular conditions (e.g., handicapped children, pregnant women, and unemployed mothers). Data for this bucket are typically collected from record reviews and community and institutional databases. Examples of the information in the provision and access bucket include the following.

- **Health provision and access**—Prevalence of pregnant women who have access to early and continuing prenatal care; prevalence of children with increased access to prenatal care; prevalence of children with health insurance; and prevalence of children who have access to regular vision and hearing screening, medical care, and well-child examinations.
- **Parenting education provision and access**—Prevalence of parents who have access to parenting classes and social supports.
- **Child care/preschool provision and access**—Prevalence of low-income, handicapped, or learning disabled children who have access to child care; prevalence of children who have

access to developmentally appropriate child care and education; and prevalence of children who have access to before- and after-school care.

- **Ready schools**—Prevalence of children who have access to schools that provide effective learning environments, engage parents, have high expectations for all children, and meaningfully acknowledge children’s individual, developmental, and cultural differences.

Bucket Four: Systems Capacity. Rather than focusing on the efficacy of service provision and access as in bucket three, bucket four suggests that attention be paid to the way services are linked and function as a system. This bucket assumes that service quality and efficiency are related to the quality and quantity of services, which, in turn, are related to children’s readiness. Far less well-developed than the other buckets, bucket four includes examinations of service redundancies, omissions, capacities, and efficiencies. Data for this bucket are collected in the aggregate and typically involve the integration of information across agencies and service providers. Examples include the following.

- **Systemic efficiency**—Degree to which the system is using its resources—fiscal, human, technical, and technological—efficiently and effectively.
- **Systemic infrastructure**—Degree to which the infrastructure (e.g., training, financing, and data gathering) functions to support efficient and effective service delivery.
- **Systemic accountability**—Degree to which accountability is dispersed across systems and degree to which agencies can build a collective accountability process.

The Buckets Together. Information regarding children’s readiness is not confined to any single bucket. The assessment of children’s readiness and society’s willingness to support assessment is reflected in the data contained in the four buckets. Indeed, if one surveys the results-oriented strategies underway in states and communities, and those proposed by scholars, all four buckets are represented in various ways. What does matter and what does differ, however, is why and how each bucket is represented. In some cases, an information item was tossed into the data collection because it was readily available, somebody always wanted to know “x” or “y,” or it made sense heuristically. Although these are not necessarily imprudent reasons, they need to be joined by a clear analysis of the purpose of the data collection. Once clarity of purpose is achieved, data from the appropriate bucket(s) can be considered. Data from all four buckets are important, but more important is the match between the purpose of the data collection and the data to be collected.

Matching Purpose and Data

Much has been written on the purposes of assessment and data collection, with Shepard suggesting that the “reason for any assessment affects the substance and the form of the assessment.”² She notes several critical purposes for assessing young children: identifying children with special needs, conducting classroom assessments, and monitoring national and state trends. Others have suggested additional reasons for moving to results-based accountability systems: replacing, or at least diminishing, the need for centralized bureaucratic micromanagement; assuring the funders and the public that investments are producing results;

facilitating cross-systems collaboration; and clarifying whether allocated resources are adequate to achieve the outcomes expected by the funders and the public.³ Still others note that assessments can establish baselines of information, inventory services, enhance understanding across cultural groups, and forge consensus.⁴

Although all of these purposes have merit and no doubt help account for the proliferation of results-driven assessment, this *Issue Brief* suggests four purposes for assessing children's readiness:

- individual child screening and placement in programs, schools, and classes;
- improvement of classroom pedagogy and program practice;
- inspiration of community awareness and activism on behalf of children; and
- local, state, or national accountability.

As the buckets form a continuum, so too do the purposes of data collection, with the first being the most germane to individual children and the last being most germane to governments. Although there is not a one-to-one correspondence between the buckets and the purposes, there is an alignment. For example, if one is most interested in assessing youngsters regarding their placement, then one would rely on bucket one indicators—those that are directly related to children's behavior and demonstrate what children know and can do. If one is more interested in improving classroom pedagogy and instructions, then a focus on buckets one and two is appropriate. If one is interested in the inspiration of community awareness and activism or on accountability, then a focus on all the buckets is appropriate, with perhaps greater emphasis on buckets two, three, and four.

As assessment efforts expand, all involved need to pinpoint the purpose of the data collection, noting the uses to which the data will be put. To streamline data collection and optimize efficiency, we need to be mindful that data from all buckets are not needed for all children for all purposes. By concentrating on the purpose(s), we can better define the nature of the relevant data, discerning the appropriate balance of items from each of the buckets.

Why Is This Important?

Early educators, psychologists, and many parents have resisted measurement over time, with varying rationales offered to explain such resistance. In part, resistance may emanate from a perception of pedagogical imprecision and distaste for accountability frequently attributed to the early childhood field. It may emerge from the field's dissatisfaction with the conversion of its historic commitment to individual observation into standardized tests that are routinely used to mislabel, miscategorize, and stigmatize children. Further, opponents of a results orientation are often disquieted by the "outcomes" being measured, believing that the narrow cognitive outcomes often tested bear little resemblance to broader developmental outcomes associated with later school and life success or to outcomes favored by early childhood educators. Others oppose a results orientation for the very young because they are offended by having to prove that services work; they suggest that children should be entitled to services and that a just society should

recognize this. Still others concerned about proving the utility of services are worried that if such services do not measure up to standards—often standards set by the dominant culture—the programs so badly needed by the poor might be jeopardized.

Such attitudinal skepticism is joined by methodological skepticism regarding the field's capacity to develop measures that meet scientific standards of validity and reliability, notably for the comprehensive domains advocated in bucket one. For example, it is comparatively easier to construct assessments in the cognitive domain than in the social and emotional domains. In addition, in all domains, there is concern about capturing elements related to cultural and contextual variation. However difficult the issues with regard to bucket one are, surely the other buckets are not immune to methodological problems. They face difficulties of nomenclature (e.g., the meaning of the term “home care”), of inconsistent data collection mechanisms and intervals, and of different intentions.

As the demand for readiness assessment gains momentum, definitional, attitudinal, and scientific hurdles remain underaddressed. Pressed by public demand, often too much attention is directed toward getting the job done, with too little time, thought, and dollars invested into creating appropriate processes and mechanisms that tackle these serious and pervasive issues. This *Issue Brief* has suggested one way to begin to do so.

- **Acknowledge multiple buckets.** We need to understand that different categories of results (buckets) coexist and that all contribute, albeit in different ways, to our understanding of children's readiness. We need to try to get beyond the input and output debate. We need to understand that where young children are concerned, results for systems (bucket four), for services (bucket three), and for families (bucket two) all impact results for children (bucket one).
- **Pinpoint the purposes of assessment.** We need to be more specific regarding the purposes of the assessments we want to create, with the understanding that the purpose will shape the nature, process, and the kind and use of data collected.
- **Engage practitioners and scientists.** We need to understand that attitudinal baggage and scientific/technical questions encase our work, demanding that time and resources be devoted to creating just and equitable assessment strategies and mechanisms. Practitioners and researchers need to be engaged in the process in order to capitalize on their knowledge and reduce historic skepticism regarding a results orientation.
- **Think inventively and strategically.** We need to capitalize on the momentum for change that exists by thinking inventively. We might consider setting up laboratories for invention within states, perhaps dividing up the work by purpose. We need to ensure that the commitment to wise and just assessments transcends the political tenure of a committed elected official. The profession needs to be involved, perhaps taking a lead on bucket one—a challenge that transcends locales—with communities and states engaged in assessments related to buckets two, three, and four.

Finally, we need to remember that assessing young children is not easy; assessing the conditions, services, and systems that foster their development is even harder. Despite the difficulty, all are necessary activities. Bucket by bucket, children's readiness for school must be assessed.

Endnotes

¹ The author is indebted to the National Education Goals Panel Goal I Technical Planning Group for its formulation of the domains of early learning and development as well as the protections it espouses to foster appropriate assessment.

² See Lorrie A. Shepard, "The Challenges of Assessing Young Children Appropriately," *Phi Delta Kappan* (November 1994).

³ See Schorr, Farrow, Hornbeck, and Watson, *The Case for Shifting to Results-Based Accountability* (Washington, D.C.: Improved Outcomes for Children Project, 1994).

⁴ See Bruner, Bell, Brindis, Chang, and Scarbrough, *Charting a Course: Assessing a Community's Strengths and Needs* (Falls Church, Va.: National Center for Service Integration, 1993).

*This issue brief was written by Sharon L. Kagan. **Editor's Note:** Sharon L. Kagan is the senior associate of the Bush Center in Child Development and Social Policy at Yale University. She is a national expert on early childhood and efforts needed to reform government services to better support young children and their families and serves as a consultant to the National Governor's Association on issues that impact children and families. The accountability issues discussed in this *Issue Brief* have been identified by the NGA State Management Task Force as a priority for 1995. The Governors' Campaign for Children, including the preparation of this *Issue Brief*, is supported by Carnegie Corporation of New York and the Foundation for Child Development.