Washington State

A Guide to Assessment in Early Childhood

Infancy to Age Eight

Office of Superintendent of Public Instruction (OSPI)

Dr. Terry Bergeson
State Superintendent of Public Instruction

Catherine Davidson, Ed. D
Chief of Staff

Martin T. Mueller
Assistant Superintendent, Student Support

Debra Williams-Appleton
Program Supervisor, Even Start Family Literacy and Early Childhood Education

Development of this guide is the result of a collaboration among:

Office of Superintendent for Public Instruction—Even Start Family Literacy, Early Childhood Education, Early Childhood Special Education, Title I, and Curriculum and Instruction Programs

Department of Early Learning

Department of Social and Health Services (DSHS)—Infant Toddler Early Intervention Program

Educational Service Districts—Early Childhood, Early Childhood Special Education, Curriculum and Instruction, and Migrant/Bilingual Programs.

Funding

Primary funding for this project is provided by the Office of Superintendent of Public Instruction's Even Start Family Literacy and Early Childhood Special Education Programs, the Department of Early Learning, and a generous contribution from Thrive By Five Washington.

©2008 Washington State Office of Superintendent of Public Instruction. This document may be freely reproduced without permission for non-profit, educational purposes. Use of or reference to this document should cite: *A Guide to Assessment in Early Childhood; Infancy to Age Eight. Washington State Office of Superintendent of Public Instruction, 2008.*

A Guide to Assessment in Early Childhood

Infancy to Age Eight

Partnership and Sponsorship

The *Guide to Assessment in Early Childhood* is sponsored by a partnership among the Office of Superintendent of Public Instruction (OSPI), the Department of Early Learning (DEL), and the nine regional Educational Service Districts (ESDs).

Prepared for

The Office of Superintendent of Public Instruction Early Childhood Assessment Workgroup

Authors

Part I: Assessment in Early Childhood

Kristine L. Slentz, Ph.D.
Western Washington University

Part II: Compendium of Assessment Tools

Diane M. Early, Ph.D.
Consultant, Education and Child Development Research

Margaret McKenna, Ph.D.
Northwest Institute for Children and Families (NICF)
University of Washington

Acknowledgements

Projects like this one come to completion only through the vision, expertise, and hard work of a large number of people. The author would like to expresses particular gratitude for the contributions of the following individuals:

Debra Williams-Appleton for her commitment to high quality assessment practices in early childhood education;

Molly Friedrichsen for her tireless, cheerful, efficient, and competent editorial and production assistance; and

Karen Walker for timely, careful, and thorough editing of multiple drafts of this document.

The Office of Superintendent of Public Instruction expresses sincere gratitude to the Early Childhood Assessment Workgroup for providing dedicated guidance and expertise. Members of the Workgroup (whose names appear below) also acted as contributing reviewers and editors.

Kelly Ferguson, Early Childhood Education and Reading, Educational Service District 189

Mary Fischer, Early Childhood Special Education and Readiness to Learn, Educational Service District 114

Carol Hall, Early Childhood Education and School Improvement, Educational Service District 112

Dave Irwin, Migrant/bilingual, Educational Service District 112

Leslie Keller, Department of Early Learning

Mary Perkins, Early Childhood Education, Educational Service District 113

Anne Renschler, Early Childhood Education, Title I, Office of the Superintendent of Public Instruction

Karen Walker, Early Childhood Special Education, Office of Superintendent of Public Instruction

Debra Williams-Appleton, Early Childhood Education and Even Start Family Literacy, Office of Superintendent of Public Instruction

Karen Woodsum, Infant Toddler Early Inter-vention, Department of Social & Health Services

Ed Yonamine, Early Childhood Education and Assistance Program

Additionally, many thanks go out to the other individuals named below who provided final review and input.

Keli Bohanon, Department of Early Learning

Lou Colwell, Special Education, Office of Superintendent of Public Instruction

Lexie Domaradski, Curriculum and Instruction, Office of Superintendent of Public Instruction

Keith Hyatt, Western Washington University

Cynthia Juarez, Childcare Licensing, Department of Early Learning

Joyce Kilmer, Early Childhood Education and Assistance Program, Department of Early Learning

Susan Sandall, University of Washington

Table of Contents

PAGE

Introduction 1

Purpose of the Guide 1

Background 3

Standards in Early Childhood Education 4

Goals 2000 4

Head Start Child Outcomes Framework 5

Good Start, Grow Smart 5

Washington's Early Learning and Development Benchmarks 6

Organization and Use of this Guide 7

Additional Readings and Resources 8

Part 1: Assessment in Early Childhood 11

Assessment = Getting to Know Children 11

Considerations in Early Childhood Assessment 12

Purposes of Assessment 14

Screening Assessment 15

Instructional Assessment 15

Diagnostic Assessment 15

Assessment for Program Evaluation/Accountability 15

Methods of Assessment 15

Types and Properties of Assessments 16

General Caveats 17

Figure 1: Assessment System 18

Additional Readings and Resources 19

Screening Assessments 21

What is Screening? 21

Purpose of Screening 21

Characteristics of Screening Assessments 22

Recommended Practices in Screening 23

State and Federal Requirements 24
School Readiness 25
Potential Screening Instruments 26
Sample Screening Instruments 26
Table 1: Characteristics of Screening Assessments 27
Additional Readings and Resources 28
Assessments to Inform and Monitor Instruction 29
What is Assessment to Inform and Monitor Instruction? 29
Purposes of Assessment to Inform and Monitor Instruction 30
Characteristics of Instructional Assessments 30
Developmental Assessments to Inform and Monitor Instruction 33
Assessments to Inform and Monitor Early Academic Instruction 33
Figure 2: Response to Intervention (RTI) Model 34
Figure 3: Recognition and Response Model 35
Recommended Practices in Assessment to Inform and Monitor Instruction 36
Methods for Collecting Ongoing Assessment Information 37
Using Instructional Assessment Effectively 38
Potential Instruments to Inform and Monitor Instruction 39
Sample Instructional Instruments 39
Table 2: Characteristics of Assessments to Inform and Monitor Instruction 40
Additional Readings and Resources 41
Diagnostic Assessments 43
What is Diagnostic Assessment? 43
Purpose of Diagnostic Assessments 44
Characteristics of Diagnostic Assessments 44
Eligibility Determination for Infants, Toddlers, and Preschoolers 44
Table 3: Washington State Eligibility Criteria for Infant Toddler Intervention and Special Education 45
Diagnosis of Early Academic Problems 47
Recommended Practices in Diagnostic Assessment 48

State and Federal Requirements 50
School Readiness 50
Potential Diagnostic Instruments 51
Sample Diagnostic Instruments 52
Table 4: Characteristics of Diagnostic Assessments 53
Additional Readings and Resources 54
Program Evaluation and Accountability Assessments 55
What is Program Evaluation and Accountability Assessment? 55
Purposes of Program Evaluation and Accountability Assessment 56
Characteristics of Program Evaluations and Accountability Assessments 57
Recommended Practices in Program Evaluation and Accountability Assessment 59
General Caveats 61
Potential Instruments for Program Evaluation 61
A Few Useful Instruments for Program Evaluation 62
Table 5: Characteristics of Accountability & Program Evaluation Assessments 63
Additional Readings and Resources 64
References for Part 1 64
Part II: Compendium of Assessment Instruments 69
Organization of the Compendium 69
How to Use the Compendium 73
How Assessment Tools Were Selected 74
A Note About School Readiness Assessments 74
Alphabetical Table of Tools with Summary Characteristics 75
Individual Tool Descriptions 89
Descriptions of Information on the Following Tables 89
Collection of Screening Tools 92
Ages & Stages Questionnaire, Second Edition [ASQ] (1999) 93
Ages & Stages Questionnaire: Social Emotional, Second Edition [ASQ:SE] (2002) 93
Battelle Developmental Inventory, Second Edition, Screening Test (2004) 94
Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III)Screening Test (2005) 94
Brief Infant Toddler Social Emotional Assessment [BITSEA] (2005) 95

Brigance Early Preschool Screen II (2005) 95
Brigance Infant & Toddler Screen (2002) 96
Brigance K & 1 Screen II (2005) 96
Brigance Preschool Screen II (2005) 97
Developmental Indicators for the Assessment of Learning, Third Edition [DIAL-3] (1998) 97
Developmental Observation Checklist System [DOCS] (1994) 98
Eyberg Child Behavior Inventory [ECBI] and Sutter-Eyberg Student Behavior Inventory-Revised [SESBI-R] (1999) 98
FirstSTEp: Screening Test for Evaluating Preschoolers (1993) 99
Fluharty Preschool Speech and Language Screening Test, Second Edition [Fluharty-2] (2001) 99
Infant Toddler Symptom Checklist [ITSC] (1995) 100
Learning Accomplishment Profile - Normed Screens [LAP-D Normed Screens] (1997) 100
Parents' Evaluation of Development Status [PEDS] (1997) 101
Pervasive Developmental Disorders Screening Test-II [PDDST-II] (2004) 101
Speed DIAL [Developmental Indicators for the Assessment of Learning] (1998) 102

Collection of Assessment Tools for Informing Instruction & Monitoring Progress 103

Assessment, Evaluation, and Programming System [AEPS] for Birth to Three Years, Second Edition (2002) 105
Assessment, Evaluation, and Programming System [AEPS] for Three to Six Years, Second Edition (2002) 105
Brigance Comprehensive Inventory of Basic Skills –Revised (CIBS-R)(1999) 106
Brigance Diagnostic Inventory of Early Development II [IED-II] (2004) 106
Carey Temperament Scales [CTS] (2000) 107
Carolina Curriculum for Infants and Toddlers with Special Needs [CCITSN], Third Edition (2004) 107
Carolina Curriculum for Preschoolers with Special Needs [CCPSN], Second Edition (2004) 108
Creative Curriculum Developmental Continuum for Ages Three to Five (2000) 108
Creative Curriculum Developmental Continuum for Infants, Toddlers & Twos (2006) 109
Developmental Reading Assessment, Second Edition [DRA 2] (2006) 109
Devereux Early Childhood Assessment [DECA] (1999) 110

Dynamic Indicators of Basic Early Literacy Skills, Sixth Edition [DIBELS-6] (2002) 110
ECLS-K Approaches to Learning Sub-Scale (1999) 111
Get It, Got It, Go (Preschool IGDIs)(2000) 111
Hawaii Early Learning Profile [HELP] (0-3 years)(2006) 112
Hawaii Early Learning Profile for Preschoolers [HELP for Preschoolers] (3-6 years)(1999) 112
High/Scope Child Observation Record for Infants and Toddlers [COR-IT] (2002) 113
High/Scope Preschool Child Observation Record [COR] (1999) 113
Individual Growth and Development Indicators [IGDIs] for Infants and Toddlers (2003) 114
Ounce Scale (2002) 114
Phonological Awareness and Literacy Screenings-Kindergarten [PALS-K] (2004) 115
Phonological Awareness and Literacy Screenings-PreK [PALS-PreK] (2004) 115
Phonological Awareness and Literacy Screenings-1-3 [PALS 1-3] (2003) 116
Qualls Early Learning Inventory [QELI] (2002) 116
Teacher Rating of Oral Language & Literacy [TROLL] (2001) 117
Transdisciplinary Play Based Assessment, Second Edition [TPBA 2] (2008) 117
Work Sampling System, [WSS] (1998) 118
Young Children's Achievement Test [YCAT] (2000) 118

Collection of Diagnostic Tools 119

Achenbach System of Empirically Based Assessment [ASEBA], Preschool Forms (2000) 121
Achenbach System of Empirically Based Assessment [ASEBA], School Aged Forms (2001) 122
Adaptive Behavior Assessment System, Second Edition [ABAS-II] (2003) 122
Arizona Articulation Proficiency Scale, Third Edition (Arizona-3)(2001) 123
Battelle Developmental Inventory, Second Edition [BDI-2] (2004) 123
Bayley Scales of Infant and Toddler Development, Third Edition [Bayley-III] (2005) 124
Behavior Assessment System for Children, Second Edition [BASC-2] (2004) 124
Behavioral and Emotional Rating Scale, Second Edition [BERS-2] (2004) 125
Bilingual Verbal Ability Test [BVAT] Normative Update (2005) 125
Boehm Test of Basic Concepts, Preschool, Third Edition (2001) 126

Boehm Test of Basic Concepts, Third Edition (2000) 126
Bracken Basic Concept Scale-Revised [BBCS-R] (1998) 127
Bruininks-Oseretsky Test of Motor Proficiency,
Second Edition [BOT-2] (2006) 127
Clinical Evaluation of Language Fundamentals,
Fourth Edition [CELF-4] (2003) 128
Clinical Evaluation of Language Fundamentals – Preschool,
Second Edition [CELF-P 2] (2004) 128
Communication and Symbolic Behavior Scale Developmental Profile,
First Normed Edition [CSBS DP] (2002) 129
Comprehensive Assessment of Spoken Language [CASL] (1999) 129
Comprehensive Test of Phonological Processing [CTOPP] (1999) 130
Conners Third Edition [Conners 3] (2008) 130
Conners Comprehensive Behavior Rating Scales [CBRS] (2008) 131
Developmental Assessment of Young Children [DAYC] (1998) 131
Differential Ability Scales – II [DAS-II] (2007) 132
Early Reading Diagnostic Assessment, Second Edition [ERDA 2] (2003) 132
Expressive One-Word Picture Vocabulary Test [EOWPVT] (2000) 133
Expressive Vocabulary Test, Second Edition [EVT-2] (2007) 133
Gilliam Autism Rating Scale, Second Edition [GARS-2] (2006) 134
Goldman-Fristoe Test of Articulation, Second Edition [GFTA-2] (2000) 134
Infant - Toddler Developmental Assessment [IDA] (1995) 135
Infant Toddler Sensory Profile (2002) 135
Infant Toddler Social and Emotional Assessment [ITSEA] (2005) 136
Iowa Test of Basic Skills [ITBS], Form A (2001), Form B (2003) 136
Kaufman Assessment Battery for Children, Second Edition [KABC-II] (2004) 137
Kaufman Brief Intelligence Test, Second Edition [KBIT-2] (2004) 137
Kaufman Survey of Early Academic and Language Skills [K-SEALS] (1993) 138
Kaufman Test of Educational Achievement, Second Edition [KTEA-II] (2004) 138
Khan-Lewis Phonological Analysis, Second Edition [KLPA-2] (2002) 139
Learning Accomplishment Profile - Diagnostic [LAP-D], Third Edition (2005) 139
Leiter International Performance Scale-Revised [LEITER-R] (1997) 140
Lindamood Auditory Conceptualization Test, Third Edition [LAC-3] (2004) 141
MacArthur-Bates Communicative Development Inventories,
Second Edition [CDI] (2006) 141

Merrill-Palmer-Revised Scales of Development [M-P-R] (2004) 142
Metropolitan Readiness Test [MRT6], Sixth Edition (1994) 142
Mullen Scales of Early Learning [MSEL] (1995) 143
OWLS: Listening Comprehension [LC] Scale,
Oral Expression [OE] Scale (1995) 143
OWLS: Written Expression Scale [WE] Scale (1996) 144
Peabody Developmental Motor Scales, Second Edition [PDMS-2] (2000) 144
Peabody Individual Achievement Test – Revised,
Normative Update [PIAT-R/NU] (1998) 145
Peabody Picture Vocabulary Test, Fourth Edition [PPVT-4] (2007) 145
Phonological Awareness Test 2 [PAT-2] (2007) 146
Pictorial Test of Intelligence, Second Edition [PTI-2] (2001) 146
Preschool and Kindergarten Behavior Scales,
Second Edition [PKBS-2] (2003) 147
Preschool Language Assessment Instrument, Second Edition [PLAI-2] (2003) 147
Preschool Language Scale, Fourth Edition [PLS-4] (2002) 148
Process Assessment of the Learner, Second Edition:
Diagnostic Assessment for Math [PAL-II Math] (2007) 148
Process Assessment of the Learner, Second Edition: Diagnostic
Assessment for Reading and Writing [PAL-II Reading and Writing] (2007) 149
Ready to Learn: A Dyslexia Screener (2004) 149
Receptive Expressive Emergent Language Scale,
Third Edition [REEL-3] (2003) 150
Receptive One-Word Picture Vocabulary Test [ROWPVT] (2000) 150
Scales of Independent Behavior - Revised [SIB-R] (1996) 151
Sensory Profile (1999) 151
Social Competence and Behavior Evaluation
[SCBE-Preschool Edition] (1995) 152
Stanford Achievement Test, Tenth Edition [SAT 10] (2003) 152
Stanford-Binet Intelligence Scales for Early Childhood,
Fifth Edition [Early SB5] (2005) 153
Temperament and Atypical Behavior Scale [TABS]
Screener and Assessment Tool (1999) 153
Test for Auditory Comprehension of Language, Third Edition [TACL-3] (1999) 154
Test of Early Language Development, Third Edition [TELD-3] (1999) 154
Test of Early Mathematics Ability, Third Edition [TEMA-3] (2003) 155
Test of Early Reading Ability, Third Edition [TERA-3] (2001) 155

Test of Language Development – Primary, Fourth Edition [TOLD-P:4] (2008) 156
Test of Phonological Awareness, Second Edition PLUS [TOPA-2+] (2004) 156
Test of Preschool Early Literacy (TOPEL)(2007) 157
Test of Word Reading Efficiency [TOWRE] (1999) 157
Toddler and Infant Motor Evaluation [TIME] (1994) 158
Vineland Adaptive Behavior Scales, Second Edition [Vineland-II] (2005) 158
Wechsler Individual Achievement Test, Second Edition [WIAT-II] (2001) 159
Wechsler Preschool and Primary Scale of Intelligence, Third Edition [WPPSI-III] (2002) 159
Woodcock-Johnson III Normative Update Complete [WJ III NU] (2006) 160
Collection of Program Evaluation Tools 161
Classroom Assessment Scoring System [CLASS] (2006) 161
Early Childhood Environment Rating Scale – Revised Edition [ECERS-R] (1998) 162
Early Language and Literacy Classroom Observation Scale [ELLCO] (2002) 162
Family Child Care Environment Rating Scale – Revised Edition [FCCERS-R] (2007) 163
Infant Toddler Environment Rating Scale - Revised [ITERS-R] (2003) 163
School-Age Care Environment Rating Scale [SACERS] (1995) 164
Assessment Tool Tables by Age Ranges
Assessment Tools for ChildrenAges Birth to 2 Years 11 Months 165
Assessment Tools for Children Ages 3 Years to 4 Years 11 Months 168
Assessment Tools for Children Ages 5 and Older 172
Information Resources for Part II 176
Appendix A: 179
Details Regarding High, Adequate, and Low Ratings of Reliability, Concurrent Validity, and Sensitivity/Specificity 179
Reliability 179
Concurrent Validity 179
Sensitivity/Specificity 180

Appendix B: 181

Position Paper on Building an Effective, Accountable System in Programs for Children Birth through Age 8, Early Childhood Curriculum, Assessment, and Program Evaluation 181

Introduction 181

The Position 181

Recommendations 182

Curriculum 182

Indicators of Effectiveness 182

Assessment of Young Children 183

Indicators of Effectiveness 184

Program Evaluation and Accountability 185

Indicators of Effectiveness 185

Creating Change Through Support for Programs 187

Introduction

The terms *assessment* and *early childhood* rarely appear together in everyday reading and conversation, so a guide on assessment in early childhood requires some explanation at the outset. Most people agree that early childhood includes the dynamic period from infancy until eight years of age, characterized by rapid and complex growth in physical, cognitive, and social domains. Assessment, by comparison, seems to be a rigid concept that conjures up visions of formalized testing inappropriate for young children.

Traditionally, assessment of young children was an expert practice limited to health and education specialists identifying special needs or gathering research data. Unfortunately, this approach has pretty much limited assessment to verifying and describing existing problems in early development instead of supporting optimal learning for all young children. This guide is designed to clarify the current role of early care and early education professionals in assessment of young children by providing background and context, practical guidance, recommendations, and resources.

The **Guide to Assessment in Early Childhood** is intended for primary use by those professionals in positions of program-level responsibility for developing comprehensive assessment plans, and selecting and administering assessment instruments. The document also can serve as a resource for designing professional development activities for program administrators and direct service staff responsible for gathering and interpreting assessment information.

Purpose of the Guide

The information provided in this guide is designed to be universally applicable in programs that serve young children with and without special needs, including English language learners, youngsters with economic and developmental risk factors, and those developing typically from birth to eight years of age.

Specific purposes of this guide are to:

1. Clarify terminology, vocabulary, and guiding principles related to assessment of all young children in early childhood programs;
2. Identify a framework for developing an assessment system that is grounded in standards, responsive to diversity among children, and connected directly to teaching and learning;
3. Describe in detail the specific types and purposes of assessment in early care and education settings, and associated best practices;
4. Provide suggestions for assessment instruments designed for screening, designing and monitoring instruction, diagnosis of special needs, and program evaluation; and
5. Provide a convenient and accessible companion listing of assessment instruments and their primary characteristics.

This guide is intended for use by early care and education program administrators, early childhood intervention specialists, early childhood special education providers, therapists, school psychologists, allied health professionals, and others who organize and coordinate assessment activities. Its contents cover the age range from birth to eight years and include all components of a comprehensive assessment system: screening, informing and monitoring instruction, diagnostic evaluation, and program evaluation/accountability.

A large number of instruments associated with each assessment purpose are described (separately in more detail and listed with defining characteristics) in **Part II** of this guide: **Compendium of Assessment Instruments.** Taken together, these materials should be a useful resource for all early childhood professionals, from those who want to learn basic information about an early childhood assessment system to those who have responsibilities for selecting tools and conducting assessments.

Background

Development and learning during the early years have become important issues in Washington State and nationwide. Professionals, parents, and policy makers alike have become increasingly interested in the quality of services and documentation of effectiveness in early childhood programs. Best practices in early care and education are currently characterized by coordinated systems of service delivery, with assessment as a central component along with curriculum and program evaluation (National Association for the Education of Young Children & National Association of Early Childhood Specialists/in State Departments of Education, 2003; National Association for the Education of Young Children & National Association of Early Childhood Specialists/in State Departments of Education, 2002).

Comprehensive systems of assessment in early childhood have been developed, predominantly in programs guided by state and federal regulations and program standards such as Head Start, state-funded preschool, infant toddler early intervention, early childhood special education, and other special services.

"The strongest effects of high-quality early childhood programs are found with children most at risk."

—Getting Ready (2006)

Historically, early care and education services have emphasized the contexts and processes of development for young children, with an associated focus on environments, interactions, materials, and activities. Recent initiatives at federal, state, and local levels have resulted in a marked increase in attention to measurable outcomes in terms of children's learning and behavior. *Increased attention to early learning outcomes, in turn, has increased the importance of assessment in early childhood programs.*

Assessment has become a pivotal component of all programs serving children ages birth to eight years, and an essential practice for all early childhood educators. Community-based and neighborhood early childhood programs provide services to children who are developing typically, those who may have specific or transitory problems in development, and those at serious risk for school failure. Regular screenings for academic problems and formative assessments of student progress in research-based core curricula are now considered critical components of high-quality instruction during primary grades. Some of the most skilled and experienced early childhood professionals, however, have had limited opportunities for professional development and training about selecting and using assessments and analyzing data about children's progress (Gettinger, 2001).

Washington State has a long history of providing services for young children and their families, training early childhood professionals, and supporting development of community resources. The state's commitment to its youngest citizens is reflected clearly in the delivery of federal programs such as the Child Care Development Fund, Head Start, and Even Start, as well as the Individuals with Disabilities Education Act (IDEA); Part C, Infant Toddler Early Intervention; and Part B, 619 Preschool Special Education programs.

In addition, the state has invested in initiatives such as the Early Childhood Educational Assistance Program (ECEAP) and a state requirement for preschool special education services in advance of the federal mandate. Washington reaffirmed a significant state level commitment to its youngest citizens with the creation of the Department of Early Learning (DEL) during the summer of 2006.

At that time, the Washington State Legislature passed legislation combining a number of existing state-level early childhood programs into a single agency, the Department of Early Learning. The new DEL is charged with coordinating and consolidating state activities related to child care and early learning programs, promoting linkages and alignment between and among the programs, and supporting the transition of children to kindergarten.

There are a number of different funding sources and administrative structures for programs serving young children and their families in the state (e.g., Child Care, Head Start/Early Head Start, Migrant and American Indian Head Start, Even Start, ECEAP, early intervention and special education, and primary school classrooms). Each program or service tends to have specific criteria for eligibility, targeted child outcomes, associated curricula, and required assessments. There are myriad specific assessment instruments in use by early educators, specialists, caregivers, and service providers.

Standards in Early Childhood Education

Early childhood programs in Washington State and nationwide are quickly evolving an emphasis on academic and social preparation of young children for success in school. The creation of academic learning standards began in the 1980s as a centerpiece of the school reform movement. Explicit and measurable standards for student learning outcomes were developed as an organizing framework for assessment and instruction—in content areas by national professional organizations and for K-12 curricula by state departments of education.

Washington State's initial standards framework, the Essential Academic Learning Requirements (EALRs), was developed in the early 1990s for reading, writing, math, science, and communication, with benchmarks at elementary, middle school, and secondary levels. Subsequently, Grade Level Expectations (GLEs) have been generated to provide more detailed benchmarks and resources for each K-12 grade level in a user friendly, on-line format. The EALRs and GLEs form a powerful infrastructure for early childhood educators to connect learning standards to assessment, curriculum, and instruction in the primary grades.

Goals 2000

The standards movement expanded to include preschool in 1994 with the passage of *Goals 2000: Educate America Act* legislation, as a means of addressing the first of eight national educational goals: "By the year 2000, all children in America will start school ready to learn."

Objectives for this ambitious goal included universal access to high quality, developmentally appropriate preschool programs; acknowledgement and support of parents as first teachers; and provision of basic prenatal, health, and nutrition services. The context of *Goals 2000* positioned preschool and primary education as a foundation for success in school, and was the genesis for the development of academic standards specific to Pre-K through grade three (Grisham-Brown, in press).

Head Start Child Outcomes Framework

A second major legislative catalyst for the development of standards at the preschool level was the reauthorization of Head Start legislation in 1998. A new requirement of the law was the creation of specific learning outcomes that preschoolers enrolled in Head Start would be expected to achieve prior to kindergarten entry. The result was the *Head Start Child Outcomes Framework* (Head Start Bureau, 2001), which includes standards for language, literacy, math and science, creative arts, social and emotional development, approaches to learning, and health/physical development.

The Head Start Child Outcomes Framework clearly expanded the developmental emphasis of preschool curricula to include early learning goals in academic content areas. Washington State followed suit in connecting infant, toddler, and preschool outcomes to academic content with the release of "Baby EALRs" in 2000: A *Framework for Achieving the Essential Academic Learning Requirements in Reading, Writing, and Communication: Birth to Five Years.*

Good Start, Grow Smart

An explicit focus on development of early learning standards was provided in 2002 in the form of a federal early childhood initiative known as *Good Start, Grow Smart.* This initiative was designed to more closely align preschool programs with the *No Child Left Behind* Act, and states were strongly encouraged to develop early learning standards as explicit expectations for what young children should know and be able to do at benchmarks from infancy through kindergarten. Most states already have or are rapidly developing early learning standards for infants, toddlers, and preschoolers, and these standards increasingly reflect academic areas aligned with K-12 curricula (NAEYC & NAECS/SDE, 2002). Recent research suggests that comprehensive curricula for preschool children include goals and objectives for emergent literacy and numeracy, along with the more conventional early motor, social, cognitive, and communication goals (Coleman, Buysse, & Neitzel, 2006).

The national context of standards and accountability for early learning outcomes has also resulted in a rapid and dramatic increase in specific outcome reporting requirements for early childhood programs. Head Start, Even Start, and similar state-funded early childhood programs now have performance standards that include aggregated reporting of child outcome data (Head Start Bureau, 2001). Preschool special education and infant toddler early intervention programs are also required to submit aggregated data on child outcomes to the Office of Special Education Programs (OSEP) starting in early 2007. Infant toddler early intervention programs also are now required to collect and report family outcome data (Hebbeler, Bailey, & Bruder, 2006).

Media attention to school readiness initiatives, kindergarten screening efforts, statewide testing against academic benchmarks, and state "report cards" in education have brought accountability squarely into the public consciousness. The emphasis on measurable outcomes has significant implications related to assessment aligned with standards, for every professional working in early childhood: "When assessment is for …accountability or to influence the curriculum, the assessment tool should be aligned with the curriculum as set forth in standards documents representing intended goals of instruction." (American Educational Research Association, 2004, pg. 2).

Washington's Early Learning and Development Benchmarks

In 2005, the Office of the Governor and the Office of Superintendent of Public Instruction (OSPI) entered into a partnership for the express purpose of creating, publishing, and disseminating *Washington's Early Learning and Development Benchmarks* (Kagan, Britto, Kauerz, & Tarrant, 2005), a "resource... to help guide children's development and learning" (pg. 1). The *Early Learning and Development Benchmarks* provide a set of general developmental expectations for what youngsters should know and be able to do at 18 months, 36 months, 60 months, and kindergarten entry, in five interrelated areas of development:

- physical/health/motor;
- social and emotional;
- approaches toward learning;
- cognition and general knowledge; and
- language, literacy, and communication.

These benchmarks are aligned with Head Start and K-12 standards, and provide a set of expectations that frame a general learning continuum for all young children (birth through kindergarten) in Washington State. They have been disseminated as a printed document and on line for use by parents and other adults who work with children at home, in childcare centers, or attending public or private preschool classrooms.

Having a concrete framework of standards for early learning and development, promotes continuity for children across early opportunities, and promotes consistency in selecting and measuring the child outcomes to be achieved across all programs in the state. Whether a program is serving toddlers or children in primary grades, children from low-income homes or those with identified disabilities, learning activities and child outcomes can be aligned with early learning standards that provide a consistent point of departure for curriculum development, instruction, and assessment.

Early learning standards tell us in a general sense what all youngsters should know and be able to do. The next step for early educators is to determine what to teach to whom (curriculum), and to measure whether or not children are learning and developing to expectations. A guide to assessment in early childhood can provide valuable information and resources to a variety of early childhood professionals.

Organization and Use of this Guide

Part I of this guide focuses on the various purposes of assessment, with an overview of assessment tools contained in the following section. Four primary purposes for gathering information on young children will be described in the overview:

- screening;
- informing instruction and monitoring progress;
- diagnosing special needs; and
- evaluating programs.

Subsequent sections provide details on recommended practices, procedures, and selected instruments for each purpose. Every section begins with questions relating to the type of assessment information being described, and ends with recommended instruments for the specific purpose.

The sections on screening, instruction, diagnosis, and program evaluation are designed to stand alone so that readers can find all the information on a specific purpose and associated practices, requirements, and appropriate instruments in one location. *The organization of the guide lends itself to selective use of the various sections, depending on which purposes of assessment are of interest.*

The sections on screening and instruction will be particularly useful for teachers; administrators might be more interested in the section on program evaluation and accountability; school psychologists and early learning assessment teams will probably find that the section on diagnostic assessment best matches their responsibilities. Some readers will want to focus on the narrative descriptions of assessment for a particular purpose; others will be searching for particular instruments.

Part II of this guide contains a compendium of assessment instruments that are designed for the purposes of screening, informing instruction, identifying children for further assessment, and evaluating programs. This section contains a table that lists numerous tools and describes important information about each, including, among other items, the primary purpose, domains assessed, publication and ordering information, and availability of technical data.

Additional Readings and Resources

A great deal of information has been summarized in this section, and there are many other resources that provide more detail on the background, contextual issues, and trends in early childhood assessment. A few of the more informative are those listed below.

Assessing young children: What policymakers need to know and do. Kagan, S., Scott-Little, C., & Clifford, R. (2003).

At the starting line: Early childhood education programs in the 50 states. American Federation of Teachers. (2002, December).

DEC recommended practices in early intervention/early childhood special education. Sandall, S., McLean, M., & Smith, B. (2000). Longmont, CO: Sopris West.

DEC recommended practices program assessment: Improving practices for young children with special needs and their families. Hemmeter, M., Joseph, G., Smith, B., & Sandall, S. (Eds.). (2001).

Developmentally appropriate practice in early childhood programs. Bredekamp, S. & Copple, C. (1997).

Eager to learn: Educating our preschoolers. Bowman, B., Donovan, S., & Burns, S. (Eds.). (2001).

Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8. National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education. (2003).

Early learning standards: Creating the conditions for success. National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education. (2002).

Framework for Achieving the Essential Academic Learning Requirements in Reading, Writing, and Communication: Birth to Five. Online: http://www.k12.wa.us/CurriculumInstruct/pubdocs/birth-to-5.pdf

From neurons to neighborhoods: The science of early childhood development. National Research Council & Institute of Medicine. (2002).

Head Start child outcomes framework. Head Start Bureau. (2001). Online: http://www.headstartinfo.org/pdf/im00_18a.pdf. National Child Care Information Center (2005, June).

Assessment and evaluation: Becoming an educated consumer. Part I: Child assessment. Retrieved January 26, 2008, from http://nccic.acf.hhs.gov/pubs/goodstart/assess-eval1.pdf. National Child Care Information Center (2005, June).

Assessment and evaluation: Becoming an educated consumer. Part II: Program evaluation. Retrieved January 26, 2008, from http://nccic.acf.hhs.gov/pubs/goodstart/assess-eval2.pdf. National Child Care Information Center (2005, June).

Assessment and evaluation: Becoming an educated consumer. Part III: Accountability systems. Retrieved January 26, 2008, from http://nccic.acf.hhs.gov/pubs/goodstart/assess-eval3.pdf

Recognition and response: An early intervening system for young children at risk for learning disabilities. Executive summary. Coleman, M. R., Buysse, V., & Neitzel, J. (2006). Chapel Hill: The University of North Carolina at Chapel Hill, FPG Child Development Institute. Online: http://www.recognitionandresponse.org/

Standards in early childhood education. Grisham-Brown, J. (in press).

Washington State early learning and development benchmarks. Kagan, S. L., Britts, P. R., Kauerz, K., & Tarrant, K. (2005). Online: http://www.k12.wa.us/EarlyLearning/Benchmarks.aspx

Because early development is complex and dynamic, the most effective assessment procedures are characterized by a combination of methods and sources of information.

Part I: Assessment in Early Childhood

The word *assessment* derives from a concept that indicates assignment of value to objects (as in a tax assessment), and has a variety of meanings for parents and professionals. Many young children are assessed only during well-child visits; others have thick files containing medical and developmental assessment reports from birth. Some early childhood professionals have devoted considerable energy to assuring that young children are spared inappropriate testing experiences; others spend a significant portion of their professional lives conducting and interpreting assessments. Assessment in early childhood can be confusing because we assess children in many different ways for many different purposes, using literally hundreds of different instruments.

Assessment = Getting to Know Children

Assessment is often used as a synonym for testing—not an appealing notion when construed as a young child sitting in silence with a paper and pencil. Early childhood has historically used informal assessments in the form of naturalistic observations and anecdotal records. Current recommendations from the field and professional literature indicate the need for assessment *systems* that use ongoing, multiple methods for gathering information (Shepard, Kagan, & Wurtz, 1998; NAEYC & NAECS/SDE, 2003). This guide will present assessment as a continuous *process* that is an integral aspect of teaching and learning, with an emphasis on the *specific purposes* for which assessments are conducted.

In one way or another, all early childhood assessments involve a process of gathering information about children in an attempt to better understand and support learning and development. It is more accurate to say that we assess young children's behaviors, skills, competencies, preferences, and interactions than to say that we assess children. Assessment results can describe some informative details of what youngsters know and can do, but can never fully portray who they are as individuals. Meaningful assessment involves thoughtful choices on the part of professionals among the many purposes, types, methods, and instruments available to assist us in getting to know more about young children.

The compendium of assessment instruments included in **Part II** of this guide addresses the following areas of development and learning. Among the most important:

- approaches to learning;
- physical/motor;
- social/emotional;
- cognitive;
- communication and language;
- literacy;
- math/numeracy;
- sensory function;
- temperament;
- behavior; and
- specific disabilities such as autism.

The instruments are intended as tools for gathering information, not as meaningful stand-alone products. Think of the actual assessment resources as analogous to the internet browsers used to find information on the computer. An assessment instrument, like a browser, is only as good as the information it yields. Assessment tools provide a structure for accessing and organizing information about early learning and development, but knowing which instrument to select can be confusing and complex. This section is designed to provide a foundation for understanding the purposes, types, and methods of assessment as a framework for developing a system that meets the needs of the children you serve.

Considerations in Early Childhood Assessment

Early childhood professionals are feeling ever-increasing pressure to document learning outcomes in an era of standards, accountability, and achievement testing. In addition to parents, we are the people responsible for the well-being of young children. Many professionals have legitimate concerns about misuse of assessment practices and instruments, and the potential for inequitable consequences for the children in our programs. Before discussing the various purposes, types, and methods of assessment, it is important to consider some implications of the unique nature of early development and learning:

1. Complete and meaningful assessment in early childhood necessitates an understanding of family context, including getting to know family language and culture, gathering developmental information from parents, and conducting home visits with parent approval. This principle applies to all youngsters and families, but is especially critical for children whose families may not share the language or some of the economic advantages of the dominant culture. Understanding family expectations and experience places a child's behavior in context and can prevent harmful decisions that result from misinterpretation of assessment data (NAEYC, 2005).

2. Younger children present some complex challenges and require flexible procedures for gathering meaningful and useful assessment information. Constitutional variables such as fatigue, hunger, illness, and temperament can easily overshadow the abilities of a young child. Time of day, setting, testing materials and other situational factors also affect performance. The younger a child, the more likely he or she is to fall asleep, become distressed, refuse to comply with directions, or be distracted from assessment activities. Professionals should be prepared to modify activities, explore alternative procedures, and/or reschedule rather than risk gathering faulty information that compromises assessment results.

3. Young children learn by doing, and demonstrate knowledge and skills through action-oriented activities. Authentic assessment of youngsters as they participate in daily activities, routines, and interactions generally produces the most valuable information for assessment. To the extent possible, assessment methods should allow for observation of young children engaged in spontaneous behaviors in familiar settings and with familiar people.

4. More assessments and increased data do not necessarily result in better assessment information. Early childhood professionals should only gather information they need, and know ahead of time how they will use all the information collected. It is generally most desirable to identify a set of appropriate methods and instruments that provide necessary information, and refine the use of those procedures over time.

5. Some assessment instruments and procedures are better than others. Factors such as purpose, content, reliability and validity, efficiency, cost, and availability of professional development are all more important than appealing packaging and effective advertising. Of primary importance is the quality of information gathered and the decisions made as a result of assessment. Ultimately, whatever assessments we conduct should benefit the children, families, and programs we serve.

Assessment tools provide a structure for accessing and organizing information about early learning and development, but knowing which instrument to select can be confusing and complex.

Purposes of Assessment

By now it is obvious that assessment in early childhood is a complicated and multifaceted enterprise. Recommendations from the field and the professional literature indicate that early care and education programs should incorporate into their services, coherent *systems* of assessment organized to address the following purposes:

- **Screening** - To identify potential problems in development; ensure development is on target.
- **Instructional** - To inform, support, and monitor learning.
- **Diagnostic** - To diagnose strengths and areas of need to support development, instruction, and/or behavior. To diagnose the severity and nature of special needs, and establish program eligibility.
- **Program Evaluation/Accountability** - To evaluate programs and provide accountability data on program outcomes for the purpose of program improvement.

These four purposes are consistently described in early childhood professional literature with some variation in organization.

- The National Education Goals Panel (Shepard, Kagan, & Wurtz, 1998), for example, combines screening and diagnostic assessment into a single category for identification of special needs, and separates program evaluation and accountability purposes.
- The National Association for Education of Young Children (NAEYC), in a joint statement with the National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE, 2003), combines assessment for instructional purposes with screening in a single section on assessment, and addresses program evaluation and accountability in another section.
- Neisworth and Bagnato (2004) separate instructional program planning and monitoring, and combine screening and diagnostic assessments.
- Coleman, Buysse, and Neitzel (2006) include screening as a primary element of the Recognition and Response system model. Recognition and Response is a conceptual model for preschool prevention and early academic interventions adapted from the Response to Intervention (RtI) model currently being implemented in schools. Assessment to inform and monitor instruction is a key aspect of increasingly intensive interventions during preschool to prevent learning disabilities later on.

Each purpose of assessment provides different levels of information to answer specific questions about early learning and development. We will describe each of the four purposes listed above more completely in subsequent sections of the guide, providing just a brief overview of each purpose here. At the end of this section is a graphic representation of an assessment system that includes screening, instructional, and diagnostic assessments.

Screening Assessment

Screening is a process designed for the purpose of identifying potential problems in learning or development. Screening instruments are quickly and easily administered to identify children who need more extensive assessment. Screening is a vital assessment activity in almost all early childhood programs because positive developmental and academic outcomes are associated with early identification of and attention to problems.

Instructional Assessment

The most important reason for assessment in early childhood is to support early learning and development. This level of assessment yields information about what children know and are able to do at a given point in time, guides "next steps" in learning, and provides feedback on progress toward goals. Assessment to support instruction is a continuous process that is directly linked to curriculum. Instructional assessments are aligned directly with curriculum goals which, in turn, are aligned with early childhood standards such as the *Washington Early Learning and Development Benchmarks.*

Diagnostic Assessment

Diagnostic assessment is a thorough and comprehensive assessment of early development and/or learning for the purpose of identifying specific learning difficulties and delays, disabilities, and specific skill deficits, as well as evaluating eligibility for additional support services, Infant Toddler early intervention, and special education. Diagnostic assessments usually are conducted by trained professionals using specific tests. When used to determine eligibility for specific support services, early intervention and special education, diagnostic assessment is a formal procedure governed by federal and state law.

Assessment for Program Evaluation/Accountability

Assessment for program evaluation and accountability addresses program-level outcomes. While instructional, screening, and diagnostic assessments address the development and learning of individual children, program evaluation and accountability assessments focus on the performance of *groups* of children. In addition, program evaluations routinely address variables, such as the quality of personnel and the physical environment, effectiveness of parent involvement, and community collaboration activities. Accountability assessment is often required by external agencies and used by policy makers to make decisions about funding, needed program supports, and program requirements.

Methods of Assessment

Methods of assessment refer to the procedures used to collect information and should be matched to the purpose for which information is being collected. *As a general rule, more formal methods and procedures are used for higher stakes decisions.* It is easy to think of assessment only as testing, but the use of formal tests with young children is generally considered inappropriate except for purposes of identifying disabilities, establishing eligibility, and documenting program accountability (NAEYC & NAECS/SDE, 2003).

Neisworth & Bagnato (2004) make a persuasive argument that "conventional testing must be abandoned within early childhood for every purpose including screening, eligibility determination, program planning, progress monitoring, and notably, program evaluation outcomes research" (pg. 199). Their indictment covers every purpose for assessment described above, and they recommend authentic alternatives for all forms of early childhood assessment instead.

Because early development is complex and dynamic, the most effective assessment procedures are characterized by a combination of methods and sources of information. Developmental checklists, rating scales, caregiver interviews, and portfolios of children's work are useful methods for gathering information about progress toward developmental goals. Increasingly, new assessment instruments incorporate caregiver reports and observation of specific developmental and content standards, and some states now allow use of authentic, observational instruments for eligibility determination.

Specific methods of collecting assessment information relevant to curriculum planning and monitoring child progress are described in the section on *Assessment to Inform and Monitor Instruction*.

Types and Properties of Assessments

The interpretation of assessment data is dependent on the purpose and method of the assessment, and the type of data collected. Assessment tools are generally designed for a single purpose and it does a great disservice to children to use them otherwise. Screening instruments, for example, cannot be used to inform or monitor instruction, or to qualify a child for Infant Toddler early intervention or special education services. Program accountability data provide little, if any, information useful for planning individual instruction. The brief review of the various types of assessments that follows is designed to provide background information, introduce terminology, and illustrate some potential areas of misuse.

Standardized assessment involves a predetermined set of assessment items that represent "standards" of knowledge and/or skills. Standardized tests may be norm or criterion referenced, and items are presented to all children in the same sequence, using the same administration procedures and materials. Scoring and interpretation of performance is also standardized. Scores on standardized tests can unfairly penalize specific groups of children, such as English language learners and youngsters with receptive language delays who do not understand verbal directions. Children with sensory and physical impairments may not be able to demonstrate skills if the standard set of materials cannot be modified to accommodate their responses.

Norm-referenced assessments compare a child's score to the scores of a group of same-age peers (norm group). Such a comparison is only meaningful if the norm group includes children who share the language, culture, and/or (dis)abilities of those being assessed. Norm-referenced assessments yield numerical scores that can underestimate the performance of young children with disabilities, those learning English, and those whose early experience differs significantly from the "norm". Norm-referenced tests are almost always standardized to preserve a consistent basis for comparison of scores.

Criterion-referenced assessments measure a child's performance against a predetermined set of criteria, generally developmentally sequenced or task analyzed skills. Criterion-referenced measures yield performance profiles and numerical scores that reflect the number of skills mastered. These instruments

may be standardized, as in the case of oral reading fluency timings in primary grades, but for developmental content usually allow flexibility in administration procedures and assessment materials.

Curriculum-referenced assessments are criterion-referenced instruments that are packaged with an aligned set of curriculum goals. Curriculum-based assessment serves to place children in a curriculum sequence and the same items are used to monitor progress toward learning objectives. These assessments often provide a logical teaching sequence, and may also include instructional activities.

Readiness assessments are tests that gather information to determine how well a child is prepared for a specific program. In early childhood, readiness assessments are most frequently used (some would say *mis*used) at kindergarten entry. Readiness assessments become problematic when the results are used to exclude children from programs rather than to identify areas where extra support is needed.

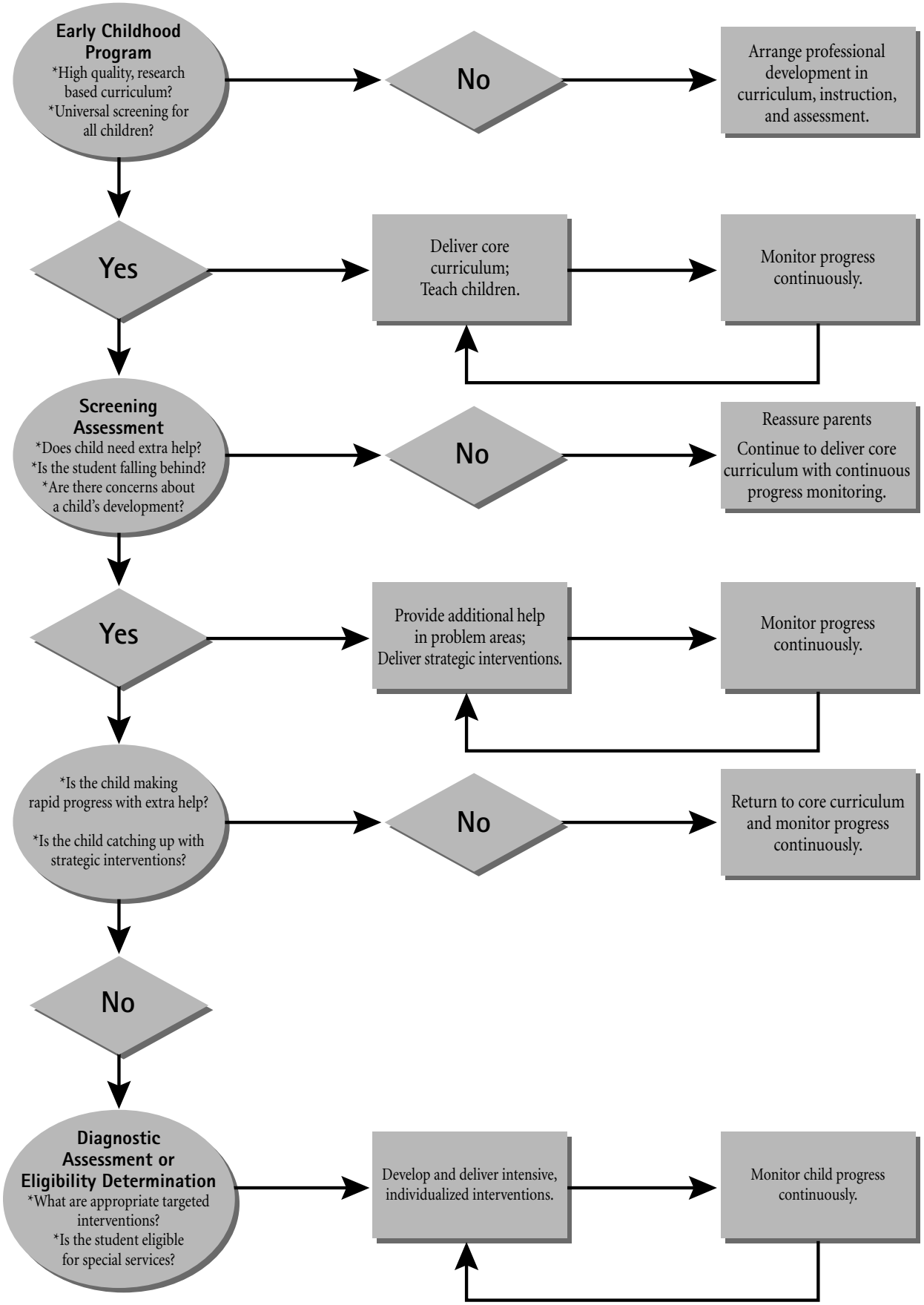
The following three terms are also important to understand and consider when selecting and using any assessment instrument:

- **Reliability** refers to the accuracy and stability of assessment scores. Every assessment contains some degree of error (in administration, scoring, interpretation) and error decreases accuracy of scores. Assessment developers ensure reliability by testing the same children twice, by having multiple people score the same child, and by statistical analysis of items.
- **Validity** is an indication of how closely the assessment measures what it is intended to measure. An assessment has to be reliable in order to be valid. Assessment developers make logical hypotheses and analyze groups of test scores to see if the hypotheses hold. Logical hypotheses for a developmental test are that older children score higher than younger children, and children with identified physical and language disabilities score lower in those domains. A screening instrument demonstrates validity if children who are identified by screening to have a problem also receive low scores on a comprehensive test of development.
- **Technical adequacy** describes the degree of demonstrated reliability and validity of a test. Technical information is often included in the assessment guide. Technical adequacy is an important consideration when selecting assessment instruments for any purpose, although norm-referenced assessments generally have more information on reliability and validity than do criterion-referenced instruments.

General Caveats

Any criticisms that can be leveled at early childhood assessments are likely to be magnified when young children from diverse language groups and cultural backgrounds are being assessed. NAEYC (2005) describes a number of problems and recommends practices to ensure fair and accurate assessment of young English-language Learners (ELL). Their recommendations are grounded in good assessment practice, and emphasize the "alignment of assessment tools and procedures with the specific cultural and linguistic characteristics of the children being assessed" (pg. 8). It is obvious that early childhood professionals must take particular care when developing assessment procedures to ensure practices that are relevant and responsive for children from all backgrounds, language groups, and cultures.

Figure 1
Assessment System



In addition, the entire endeavor of assessment carries a certain judgmental connotation that can be especially threatening to parents. Assessment results are often presented as numbers and youngsters are routinely described as "being" high or low scorers. "You should never assume that scores based on even the most thorough measurement procedure can actually summarize the total person." (Howell & Nolet, 2000, pg. 106). It should go without saying that assessment results never indicate the value of a child; a test score is only a number and not a determination of a child's worth.

Additional Readings and Resources

Eager to learn: Educating our preschoolers, Bowman, B., Donovan, S., & Burns, S. (Eds.). (2001).

Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8. National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education (2003). Joint position statement. Washington, DC: National Association for the Education of Young Children.

Principles and Recommendations for Early Childhood Assessments. Shepard, L., Kagan, S., & Wurtz, E. (1998). Washington, D.C.: National Educational Goals Panel. Online: http://www.negp.gov

Recognition and response: An early intervening system for young children at risk for learning disabilities. Executive summary. Coleman, M. R., Buysse, V., & Neitzel, J. (2006). Online: http://www.recognitionandresponse.org/

Screening and Assessment of Young English-Language Learners: Draft Recommendations. NAEYC and the National Association of Early Childhood Specialists in State Departments of Education (January 2005). Joint position statement in supplement to *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8.*

The mismeasure of young children: The authentic assessment alternative. Neisworth, J. T., & Bagnato, S. J. (2004).

The words we use: A glossary of terms for early childhood education standards and assessments. Council of Chief State School Officers (2004). Retrieved December 14, 2006. http://www.ccsso.org/eceaglossary

Screening Assessments

- My other children crawled and walked much earlier. **Is there something wrong** with my youngest son?
- **Should I be worried** because one of the toddlers isn't talking like the other kids?
- One little girl in the preschool is so often sick. **Is her development on track?**
- **Is my child ready** for kindergarten?
- **Is my child learning to read?**

Parents, family members, teachers, and other caring adults often have questions about the development of young children they know and care for. Pediatricians routinely document the developmental status of large numbers of youngsters during well-child visits. Parents and early childhood professionals may suspect developmental problems but not be able to provide specific descriptions. Kindergarten and primary teachers need to know which students may not be proficient in the understanding and use of the English language, or are struggling in academic areas. Comprehensive developmental and academic assessments are expensive and time-consuming, however, and considered too intensive to address general questions about learning and development. Screening assessments are the best choice for an initial look at a child's learning or development, to document typical development and identify youngsters who might be delayed in academic or developmental areas.

What is Screening?

Screening is a very general type of assessment that addresses common questions parents and professionals have about the development of young children. Screening assessments are designed to efficiently identify those youngsters who need more thorough and detailed assessment. Such screening is ideally brief and cost-effective so that large numbers of children can be assessed in a relatively short period of time. The procedures and tests used in screening are developed to be quickly and easily administered without highly specialized training.

Some common examples of screening activities are child-find clinics in the community, kindergarten screening clinics at schools, and the home language survey completed for all English Language Learners at school registration. Screenings for problems in learning reading and math are becoming commonplace in primary school classrooms. In addition to such broad-based efforts to identify developmental problems among large groups of children, child care programs are being encouraged to conduct periodic screenings of all children served. Chances are good that if you are reading this guide, you have already been involved in community-based and/or program-based screening assessment activities.

Purpose of Screening

The ultimate purpose of screening in early childhood is rapid assessment of large groups of children to identify those who need more in-depth assessment of special needs. It is well-documented in educational

and medical professional literature that developmental outcomes for young children with delays and disabilities are improved with early identification and intervention (Squires, Nickel, & Eisert, 1996; Shonkoff & Meisels, 2000).

Comprehensive, in-depth assessment of any sort is generally expensive, time-intensive, and requires the expertise of specialists, however, and the majority of young children will demonstrate typical growth and learning patterns. Screening permits a quick assessment of many children, and systematically limits more extensive and expensive assessment to those few who are likely to need it most (Meisels & Fenichel, 1996).

Characteristics of Screening Assessments

Questions like the ones at the beginning of this section are not very specific, so the information gathered during screening assessments is relatively general, also. The idea behind screening is to assess large groups of children quickly, to identify the few who will benefit from a more comprehensive and thorough assessment in the areas where problems are identified. Screening tests have relatively few items, so information is gathered only on major indicators of development and learning.

Comprehensive developmental screening instruments are generally norm-referenced, comparing a child's performance in physical, social, cognitive, and communication domains to a group of same age peers. Many kindergarten programs have criterion-referenced lists of skills to screen children's pre-academic knowledge, such as recognition of numerals and letters. Vision and hearing screens are also criterion-referenced, with a standard criterion such as 20/20 vision used as a basis for comparison. In primary grades children are routinely screened in reading and math to identify those who are falling behind grade level expectations and need additional instruction to catch up.

Screening assessment results are only a sample of academic or developmental level at a point in time. Screening tests generally have a cut-off score, with scores below the cut-off indicating lower than expected performance. For example, a toddler might score above the cut-off score in motor and social domains, but fall below in cognitive and communication, indicating a potential problem in language development. A kindergarten teacher can use screening results to identify those children who need extra help learning letters and numbers. Students who score below the cut-off point for grade level expectations in reading are referred for more intensive assessment and instruction.

Screening is only the first step in answering questions about children's development and learning. The results of screening assessments are best used to sort children into groups or categories relative to cut-off scores. Because screening results are based on a minimum of detail, results can only indicate, for example:

1. "Yes, development is on track"; "Your child's hearing is fine"; "He is reading at grade level" (scores above cut-off)
2. "No, language development seems delayed"; "Your child seems to have a problem hearing"; "She could use some extra help in reading" (scores below cut-off)
3. "Maybe; it's unclear whether there might be a problem"; "We aren't sure what your child is hearing"; "He is almost at expected levels in math" (scores right at cut-off, scattered, unclear)

The quality of a screening instrument depends on the extent to which it sorts groups of children accurately, represented by the technical terms *sensitivity* and *specificity*.

- *Sensitivity* means that a screening test is sensitive enough to identify those children whose development is at-risk or progressing below expectations. A screening test that is 100% sensitive will not miss any children who are having developmental or learning problems.
- *Specificity* refers to how selective the test is at identifying *only* those children whose performance is truly problematic. A test that is 100% specific would never indicate problems for children who are developing typically.

As you might expect, there are no tests that are 100% accurate, but information about sensitivity and specificity should be reported for screening tests, and is an important consideration in selecting screening instruments. *See the end of this section for a table that summarizes characteristics of screening instruments.*

Recommended Practices in Screening

One major appeal of screening instruments is that they are short and easy to administer. When conducted properly, however, screening is a set of procedures rather than simply a test. High-quality screening practices improve the validity of results, and ensure that instruments are used consistently with the purposes of screening assessment:

1. **Screening procedures should include multiple sources of information, with special attention to the family's perspective in gathering information and reviewing results.** Screening is never conducted as an end in itself, but rather to document those youngsters whose development is on track, and to identify concerns. The implication for practice is thoughtful attention to combining screening tests with observation and interview of parents, and reviewing results in light of family input and feedback.
2. **The best screening procedures have predetermined decision rules to guide follow up of results.** Screening procedures should always include referral and follow-up guidelines. For example: When scores indicate typical or normal development, parents are reassured and the next screening is scheduled for 6 months – one year. When scores suggest a concern, children are referred for more extensive assessment and parents receive information and counseling about the process. When scores are equivocal or borderline, parents are notified of results and the next screening is scheduled in three months.
3. **Screening results should only be used for the purpose they are developed: to identify children who will benefit from further assessment.** It is no doubt clear by this point that the results of screening tests are too broad and general to be used to diagnose or label a child, to determine eligibility, or to identify instructional goals. And there is obviously little sense in conducting screenings for children who have a diagnosed disability or delay, nor for those who have already been determined to be eligible for special services.
4. **Developmental screening instruments should be norm-referenced; sensory and early academic screenings that are criterion-referenced should have explicit standards for comparison.** All screening instruments should be standardized in their administration and scoring. For a meaning-

ful comparison of individual scores to a norm or pre-set criteria, all youngsters must have the same experience during the screening assessment.

5. **Screening instruments must have data available to document reliability and validity, as well as sensitivity and specificity.** Data about the technical properties of the test are necessary for professionals to have confidence in the results of screening. Identification of special needs is a high-stakes endeavor, and professionals need to know that the results of screening assessments are accurate and meaningful.

6. **Screening procedures must be culturally and linguistically relevant.** Results of screenings are only valid if the procedures and methods are appropriate for a given child's culture and language background. Including parent input about developmental expectations and the child's behaviors over time and settings is one of the best ways to ensure congruence with the child's life context.

A word about equivocal or inconsistent screening results: Anyone who has been involved in large- scale screening efforts knows that the results can be inconclusive or confusing for some children. Sometimes a score will just barely be above the cut-off that indicates a developmental problem. Other times a child's score will indicate a developmental concern at one screening point, and then no problems the next time, only to show a different concern at the third assessment. There is some indication that screening assessments sometimes identify children for further assessment who may not be eligible for special services, but are none-theless performing significantly below their peers in adaptive, academic and language areas (Glascoe, 2001).

When young children with inconsistent patterns of screening scores are from families challenged by poverty, a chronic lack of basic resources may explain depressed scores on screening assessments. For example, infants and toddlers who are malnourished or have chronic, untreated health conditions may demonstrate developmental delays better treated by food and medical care rather than early intervention services. Kindergarten children who have not had the benefit of preschool often perform poorly on academic screenings, as do primary school students who have experienced substandard instruction. Children who perform inconsistently on screening assessments should be monitored closely and provided additional services as necessary to support development and prevent the need for special education.

State and Federal Requirements

More than any other type of assessment, screening has become a centerpiece of recommended practice in early childhood programs across both regular and special education programs. All early childhood programs provided under federal and state regulations require screening of children as a condition for program approval and funding. Early identification of developmental and/or academic problems has become an important outcome for early childhood programs.

Federal regulations require screening for all preschoolers in Head Start programs and for infants and toddlers in Early Head Start. Performance Standards indicate that comprehensive screening is to be conducted for every child within 45 days of enrollment, including all areas of development, hearing, vision, and behavior. The Even Start family literacy program, regulated under No Child Left Behind (NCLB) legislation, also requires developmental screening and associated referrals for every child in the early childhood component of the program. A similar requirement for screening is contained in

the Washington State regulations that govern the Early Childhood Education and Assistance Program (ECEAP). State and federal agencies acknowledge the importance of early identification of developmental and academic problems by including screening in program regulations.

State and federal laws also require early intervention programs, local lead agencies, local school districts, and other agencies to engage in child find activities to identify birth to 21-year-olds who are eligible for services. In Washington State, early intervention and school district special education programs have a long history of providing public awareness information, developmental screening clinics, and coordination of referrals to appropriate services.

A recent trend in primary school academics is Response to Intervention (RTI), a three-tiered model of instruction where screening assessments are used periodically to identify students who are not making expected progress. Assuming that all children in a class are receiving high quality instruction in the core curriculum (tier I), screening assessments are expected to indicate that approximately 15% of students are not learning at the expected rate. This subgroup of students ideally receives more specific assessment and more frequent instruction (tier II) for the purpose of remediating academic problems early and preventing referral to special education (tier III).

The RTI model has been modified for preschool programs in a model called Recognition and Response, where screening for early literacy and math skills is a central component of early identification of children at risk for learning disabilities. Screening results that "recognize" children as behind in acquiring early academics skills indicate a need for a "response" in the form of increased time and attention to problem areas, to prevent more serious delays in learning.

School Readiness

Almost all American children attend kindergarten before entering first grade. A child's knowledge and skills at age 5 are the cumulative result of maturation and experience, and there is great variability among children entering kindergarten. Research studies indicate that at least half of the achievement gaps between poor and non-poor children in American education exist when children enter kindergarten. Many communities hold kindergarten screening clinics, sometimes designed to assist children and families to prepare for kindergarten entry, and sometimes to identify children who are and are not "ready" for school. There is not a unified and consistent opinion among experts about the knowledge and skills that are required for kindergarten entry. Typical kindergarten screening assessments include beginning knowledge of personal information, concepts, basic fine and gross motor skills, letters and numbers, communication, and social behavior.

There are many legitimate concerns about the concept of "kindergarten readiness" when the interpretation is that youngsters must demonstrate a certain level of knowledge and skills to enter kindergarten. Many experts suggest that the onus for a successful beginning in school rests with the school, families, and communities rather than being an inherent aspect of child development. From this perspective, schools need to be ready to address variability in experience and maturity of kindergarten children. Kindergarten readiness checklists are best used as curriculum guides for preschool programs, to ensure that children are being introduced early to the concepts and skills they will need for a smooth transition to school.

Just prior to kindergarten entry, screenings are most appropriate for identifying children who require supplementary supports and services such as every day classroom experiences to ensure success in school. *Early childhood professionals should never use kindergarten readiness assessments to screen children out and delay kindergarten entry.* Considered logically, those youngsters who have the most to learn should be afforded the opportunity to begin school as soon as possible, and provided the additional supports needed to be successful.

Potential Screening Instruments

Selecting appropriate screening instruments can be a daunting task; there are literally hundreds of commercially available screening tests for children birth to age eight. Screening for problems in early academics often involves curriculum-based general outcome measures such as one-minute timings for oral reading fluency in grade level texts, or one-minute think/three-minute writing samples.

The instruments listed here are selected examples of early developmental and academic screening tests that are widely used and consistent with the recommended practices described in the previous section. Additionally, each instrument listed has the following characteristics:

1. Developed for the express purpose of screening.
2. Adequate technical information to demonstrate reliability and validity.
3. Administration time of 30 minutes or less.
4. Comprehensive administration and scoring instructions.
5. Available from a reputable publishing company.

Sample Screening Instruments

- *Ages and Stages Questionnaires (ASQ)*, Brookes Publishing Company (available in Spanish, French, and Korean)
- *Battelle Developmental Inventory Screening Test*, Riverside Publishing
- *Developmental Indicators for Assessment of Learning (DIAL) III*, Pearson Assessments (includes Spanish materials)
- *Dynamic Indicators of Basic Early Literacy Skills (DIBELS)*, University of Oregon Center on Teaching and Learning
- *Early Screening Inventory-Revised (ESI-R)*, Pearson Early Learning (includes separate scoring for preschool and kindergarten)

Characteristics of Screening Assessments
Table 1

	Developmental (Birth – Primary Grades)	Early Academic (PreK – Primary Grades)
Question(s) Asked	**General Questions** • Does this child require further assessment? • Is the young child developing typically? • Is there a developmental problem?	**General Questions** • Does this child need additional services or extra help? • Is this student falling behind in academics?
Results	**Yes/No Answers and Categorical Results** • Below cut-offs = further evaluation • Above cut-offs = reassure • Close to cut-offs = frequent monitoring	**Yes/No Answers and Categorical Results** • Below cut-offs = further evaluation/extra help • Above cut-offs = core curriculum • Close to cut-offs = frequent monitoring
Type of Test	**Norm Referenced** • Compares child's scores to scores from a representative group of same age peers	**Criterion Referenced** • Curriculum-based measures, compares a student's performance to the "standard performance" in the local curriculum
Who Administers?	Most often administered by parents, paraprofessionals, professionals, or pediatric health care practitioners.	Most often administered by para-professionals or education professionals.
Administration	**"Quick & Comprehensive"** • Designed to be administered to large numbers of children at low cost; often covers all major domains	**"Quick & Broad"** • Designed to be administered to large numbers of children at low cost; often covers early reading, writing, and mathematics
Technical Adequacy	**Check for evidence of:** • Reliability and validity • Sensitivity and specificity	**Check for evidence of:** • Reliability and validity • Sensitivity and specificity
Instruments	• Comprehensive across all developmental areas	• Survey level curriculum-based measures, in content areas

Additional Readings and Resources

Developmental Screening in Early Childhood: A Guide. S. J. Meisels & S. Atkins-Burnett. (2005). Washington, D.C.: National Association for the Education of Young Children.

Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8. National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education (2003). Joint position statement. Washington, DC: National Association for the Education of Young Children.

Early Childhood Measures Profiles. D. Berry, L. Bridges, & M. Zaslow (2004). Child Trends. Online: http://aspe.hhs.gov/hsp/ECMeasures04/report.pdf

Evaluation and Assessment in Early Childhood Special Education: Children Who Are Culturally & Linguistically Diverse. Office of Superintendent of Public Instruction (1997). Online: http://www.k12.wa.us/SpecialEd/pubdocs/CLD.doc

Principles and Recommendations for Early Childhood Assessments. By Shepard, L., Kagan, S., & Wurtz, E. (1998). Washington, D.C.: National Educational Goals Panel. Online: http://www.negp.gov

Recognition and response: An early intervening system for young children at risk for learning disabilities. Executive summary. By Coleman, M. R., Buysse, V., & Neitzel, J. (2006). Chapel Hill: The University of North Carolina at Chapel Hill, FPG Child Development Institute. Online: http://www.recognitionandresponse.org/

Screening and Assessment of Young English-Language Learners: Draft Recommendations. National Association for the Education of Young Children & National Association of Early Childhood Specialists in State Departments of Education (January 2005). Joint position statement in supplement to *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8.*

The ABCs of CBM: A Practical Guide to Curriculum-Based Measurement. Hosp, M., Hosp, J., & Howell, K. (2007). New York: Guilford.

The words we use: A glossary of terms for early childhood education standards and assessments. Council of Chief State School Officers (2004). Retrieved December 14, 2006. http://www.ccsso.org/eceaglossary.

Using Response to Intervention for Washington's Students, includes manual, Powerpoint presentations. Office of Superintendent of Public Instruction (2006). Online: http://www.k12.wa.us/SpecialEd/RTI.aspx

*Washington State Infant Toddler Early Intervention Program.*Online: http://www1.dshs.wa.gov/iteip/sicc1.html

Assessments to Inform and Monitor Instruction

- My youngest son seems so content to sit and look at books instead of joining physical play. **Is this just a preference or is he having trouble with his motor skills?**
- The toddlers in my child care program exhibit a wide range of language and communication skills. **What sorts of activities and experiences will promote and support language development for the group?**
- Some of the preschoolers pick up new skills so quickly and others take much longer to learn. **How can I determine if all the children are learning what they should be?**
- **How can I document the progress of all first graders in reading?**
- This second grader is below grade level in math. **Where should we start instruction?**

The ultimate purpose of all assessment in early childhood programs is to support growth, learning, and development of young children. The undisputed centerpiece of an early childhood assessment system is the gathering of information that is directly connected to young children's daily learning experiences. Assessment to support teaching and learning is grounded in children's involvement in daily routines and interactions of home, child care, community, and classroom. This level of assessment provides early childhood professionals with information on each child over time to summarize the complexities of individual competencies and approaches to learning. This section of the guide describes characteristics of assessment to inform and monitor instruction, and makes recommendations for systematic and feasible incorporation of ongoing assessment into curriculum planning and delivery.

What is Assessment to Inform and Monitor Instruction?

The terms *classroom assessment*, *programmatic assessment*, *instructional assessment*, *curriculum-based measurement,* and *curriculum-based assessment* are all used to indicate assessments that are closely connected with teaching and learning. The term *instructional assessment* will be used in this section as a matter of efficiency to describe practices and instruments closely associated with teaching and learning. The terms *curriculum-based assessment* and *curriculum-based measures* will appear later in the section to describe specific instruments that assess early development and early academic skills, respectively.

Instructional assessment is a process that informs parents, child care providers, classroom personnel, and specialists about what children already know and are able to do, what they are expected to be learning next, and how quickly they are progressing. Although commonly listed and described as a particular category of assessment activity, in everyday practice instructional assessment is more aptly considered to be a vital component of curriculum and instruction. Instructional assessment is necessary to provide a starting point, continual feedback, and periodic review of the effectiveness of teaching and learning.

Familiar examples of instructional assessments are portfolios of children's work, anecdotal records, developmental checklists, and timings for oral reading fluency and math computation. Quality assessment for instructional purposes is a matter of collecting information that is directly relevant and useful

for planning curriculum and activities, designing plans for individual children, and monitoring progress toward learning goals.

The screening, diagnostic, and program evaluation assessments described in other sections of this guide are also important components of assessment systems in early childhood, but play a less central role in the actual processes of teaching and learning. For example, children who have existing diagnostic labels do not need to be included in screening assessments and bilingual youngsters often require more complex diagnostic procedures. Assessment to inform and monitor instruction, however, is universal and ongoing for ALL children in a program without contingencies, criteria, or exclusions.

Purposes of Assessment to Inform and Monitor Instruction

The primary purpose for instructional assessments is identification of developmental and academic competencies for individual children and monitoring of progress over time. Constant awareness of each child's acquisition and use of knowledge and skills, along with attentiveness to dispositions and attitudes, provides a point of departure for design and modification of curriculum activities. In addition, classroom- level assessments supply immediate feedback on the extent to which children are making progress in meeting developmental and academic standards. Data on the child growth and learning outcomes provides the best information for continual improvement of teaching, allowing teachers to revise curriculum and modify instruction in response to children's rates of learning and needs for support.

Classroom or instructional assessments also serve the purpose of providing a foundation for development of the individualized plans required in Head Start, ECEAP, early intervention, and special education programs: Individual Learning Plans (ILPs), Individualized Family Service Plans (IFSPs), and Individualized Education Programs (IEPs), respectively. The specific information gathered for instructional assessment supplies a rich description of present levels of developmental and/or academic performance, and indicates clearly which youngsters are acquiring new knowledge and skills as expected. Ideally, the core curriculum is also reflected in individualized learning goals, so instructional assessment data targets next steps in learning plans and guides early intervention plans or specially designed instruction.

A third purpose of instructional assessment is to yield information that is immediately relevant for children and parents. Children can review permanent products to see concrete representations of their own progress, and ultimately become participants in self-assessment. Classroom assessment data can likewise assist parents to understand their children's progress over time and in the context of a program's curriculum. Assessment that describes a child's development and learning in terms of recognizable and functional daily skills and behaviors is particularly meaningful to parents, and helpful for comparing and contrasting home and school points of view. This is an especially important purpose of assessment for infants and toddlers, because interventions are so often delivered within the context of home and family.

Characteristics of Instructional Assessments

Questions like the ones at the beginning of this section address the specific performances, preferences, profiles, and abilities of children individually and in groups. As a result, the information gathered to inform and monitor instruction is very specific to individual children, reflecting every child's unique growth and learning in relationship to the program's curriculum. Instructional assessments ideally use

multiple sources of information from multiple perspectives to build a distinctive profile of growth and development for each child in a program. In theory, there is no such thing as too much instructional assessment information; nonetheless, there are pragmatic limits on the amount of instructional time professionals wish to divert to collecting assessment data. *The secret of effective instructional assessment is to systematically sample important aspects of learning and development in ways that are efficient, yet representative of children's complete repertoires.*

Developmental and academic instruments to inform and monitor instruction are criterion-referenced, comparing a child's performance with a pre-specified set of performance standards. Results of assessments portray the complexities of child growth, development, and learning as individualized descriptions of skills and behaviors. Parent/child interaction records, narrative observational summaries, lengthy checklists across multiple areas of development, anecdotal records compiled over time, skill profiles, environmental inventories, parent ratings of child performance, running records of reading performance, oral reading fluency and math computation timings, and writing samples exhibit the range of instructional assessment formats. A hallmark of instructional assessment is frequent repeated measurements that indicate the level and rate at which children are progressing.

Instructional assessments should be curriculum-referenced, directly reflecting the content children are learning. In other words, the curriculum provides the criteria for performance assessments, and we assess what we teach, and vice versa. *The alignment between curriculum and assessment is an essential feature of assessment to inform and monitor instruction.* When the content of early childhood curriculum and instructional assessment is parallel, teachers and child-care providers can more easily integrate assessment within daily routines and schedules, improving the authenticity of results. It is essential, therefore, that instructional and classroom assessment items portray the skills necessary for young children to function spontaneously and independently in daily environments, and include all areas of a general developmental and/or early academic curriculum.

For infants, toddlers, and preschoolers, instructional assessments address the following areas of development:

- physical and motor;
- social and emotional;
- approaches to learning;
- language and communication;
- cognitive; and
- general knowledge.

Comprehensive assessment is desirable because of the arbitrary divisions among developmental domains and the interrelated nature of development and learning in different areas. Generally, early instructional assessments take the form of a developmental continuum of skills that suggests a logical teaching sequence.

Primary school assessments will obviously have a relative emphasis on language, as well as early reading, writing, and math, but should ideally continue to include physical and social/emotional areas. The preschool years are a time of transition from a developmental curriculum to incorporation of specific emphasis on early literacy and numeracy, and specific content areas such as science.

Instructional assessments are conducted and interpreted by the people who are responsible for design and delivery of an early childhood curriculum: teachers, paraprofessionals, specialists, and parents. Since instructional assessments are intended to be integrated with the curriculum and embedded in the routines and activities of classrooms, child care centers, and homes, it makes sense that the same people who provide instruction are those who collect, organize, interpret, and share assessment information.

There are generally no testing kits or specialized materials for instructional assessments. Data collection occurs within the usual care-giving and teaching routines of home and classroom, and seldom involves individual testing of birth to five-year-olds. When instructional assessment is conducted properly, infants, toddlers, and preschoolers should not be able to distinguish assessment from other interactions and activities during the day. Likewise, classroom assessments in primary grades are ideally conducted by teachers and require students to perform the same tasks with the same materials used during instruction.

There are different opinions on requirements for technical adequacy relative to instructional and classroom assessments. Authors of an influential assessment booklet for early childhood educators hold that reliability and validity are not as important for instructional assessment as for screening, diagnosis, and program evaluation (Shepard, Kagan, & Wurtz, 1998). This perspective emphasizes the ongoing nature of assessment to monitor and inform instruction, and assumes that decisions about daily instruction are relatively low-stakes because mistakes can be quickly and easily rectified. This view also assumes that screening, diagnostic, and instructional assessments involve separate and distinct procedures and instruments, as is often the case for infants, toddlers, and preschoolers.

A different perspective is widely held by authors who write about instructional assessment of academic skills (Hosp, Hosp, & Howell, 2007). Their opinion is that formative, curriculum-based measures must be reliable and accurate because the assessments are also used to inform high-stakes decisions such as identification of students in need of additional or alternative instruction, and selection of goals and objectives. As described in other sections of this guide, curriculum-based measures for early literacy and math are used for multiple purposes of screening, diagnosis, and progress monitoring. In primary grades, the distinctions among assessment purposes are often blurred in comparison to earlier years, with single instruments often serving multiple purposes.

One way to think about instructional assessment and requirements for technical adequacy is to consider the frequency, method of assessment used, and the types of decisions being made. Multiple perspectives might be more valuable than inter-observer reliability during observations to document growth and learning of typically developing youngsters. On the other hand, evidence of adequate reliability and validity is unquestionably important when assessment results are used to select and monitor progress toward IEP and IFSP goals.

Developmental Assessments to Inform and Monitor Instruction

Curricula for infants, toddlers, and preschoolers are often based on the patterns and sequences of typical development. Many commercially available developmental curricula are packaged with assessment instruments designed to identify curriculum placement and monitor progress for individual children. These packages are familiar to early educators as Curriculum-based Assessments (CBAs), instruments and procedures for initial information gathering, as well as repeated measurement of child performance across a continuum of curriculum goals and objectives. A number of CBAs (designed for use with children with and without special needs) are listed under Potential Instruments toward the end of this section.

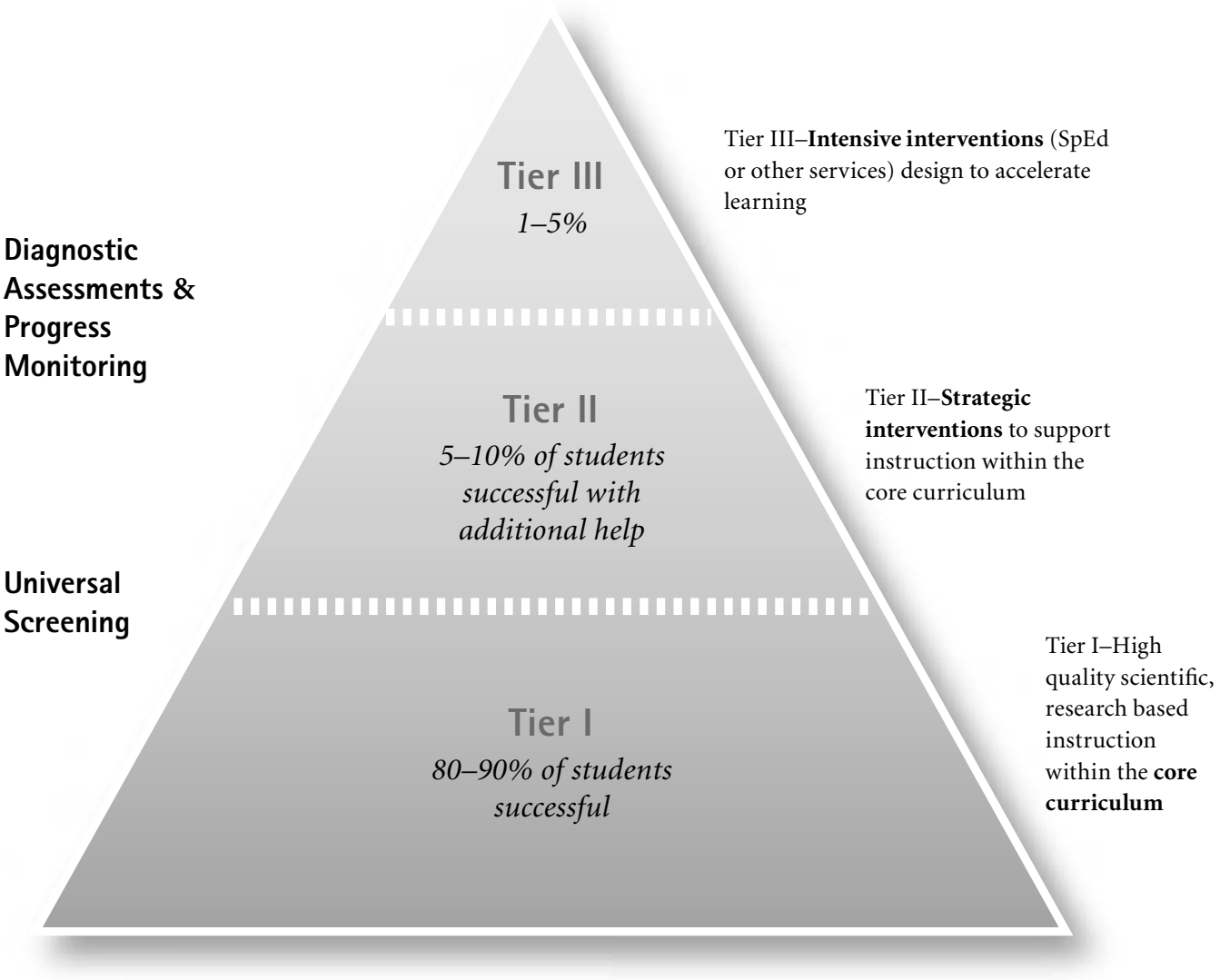
Almost all available CBA packages for infants, toddlers, and preschoolers address comprehensive content across developmental domains, some with measures sensitive enough to inform and monitor specially designed interventions (e.g. AEPS, Carolina, HELP) and others designed to address more generic outcomes for typically developing children (e.g. Creative Curriculum). For preschoolers, administration of curriculum-based measures in academic areas should be considered a supplement rather than a replacement for comprehensive, developmental assessment.

Assessments to Inform and Monitor Early Academic Instruction

During the primary school years, the focus of the curriculum shifts to early academics from development in physical, social, cognitive, and communication domains. Identification of students who need extra help in acquiring early math and literacy skills is evolving from a traditional model of diagnostic testing against criteria for learning disabilities to a model called Response to Intervention (RTI). RTI is based on a three-tiered model of instruction where tier I represents a high quality core curriculum. Tier II involves provision of additional or alternative instruction for those students who are identified by screening as having early academic problems. Children's progress is monitored frequently in tier II instruction, and the data are used to determine whether or not a referral to tier III, more intensive intervention or special education, is appropriate.

RTI is often represented as a triangle (see next page) with approximately 80-90% of children performing successfully in the core curriculum, approximately 5-10% of students achieving success with core curriculum plus tier II instruction, and about 1-5% of students being served in tier III instruction. These students will need intensive interventions if their learning is to be appropriately supported (Tilly, 2006).

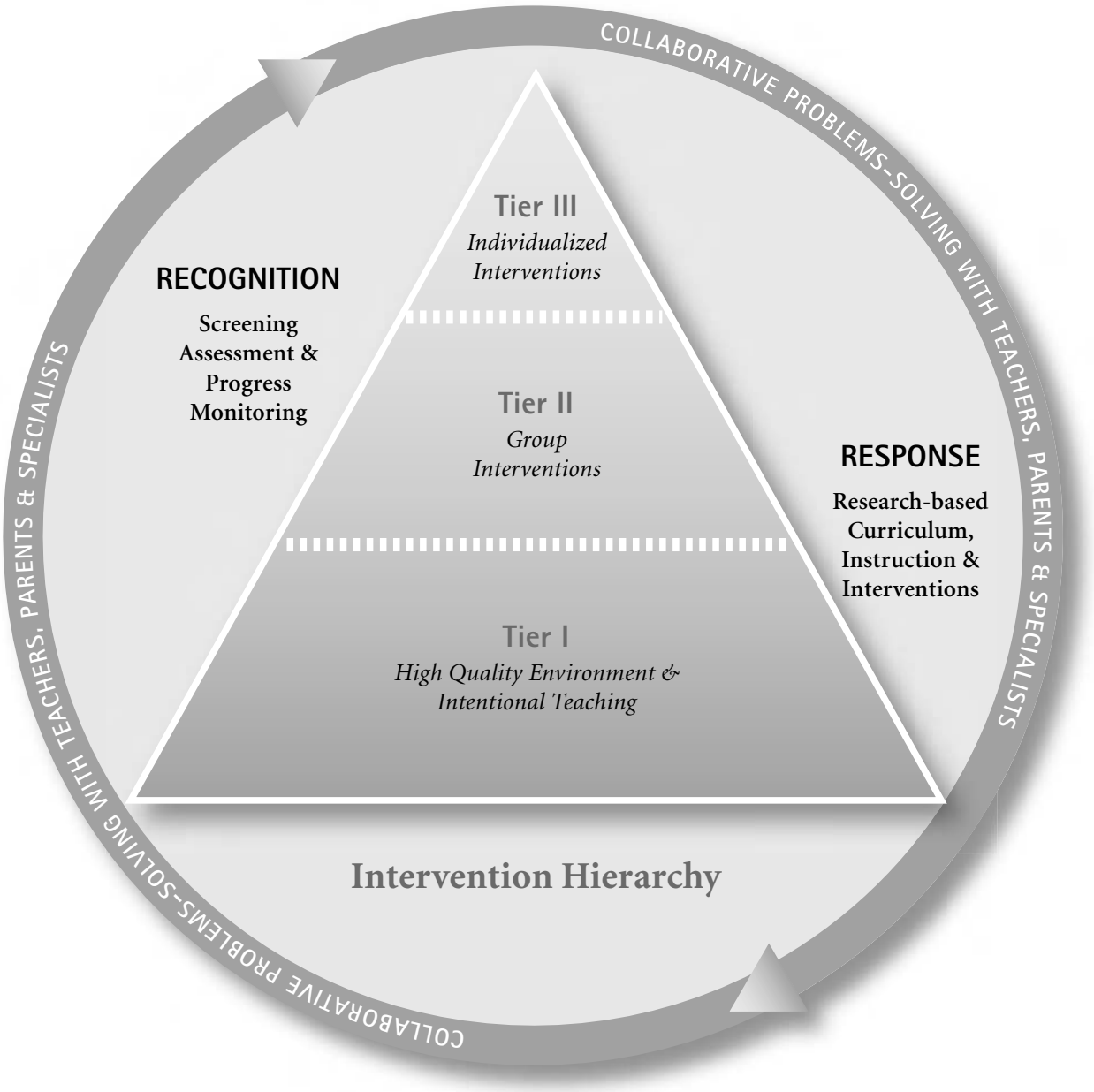
Figure 2: Response to Intervention (RTI) Model



Curriculum-based measures (CBMs) are used for purposes of screening, diagnosis, and monitoring in RTI models. Many districts have begun using comparisons to grade level academic expectations to identify problems in acquisition of early academic skills in reading, writing, and math. Individual student performance is compared to local norms, curriculum benchmarks, and/or the performance of peers who are demonstrating success in a particular academic area. Results of CBMs are combined with information gathered in interviews, record reviews, and observations in a problem-solving process to hypothesize the causes of and best interventions for academic delays. Individualized instruction is selected or designed accordingly, students receive additional help, and progress toward goals is monitored frequently (Hosp, Hosp & Howell, 2007).

The RTI model has been modified for preschool programs in a program called *Recognition and Response*, a system designed to identify 4- and 5-year-old children who show indications of being at risk for difficulties in early academics and/or learning disabilities. Beginning in preschool universal, classroom-based screening is conducted to identify individual children who are falling behind, triggering an immediate alteration of curriculum and instruction to provide additional help in problem areas. In this model, young children receive instructional support *before* their learning delays become serious enough to require special education. *Recognition and Response* is notable as a preventive approach that combines screening and diagnostic assessment with ongoing support for learning, blurring the distinction between early childhood general and special education.

Figure 3: Recognition and Response Model



Recommended Practices in Assessment to Inform and Monitor Instruction

The challenge for teachers and others who plan early childhood curriculum is to determine how much assessment data to gather and how to use the resulting information. The key is to organize a sustainable infrastructure that facilitates gathering meaningful information on every child regularly and systematically.

1. **The best instructional assessment information is collected frequently and used continuously to inform curriculum and instruction decisions.** Young children learn quickly, and regular monitoring of progress is necessary to document incremental improvements in skills and behaviors. Teachers and other adults are more likely to collect, interpret, and apply assessment data if the process is built into the daily classroom schedule and responsibilities.

2. **Assessment to inform and monitor instruction is most accurate if multiple sources are considered and multiple methods used.** Because much of young children's learning and development occurs outside classrooms, it is important that competencies and progress are considered in all contexts and environments. The richest results are obtained when the perspectives of parents, caregivers, teachers, specialists, and children themselves are combined. Assessment instruments that gather information by observation and report, and include components for parents to report on children's developmental and academic progress, provide a helpful structure.

3. **Methods used to collect instructional assessment information should resemble ongoing instruction and the interactions of home, child care centers, and classrooms.** There are valid concerns about the negative effects of testing on young children. One of the biggest advantages of assessment to inform and monitor instruction is that the process seldom necessitates individualized testing. Adults should be able to collect instructional assessment data using the same experiences designed to facilitate preschool learning in group activities, support parent/infant interactions at home, or provide instruction in primary classrooms.

4. **Assessment instruments should be criterion-referenced with items that reflect functional skills.** Actual assessment items and administration procedures will vary depending on whether the focus is in development or early academics. Nonetheless, repeated measures of performance across a sequence of skills that reflects curriculum goals provides the best documentation of children's progress. Assessment to inform and monitor instruction is only as good as the curriculum and instruction young children receive. Assessing what we teach and teaching what we assess is essential for meaningful instructional assessment.

5. **Instructional and classroom assessment instruments should ideally reflect a logical teaching sequence.** Assessments to inform and monitor instruction are most useful if items are organized in a sequence that reflects major skills in the curriculum, along with prior knowledge and/or prerequisite skills. This organization maps the curriculum for teachers and provides guidance for selection of subsequent learning goals.

6. **If results are being used to develop plans and monitor progress toward IEP goals or IFSP outcomes, instructional assessments must have data available to document reliability and validity.**

Data about the technical properties of tests are necessary for professionals to have confidence that individualized plans of instruction will be accurate and effective. Children receiving additional help cannot afford to spend instructional time or everyday learning activities working on goals, objectives, or outcomes that are too easy or too difficult. In addition, it is critical that changes in assessment results reflect real progress rather than inconsistencies in administration, scoring, or interpretation of assessment results.

7. **Instructional assessment instruments and procedures must be culturally and linguistically relevant.** If the basic skills represented in early developmental and academic curricula are functional and appropriate for young children, instructional assessments have the advantage of being able to incorporate familiar materials, people, routines, and important events of a child's life. On developmental CBAs, items that are culturally inappropriate can be revised, and parents can suggest more familiar and appropriate materials and behaviors. It is especially important that English Language Learners not be penalized by test materials or directions that confuse cultural and language differences with cognitive or academic delays.

Selecting methods and organizing a system for gathering instructional assessment data for even a small group of young children might initially seem to be an intimidating undertaking. The good news is that there are many existing models and instruments to use in designing an approach that will work best for any given program, many of which are listed at the end of this section under Potential Instruments and Additional Resources.

Including CBAs in assessment systems for birth to five-year-olds offers a structured and consistent framework for periodic collection of instructional assessment information. Most early childhood CBAs can be scored primarily or totally via observation of children engaged in activities and interactions at home, in child care settings, or in the classroom. Detailed results identify precisely the skills and behaviors that have been acquired, those that need more practice, next steps in learning, and problematic skill and knowledge areas. Similarly, the use of CBM is an efficient, reliable, and valid approach to identification of early academic goals, diagnosis of academic problems, and progress monitoring. A summary of characteristics of instruments to inform and monitor instruction is included at the end of this section.

Methods for Collecting Ongoing Assessment Information

Effective instructional assessment does not require any one particular method to collect, interpret, or share the information. Any of the assessment methods described below can be employed, preferably in combination, to gather instructional assessment data. Selections of specific methods and instruments will depend on the size of the group, ages and characteristics of children, areas being assessed, and program resources.

The most popular and useful methods include:

- **Observations** – A long-standing tradition in early childhood education is the practice of observing young children as they play and interact with one another during daily activities and routines. Virtually all early childhood professionals make use of their observations of behaviors, skills, knowledge, and attitudes of children in their programs to guide curriculum design, activity

planning, interactions, and instruction. Running records and anecdotal reports of children's behavior have been characterized as informal documentation of developmental progress. A current trend is the more structured and systematic use of observation as the primary method to collect assessment data to complete assessment components of curriculum-based assessment instruments (CBAs).

- **Interviews** – Interviews involve collecting child assessment data that is reported by people familiar with children's skills and behaviors. Interviews are useful for including the multiple perspectives of parents, other adults, and children themselves. Interviews are the preferred method for gathering contextual information to clarify a child's history, describe variability across settings, and identify family priorities. Some CBAs include structured components for parents to report directly on child development and learning.

- **Permanent Products of Children's Work** – A popular method for documenting certain types of skills is to collect and display children's work over time. Viewing progress of skills in drawing and writing is especially efficient and informative. Audiotapes of language samples and oral reading are also very illustrative. Permanent products are often included in portfolios that are organized to include observation records, photos, interview excerpts, and other sources of assessment information.

- **Direct Assessment** – Direct assessment has traditionally meant removing children from familiar environments and testing them individually. Currently, direct testing for purposes of informing and monitoring instruction of infants, toddlers, and preschoolers is more likely to involve observation of individual children as they play and interact at home and school. In primary grades, direct assessment of early academic skills is usually conducted at the classroom level and interpreted individually.

Using Instructional Assessment Effectively

Every method of assessment has advantages and drawbacks. Repeated observations of youngsters over time and across settings provides rich and detailed information on authentic performance, but requires design of a system for organizing and summarizing information, as well as a significant time investment for both collecting and interpreting data. Interviews with parents and observations of children at home are especially informative for English Language Learners, but often entail the services of an interpreter. Permanent products of children's work provide concrete samples of progress over time, but are only representative of specific types of skills.

Mixing and matching methods can create flexibility in collecting information to accommodate the characteristics of individual children within a systematic plan that is realistic for an entire group of youngsters. In collecting samples of early writing, for example, a teacher could decide to sample twice as many products from a child growing up in a bilingual home. Likewise, an interview in the home might augment a brief written survey that all parents fill out.

A good general strategy is to identify a manageable set of assessment methods that are efficient and yield an adequate level of information for *all* children as the basis for an assessment system. For some children, more time-consuming and in-depth methods might be necessary to achieve the same level of understanding. And the services of a specialist may be needed to supplement assessment for those few children whose development and behavior is especially complex.

Early childhood personnel must constantly balance the need for instructionally relevant assessment information against the time and resources necessary to collect, organize, interpret, display, and store data. The most defensible assessment systems are those that deliver the maximum amount of information and take the least time away from instruction.

Potential Instruments to Inform and Monitor Instruction

The instruments listed here are selected examples of instructional assessment for children with and without disabilities. The selections represent those that are widely used and each is consistent with most of the recommended practices described in the previous section. In addition, the list of potential instruments shares the following criteria:

1. Includes explicit connections to curriculum;
2. Explicitly designed to inform instruction and monitor progress;
3. Facilitates assessment in familiar daily environments and activities;
4. Criterion-referenced;
5. *Assessment component supported by reliability and validity studies.

Sample Instructional Instruments

- *Assessment, Evaluation, and Programming System*, Second Edition (AEPS)**, Brookes Publishing Company
- *Carolina Curriculum for Infants and Toddlers*, Brookes Publishing Company
- *Carolina Curriculum for Preschoolers*, Brookes Publishing Company
- *Creative Curriculum Developmental Continuum*, Teaching Strategies, Inc.
- *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS)*, University of Oregon Center on Teaching and Learning
- *Hawaii Early Learning Profile (HELP)*, VORT
- *High Scope Infant Toddler COR*, High/Scope Press
- *High Scope Preschool COR, High/Scope Press*
- *The Work Sampling System*, Pearson Early Learning

Characteristics of Assessments to Inform and Monitor Instruction
Table 2

	Developmental (Birth – Primary Grades)	Early Academic (PreK – Primary Grades)
Question(s) Asked	**General Questions** • What are the child's unique developmental and early learning strengths and needs? • What are appropriate goals and outcomes for the child?	**General Questions** • What are the child's unique academic/behavioral strengths and needs? • What are appropriate goals and outcomes for the child?
Results	**Formative Data** • Present levels of developmental performance • IEP/IFSP development • Skills checklists • Progress monitoring data	**Formative Data** • Present levels of educational performance • IEP goals and objectives • Skills checklists • Progress monitoring data
Type of Measures	**Criterion Referenced** • Curriculum-based measures, compares a student's performance to repeated measures for ongoing progress monitoring in goals and objectives	**Criterion Referenced** • Curriculum-based measures, compares a student's performance to repeated measures for ongoing progress monitoring in goals and objectives
Who Administers?	Most often administered by parents, paraprofessionals, educational professionals, and specialists	Most often administered by paraprofessionals, educational professionals, and specialists
Administration	**"Functional & Individualized"** • Designed to produce an ongoing record of what children know and are able to do in daily environments (home, school, community) • Skills included should reflect culturally relevant and age appropriate goals	**"Functional & Individualized"** • Designed to produce an ongoing record of basic skill acquisition (e.g. reading, writing, math, social) • Assessments should be unbiased and interpreted in light of culture and language differences
Technical Adequacy	**Check for evidence of:** • Reliability and validity	**Check for evidence of:** • Reliability and validity
Sample Instruments	Comprehensive, criterion-referenced instruments	Specific level curriculum-based measures in content areas

Additional Readings and Resources

Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8. National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education (2003). Joint position statement. Washington, DC: National Association for the Education of Young Children.

Early Childhood Measures Profiles. D. Berry, L. Bridges, & M. Zaslow (2004). Child Trends. Online: http://aspe.hhs.gov/hsp/ECMeasures04/report.pdf

Evaluation and Assessment in Early Childhood Special Education: Children Who Are Culturally & Linguistically Diverse. Office of Superintendent of Public Instruction (1997). Online: http://www.k12.wa.us/SpecialEd/pubdocs/CLD.doc

Principles and Recommendations for Early Childhood Assessments. Shepard, L., Kagan, S., & Wurtz, E. (1998). Washington, DC: National Educational Goals Panel. Online: http://www.negp.gov

Recognition and response: An early intervening system for young children at risk for learning disabilities. Executive summary. Coleman, M. R., Buysse, V., & Neitzel, J. (2006). Chapel Hill: The University of North Carolina at Chapel Hill, FPG Child Development Institute. Online: http://www.recognitionandresponse.org/

Screening and Assessment of Young English-Language Learners: Draft Recommendations. National Association for the Education of Young Children & National Association of Early Childhood Specialists in State Departments of Education (January 2005). Joint position statement in supplement to *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8.*

The ABCs of CBM: A Practical Guide to Curriculum-Based Measurement. Hosp, M., Hosp, J., & Howell, K. (2007). New York: Guilford.

The words we use: A glossary of terms for early childhood education standards and assessments. Council of Chief State School Officers (2004). Retrieved December 14, 2006. http://www.ccsso.org/eceaglossary

Using Response to Intervention for Washington's Students, includes manual, Powerpoint presentations. Office of Superintendent of Public Instruction (2006). Online: http://www.k12.wa.us/SpecialEd/RTI.aspx

Washington State Infant Toddler Early Intervention Program. Online: http://www1.dshs.wa.gov/iteip/sicc1.html

Diagnostic Assessments

- My infant son isn't able to sit without support. My other children were able to sit unsupported at his age. **How serious is his delay?**
- One of the toddlers in my child care isn't talking like the other kids. **Why isn't she talking and how much is she understanding?**
- My niece acts so young for her age and needs a lot of help to learn new things. **Is she eligible for special education preschool?**
- **Why isn't my child learning to read like the other first graders?**
- This second grader is below grade level in math. **What does he need to learn in order to catch up to his peers?**

Health, education, and related services professionals need more detailed information than screening assessments can provide, in order to make appropriate referrals and select effective interventions. A complete description of a young child's delay or disability is necessary in order for parents to partner with professionals and incorporate individualized intervention strategies into family routines. Eligibility criteria for infant toddler early intervention and preschool special education services define developmental delay or disability in terms of standard scores on comprehensive developmental assessment instruments. Kindergarten and primary teachers must identify the specific components of early reading and math skills that are problematic for students who are struggling, in order to tailor effective instruction. *Diagnostic assessments are designed to provide detailed information about developmental delays or disabilities and early academic problems.*

What is Diagnostic Assessment?

Diagnostic assessment is a comprehensive procedure that addresses specific questions about the development, knowledge, and skills of young children. A careful and systematic process is used to diagnose problems in a particular area of development or academics, and a relatively large amount of information is used to build a fine-grained understanding of a child's problem. The results of diagnostic assessments are designed to describe a developmental or academic problem with some precision. Current best practices in assessment recommend the use of diagnostic assessment procedures to guide targeted interventions, as well as to determine eligibility for special services.

A common example of diagnostic assessment is eligibility determination by a multi-disciplinary team (MDT): infants and toddlers for early intervention; preschoolers for special education, mental health and/or related services. Until recently, eligibility determination for infant toddler early intervention and special education has been the centerpiece of diagnostic assessment in early childhood. Typically, screening assessments establish the existence of a developmental or academic problem and trigger a referral to the MDT for a comprehensive evaluation in the areas of concern or suspected disability. The eligibility determination process combines review of records, interviews, observation, and testing to reveal a comprehensive summary of the developmental or academic problem, and render a decision about eligibility.

An undeniable problem of diagnostic assessment grounded in eligibility determination is a lack of connection between the information gathered to determine eligibility, and information that is relevant for instruction and intervention. Students identified as having delays by screening/benchmarking in early academic areas, especially reading and math, may not exhibit problems severe enough to qualify for special education. These students remain at a continued disadvantage if the only outcome of diagnostic assessment is a yes/no decision about eligibility. To address this problem, criterion-referenced diagnostic assessment of academic problems is increasingly being employed in primary school general education classrooms.

Purpose of Diagnostic Assessments

The purpose of diagnostic assessment in early childhood is to identify and secure appropriate intervention services for children whose development and learning are delayed. Access to targeted interventions involves identifying the nature and severity of developmental, academic, or learning problems comprehensively and systematically. Comprehensive, in-depth testing of any sort is relatively expensive, time-consuming, and often requires the expertise of specialists. As a result, diagnostic assessment of young children is reserved for those few who do not demonstrate typical growth and learning trajectories, for example those identified by screening assessments as having potential developmental or academic problems. Diagnostic assessment may determine eligibility for a number of programs such as infant toddler early intervention, preschool special education, and mental health; or it may indicate the need for additional or alternative instruction in pre- or early academic areas.

Characteristics of Diagnostic Assessments

Questions like the ones at the beginning of this section are quite specific, so the information- gathering process is relatively lengthy and detailed, often including medical and social histories, comprehensive developmental testing by a variety of professionals, and team meetings with professionals and parents. Diagnostic tests have many fine-grained items that assess very specific indicators of development and learning. Performance is generally summarized numerically in standard scores such as percentile ranks, standard deviations in relation to the mean, age or grade equivalencies, or percent delay from an age or grade norm. Diagnostic assessments are usually conducted by professionals who have been specially trained.

Eligibility Determination for Infants, Toddlers, and Preschoolers

Comprehensive diagnostic instruments are generally norm-referenced and are required in Washington State to determine if an infant, toddler or preschooler has a developmental delay or disability, and is eligible and in need of services. Diagnostic assessment results for this group are generally presented as a standard score that compares a child's performance in physical, adaptive, social/emotional, cognitive, and communication domains to a group of same age peers.

In Washington and many other states, early intervention and special education preschool eligibility require test scores that summarize performance in terms of standard deviations below the mean, or percent delay. Children whose test scores fall at or below stated eligibility criteria are determined to be eligible for special services. Eligibility decisions are ideally based on scores from norm-referenced tests

that demonstrate adequate reliability and validity, and have large norm samples that reflect the characteristics of the youngsters being evaluated. Eligibility determination criteria for infants and toddlers, preschoolers, and primary school children in Washington State are summarized in the Table 3 below.

Washington State Eligibility Criteria for Infant Toddler Intervention and Special Education Table 3

Program/Age Range	Regulatory Authority	Eligibility Criteria
Early Intervention **Infants & Toddlers** (birth–three years)	**Federal:** Individuals with Disabilities Education Act (IDEA) of 2004, Part C **RCW:** Chapter 70.195 Early Intervention Services - Birth to Six **State of Washington**'s Federally Approved Plan, IDEA, Early Intervention Section	**To be eligible, a child must:** • Have a 25% delay or score 1.5 standard deviation below the mean in one or more of the developmental areas OR • Have a physical or mental condition such as Down syndrome that has a high probability of resulting in delay
Special Education Preschool **Preschool** (three–five years)	**Federal:** IDEA '04, Part B, Section 619 **State of Washington** Rules for the Provision of Special Education to Special Education Students Chapter 392-172 WAC	**To be eligible, a child must:** • Score two standard deviations below the mean in one or more of the five developmental areas OR • Score one and one-half standard deviations below the mean in two or more developmental areas OR • Meet one of the other eligibility criteria AND • Need specially designed instruction
Special Education Primary **Primary Grades K–3** (five–nine years)	**Federal:** IDEA '04, Part B **State of Washington** Rules for the Provision of Special Education to Special Education Students Chapter 392-WAC	**To be eligible, a child must:** • Score two standard deviations below the mean in one or more of the five developmental areas OR • Score one and one-half standard deviations below the mean in two or more developmental OR • Meet one of the other eligibility criteria AND • Need specially designed instruction

The quality of a comprehensive developmental instrument depends on the extent to which it accurately identifies young children who in fact need special services, and provides information about the areas of need. Multiple items in each developmental area for every age level are desirable in order for thorough evaluation of skills. Items on norm-referenced diagnostic tests are selected as developmental indicators that discriminate among children's behaviors at different ages, and are always administered in the same order, using the same directions and a standard set of materials. Items that all children pass or fail are eliminated during test development because they do not reflect variability.

Because items are selected statistically to "spread out" scores of youngsters being tested, behaviors evaluated on norm-referenced tests are not necessarily functional or teachable skills for young children. For example, standing on one foot might be a great item for testing balance but is not a skill that is

required for successful participation in most preschool games and activities. Items on norm-referenced tests of social behavior often describe maladaptive behaviors or characteristics of temperament that may well indicate a need for mental health services, but are not appropriate goals for teaching.

Test materials in diagnostic assessment kits are specialized and designed to be unfamiliar, in order to provide an equitable experience for all youngsters who take the test. The blocks, form boards, shape puzzles, sticks, dolls, vehicles, cups, and utensils provided in most test kits are intended to test specific skills without distracting the child to extraneous play activity. Additionally, the same materials may be used to test a wide range of skills. The one-inch cubes common to many test kits, for example, are used to evaluate simple grasping skills, spatial relations for filling and dumping a cup, and more complex concepts such as stacking and aligning. Tiny pills and small glass bottles can be used to evaluate shifts in visual attention, fine motor control, and problem solving. Again, the materials are designed to be equally unfamiliar and prevent any group of youngsters from being advantaged or disadvantaged by experience, rather than relevant to instruction or play. One-inch cubes may not be the most interesting toys for building and play, and we certainly do not want to teach young children to dump pills out of bottles!

The obvious disadvantage of diagnostic assessment that yields scores designed to inform a yes/no decision about eligibility is that there is little connection between the testing instruments and subsequent intervention efforts. To incorporate instructionally relevant information into the diagnostic process, a number of states have also begun allowing use of criterion-referenced tests that meet specific psychometric criteria for determining eligibility of infants, toddlers, and preschoolers.

One promising and progressive alternative is the use of authentic, criterion-referenced assessments that are connected directly to curriculum and have validated cutoff scores. Although such cut-offs are calculated differently than standard deviation or percent delay, research indicates they are as accurate as norm-referenced tests at identifying those children who should be eligible for services (Macy, Bricker, & Squires, 2005). In addition, the content of these instruments has the advantage of being aligned with curriculum.

Diagnostic assessments are scored in various ways, but raw scores are generally converted to standard scores that are used to inform decisions about access to specialized services. Results supply information about relative areas of developmental strengths and delays, and further describe how serious the problems are. Whether diagnostic assessment for eligibility is norm-referenced or criterion-referenced, the nature and severity of a delay is generally represented by a single cut-off score that determines eligibility, for example:

1. "The infant is eligible because her scores in fine motor and cognition are below the cut-off for a child her age."—"The toddler is eligible for early intervention; she has a 40% delay in language and cognition."—"The preschooler scored 2 SD below the mean compared to his peers in the motor domain and 1.5 SD below the mean in the social domain."

   OR

2. "The infant isn't eligible for early intervention because her score in the communication domain indicates less than a 25% delay"—"Your child is not eligible for special education preschool because his scores in the motor and social domains are less than 1 SD below the mean"—"The

four-year-old is not eligible for special education preschool because he scored above the cut-off in each domain."

Diagnosis of Early Academic Problems

Many districts have begun using criterion-referenced, grade-level academic expectations to diagnose problems in learning reading and math. Individual student performance is compared to local norms, curriculum benchmarks, and/or the performance of peers who are demonstrating success in a particular academic area. Results of survey and specific-level assessments are combined with information gathered in interviews, record reviews, and observations in a problem-solving process to hypothesize the causes of and best interventions for academic delays. Individualized instruction is selected or designed accordingly, students receive additional help, and progress toward goals is monitored frequently (Hosp, Hosp, & Howell, 2007).

Proponents of Response to Intervention (RTI), a three-tiered model of instruction, recommend use of formative curriculum-based measures (CBMs) rather than a single point in time assessment to diagnose learning disabilities in primary school students. Screening or benchmarking is used to identify a subgroup of students who are falling behind in the core curriculum (tier I). These children receive more specific assessment and more frequent or alternative instruction (tier II) to address their specific academic needs. The decision to make a referral to special education, or other more intensive interventions (tier III) is made when progress monitoring indicates that students are not progressing even with additional support and instruction.

In contrast to a single test score that determines eligibility, ongoing responses to interventions (RTIs) are evaluated over time using curriculum-based measures (CBMs), and progress monitoring data are used to determine whether or not a special education referral is appropriate. The use of CBMs to diagnose academic problems and monitor progress toward goals is discussed more completely in the section on Assessment to Inform and Monitor Instruction.

The RTI model has been modified for preschool programs in a similar model called *Recognition and Response*. Regularly scheduled, universal screening in early academic areas "recognizes" children who are behind and indicates the need for a "response": more thorough assessment and additional help in problem areas. In this model, young children receive instructional support *before* their academic problems become serious enough to require special education. *Recognition and Response* is notable as a preventive approach that combines diagnostic assessment with ongoing support for learning, and blurs the distinction between early childhood general and special education. See the end of this section for a summary table of characteristics of diagnostic instruments.

Recommended Practices in Diagnostic Assessment

One major advantage of diagnostic assessment instruments is that they come with specific directions for administration and scoring, and often include a test kit of materials. Careful selection of assessment instruments is important because good decisions can only be made on the basis of good information. When conducted properly, diagnostic assessment involves systematic decision-making rather than administration of a single test:

1. **Diagnostic procedures should always include multiple sources of information, with special attention to the family perspective in gathering information and interpreting results.** Diagnostic testing is never conducted as an end in itself, but rather to determine the need for early intervening, or to determine eligibility or access to special services. The implication for practice is careful and systematic efforts to interpret test results in the context of information gathered via direct observation of a child in familiar environments, interviews of parents and other caregivers, and review of pertinent screening results, portfolio artifacts, as well as medical and educational records. The more a child's language, early experience, and family background differs from the majority culture, the more lengthy and complex diagnostic assessment may need to be in order to yield accurate and valid results.

2. **Diagnostic assessment is used to include children in placements and services, never to exclude children.** In particular, assessment of early learning and academic skills should never be used as the sole criterion to deny children kindergarten entrance. When scores on school readiness tests suggest a concern, children should be referred for additional help to accelerate their learning.

3. **Use of individual norm-referenced diagnostic tests that are not directly connected with curriculum should be limited with young children.** Diagnostic assessment is time consuming and diverts time from teaching and learning opportunities. Administration of individual developmental and academic assessments should be limited to those children for whom screening results have indicated a potential developmental delay or concern.

4. **Diagnostic results should only be used for the purpose they are developed: to identify children who are eligible for additional help, special services, or specific interventions.** It is no doubt clear by this point that diagnostic tests are too lengthy and costly to be used with large groups of children, so are inappropriate for screening purposes. Additionally, the results of diagnostic developmental assessments are inappropriate for identification of intervention or instructional goals because test content is not functional or instructionally relevant. In Washington, as in many states, there is little to be gained by conducting additional diagnostic testing for eligibility when a child has an established physical or mental condition that has a high probability of resulting in a development delay such as Down syndrome or Autism.

5. **Developmental and academic diagnostic instruments should be norm-referenced (pg. 16), with explicit standards for comparison.** Norms may be derived from a conventional norming process for developmental tests, or set via determination of local norms or accepted grade level expectations for academics. All diagnostic instruments should be standardized in their administration

and scoring. For a meaningful comparison of individual scores to a norm or pre-set criteria, all youngsters must have the same experience during the diagnostic assessment.

6. **Diagnostic instruments must have data available to document reliability and validity (see pg. 17), and/or data to support cut-off scores for eligibility decisions.** Data about the technical properties of tests are necessary for professionals to have confidence in the results of diagnostic assessment. Eligibility determination for special services is a high stakes endeavor, and professionals need to be confident that the results of diagnostic assessments are accurate.

7. **Diagnostic procedures must be culturally and linguistically relevant.** Results of diagnostic assessments are only valid if the procedures and instruments used are appropriate for a given child's culture and language background. Lack of proficiency in English may be misinterpreted as cognitive or academic delay, and unfamiliar patterns of behavior may be misinterpreted as behavior problems. It is especially important that English Language Learners be screened for language proficiency in both English and their primary language, to prevent misinterpretation of assessment results. Testing for developmental delay or disability should be conducted to the extent possible in the child's primary language, which may involve the use of interpreters, alternate forms of information gathering, and professional judgment. Care should be taken to ensure that all children understand directions for test administration.

Anyone who has administered diagnostic assessments has experienced the frustration of interpreting scores that are ambiguous or do not seem to adequately reflect a child's abilities. The performance of young children is highly variable even within a single assessment session.

Sometimes a child will perform very poorly on a screening test and then score well on more comprehensive assessments. Did s/he just have a bad day previously, or is her/his behavior generally inconsistent? Other times, children may receive a score that is just above or below the cut-off for eligibility. Will the potential benefits of special services outweigh the disadvantages of a more complex schedule and the stigma of being labeled? A child may receive a low score that is nonetheless characterized by inconsistently sophisticated skills in some areas. Will additional services really help or did s/he not understand some of the directions during the test?

It is a difficult venture to make a decision about eligibility or the need for additional services based on a single test score when children's behaviors are variable or inconsistent. Familiar examples are the young English Language Learner (ELL) who is suspected to also have delays in cognition and communication, and youngsters whose behavior interferes with diagnostic assessment procedures. In the first case it is probably unreasonable to expect that one assessment will suffice to distinguish between a language disorder and typical bilingual acquisition. In the second, it may not be possible to determine if behavior issues are specific to the assessment situation or evidence of a more serious social-behavioral problem.

Additional assessment will undoubtedly be necessary in both scenarios, and professional judgment can also supplement test results by integrating interview and observation data that might lead to different conclusions about the nature and severity of developmental problems.

State and Federal Requirements

A number of early childhood programs have responsibilities related to diagnostic assessment. Federal regulations require Head Start and Early Head Start programs to initiate referrals and coordinate diagnostic assessments for children whose screening results indicate a possible disability or delay. Head Start staff must also provide support for parent participation in eligibility determinations for early intervention, special education, and mental health services. When infants and toddlers are determined to be eligible for early intervention services, Early Head Start staff work in partnership with parents during development of the IFSP.

The Even Start family literacy program, regulated under No Child Left Behind (NCLB) legislation, also requires staff to make diagnostic assessment referrals for children suspected of having developmental delays. Washington State regulations that govern the Early Childhood Education and Assistance Program (ECEAP) include a similar requirement for staff to work collaboratively with local programs to coordinate diagnostic assessments and support parents during the eligibility determination process.

State and/or federal laws also regulate determination of eligibility for support programs under NCLB, such as Title I, Migrant and Bilingual, and American Indian services, as well as early intervention and special education services. Procedures, methods, criteria and timelines for determining eligibility for children ages birth to three are described in the Washington State Federally Approved Plan for the Individuals with Disabilities Education Act (IDEA) Early Intervention Section at http://www1dshs.wa.gov/iteip/FedAppPolicies.html. .

Procedures, methods, criteria, and timelines for determining eligibility for children ages three to 21, as well as periodic evaluation for continuing eligibility, are regulated through Washington Administrative Code (WAC 392-172a). Rules for the Provision of Special Education to Special Education Students are available at http://www.k12.wa.us/SpecialEd/pubdocs/wac/WAC_392-172a.doc.

School Readiness

Relatively few kindergartners come to school with Individual Educational Plans (IEPs) from preschool special education programs. Those who do tend to have fairly serious and pervasive disabilities or delays that disrupt basic early development in communication, motor, social, and cognitive areas. Sometimes these youngsters have been receiving special services from the time they were infants.

Perhaps because IEPs make special needs of this group readily apparent, we tend to think of youngsters who are eligible for special education and those who are not as two separate groups. In reality, the abilities of young children fall along a continuum, and the majority of students who are referred to special education during primary grades enter kindergarten with unidentified problems in learning and development.

Nationwide, rather few children are determined to be eligible for special services while in school. There is a lot of variability in performance based solely on whether or not a child has attended preschool. Maturation is a powerful force at this age, and the typical range of knowledge and skills from "just five" to "almost six" is quite substantial. Many of the children who struggle early in their kindergarten year do, in fact, mature by spring. But other youngsters enter kindergarten with serious learning problems,

developmental delays, and disabilities that are evident from the first days of school. It goes without saying that these children are candidates for diagnostic assessment during kindergarten, so that they can receive the benefits of specially designed instruction before they go on to first grade.

There is, however, reasonable reluctance to conduct expensive diagnostic testing with children who might just need some time to negotiate the social demands and settle into the regular routines of school. Perhaps diagnostic assessment does not occur very often during kindergarten because there is a "wait and see" attitude, in hopes that the variability among children will decrease as the less capable catch up to their peers. This is somewhat understandable since there is sometimes only a few points difference in diagnostic assessment scores between children who are eligible for special services and those who are not.

Eligibility determination for special services, therefore, is not the only or best solution to the problem of early learning problems, because not all children who have trouble with early academics will have serious enough problems to be eligible. But a "wait and see" attitude further compromises progress for those children who are at risk for learning disabilities without extra support. Kindergarten students who are falling behind academically need extra help, but diagnostic assessment for eligibility determination is not usually the help they need.

One sensible approach for early identification of academic problems is to use universal screening assessments to identify those children who have potential problems in early literacy and math specifically, and to rescreen after a relatively short interval. Any child whose results indicate concerns for two successive screenings then receives targeted attention to skill development in the areas of concern.

For example, if a youngster's early literacy screening indicated problems with phonemic awareness skills and letter/sound association, extra help and additional practice in those areas would be provided. If extra help and practice are unsuccessful in catching the child up, diagnostic assessment can be used to identify specific instructional interventions rather than being used only for eligibility determination. With this approach, kindergarten children at risk for learning disabilities receive the additional help they need without having to be determined eligible for special services. Continuous progress monitoring provides information about the child's response to targeted interventions, and diagnostic assessment for eligibility determination is reserved for those situations where children continues to fall behind and not progress with the extra intervention.

Potential Diagnostic Instruments

Selecting appropriate diagnostic instruments can be a daunting task; there are scores of commercially available norm-referenced tests for assessing the developmental status of children birth to age eight. Diagnosing problems in early academics often involves multiple CBMs such as survey and specific-level assessments of early reading, math, and writing skills.

It is desirable to have an identified battery of diagnostic assessment instruments available and approved for use with birth to eight-year-olds. The enormous range of diagnostic assessment information required for groups of young children requires a variety of available instruments to address developmental and academic, sensory, and mental health areas at a minimum. While few children will require a full battery of tests, it is most efficient and effective to have the instruments available with guidelines for implementing supplemental testing in addition to a standard core set of assessments.

The instruments listed here are selected examples of diagnostic tests that are widely used and consistent with the recommended practices described in the previous section. Additionally, each instrument listed has the following characteristics:

1. Developed for the express purpose of diagnosing developmental and/or academic problems in multiple or single areas.
2. Adequate technical information to demonstrate reliability and validity, or to support cut-off scores.
3. Comprehensive and detailed administration and scoring instructions.
4. Large and representative norm group and current norms.
5. A test kit or detailed listing of materials for assessment.
6. Available from a reputable publishing company.

Sample Diagnostic Instruments

- *AIMSweb (Academic Information Management System)* web-based, curriculum-based measures and data management system (subscription based). Online at www.aimsweb.com
- *Assessment, Evaluation, and Programming System Second Edition (AEPS),* Brookes Publishing Company (criterion-referenced)
- *Battelle Developmental Inventory Second Edition (BDI-2)*, Riverside Publishing
- *Bayley Scales of Infant Development, Third Edition (Bayley-III)*, Harcourt Assessment (formerly The Psychological Corporation)
- *Intervention Central*, RTI resources including CBM warehouse. Free online at www.intervention-central.org
- *Kaufman Assessment Battery for Children (K-ABC)*, Pearson Assessments
- *Peabody Developmental Motor Scales, Second Edition (PDMS-2)*, PRO-ED
- *Preschool Language Scale, Fourth Edition (PLS-4)*, Harcourt Assessment (formerly The Psychological Corporation)
- *Social Skills Rating System (SSRS),* Pearson Assessment

Characteristics of Diagnostic Assessments
Table 4

	Developmental (Birth – Primary Grades)	Early Academic (PreK – Primary Grades)
Question(s) Asked	**General Questions** • What is the nature & severity of the problem? • Is this student eligible for special services?	**General Questions** • What is the nature & severity of the problem? • Is this student eligible for special services?
Results	**Numerical Scores, Standard Scores** • DQs, IQs, • Quartiles, percentiles	**Numerical Scores, Error Analysis** • Correct word sequences • Computation errors • Oral reading fluency rate • Writing samples
Type of Measures	**Norm Referenced** • Compares child's performance to a representative group of same age peers	**Criterion Referenced** • Curriculum-based measures; compares a student's performance to the "typical" performance in the local curriculum (may be derived from local norms)
Who Administers?	Most often administered by health and education professionals, often specialists.	Most often administered by education professionals.
Administration	**"Narrow & Deep"** • Designed to obtain specific information on a suspected delay; may cover all major content areas or be specific to suspected problems	**"Narrow & Deep"** • Designed to obtain specific information on a suspected problem
Technical Adequacy	**Check for evidence of:** • Reliability and validity	**Check for evidence of:** • Reliability and validity
Sample Instruments	• Comprehensive norm-referenced instruments	• Specific level curriculum-based measures in content areas

Additional Readings and Resources

Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8. National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education (2003). Joint position statement. Washington, DC: National Association for the Education of Young Children.

Early Childhood Measures Profiles. D. Berry, L. Bridges, & M. Zaslow (2004). Child Trends. Online: http://aspe.hhs.gov/hsp/ECMeasures04/report.pdf

Evaluation and Assessment in Early Childhood Special Education: Children Who Are Culturally & Linguistically Diverse. Office of Superintendent of Public Instruction (1997). Online: http://www.k12.wa.us/SpecialEd/pubdocs/CLD.doc

Validity and reliability of a curriculum-based assessment approach to determine eligibility for Part C services. Macy, M. G., Bricker, D. D., & Squires, J. K. (2005). *Journal of Early Intervention, 28*, 1-16.

Using curriculum-based assessment to determine eligibility: Time for a paradigm shift? McLean, M. (2005). *Journal of Early Intervention, 28*, 23-27.

Principles and Recommendations for Early Childhood Assessments. Shepard, L., Kagan, S., & Wurtz, E. (1998). Washington, D.C.: National Educational Goals Panel. Online: http://www.negp.gov

Recognition and response: An early intervening system for young children at risk for learning disabilities. Executive summary. Coleman, M. R., Buysse, V., & Neitzel, J. (2006). Chapel Hill: The University of North Carolina at Chapel Hill, FPG Child Development Institute. Online: http://www.recognitionandresponse.org/

Screening and Assessment of Young English-Language Learners: Draft Recommendations. National Association for the Education of Young Children and the National Association of Early Childhood Specialists in State Departments of Education (January 2005). Joint position statement in supplement to *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8.*

Special Education: Rules for the Provision of Special Education to Special Education Students. Chapter 392-172a WAC (July, 2007). It can be accessed online: http://www.k12.wa.us/SpecialEd/pubdocs/wac/WAC_392-172a.doc

The words we use: A glossary of terms for early childhood education standards and assessments. Council of Chief State School Officers (2004). Retrieved December 14, 2006. http://www.ccsso.org/eceaglossary.

Using Response to Intervention for Washington's Students, includes guide, Powerpoint presentations. Office of Superintendent of Public Instruction (2006). Online: http://www.k12.wa.us/SpecialEd/RTI.aspx

Washington State Infant Toddler Early Intervention Program. Online: http://www1.dshs.wa.gov/iteip/sicc1.html

Program Evaluation and Accountability Assessments

- Interventionists are frustrated about home visits. **How many no-shows and cancellations are there, and why aren't families participating?**
- I want to choose the best preschool for my son. **How can I judge the quality of an early childhood program?**
- Our district is trying to decide whether to offer full day kindergarten. **Does full day kindergarten really produce better results?**
- We have to report child progress data to our funding agency. **How do we measure child progress toward early academic and social goals**?
- **How can we collect the data we need to demonstrate that children are learning at grade level, and identify how many are falling behind?**

Previous sections of this guide have focused on various types and purposes of assessment to answer questions and make decisions about the development and learning of individual children. Administrators, policymakers, and the general public also need information about entire groups of children at program, state, and national levels. Program personnel collect data on program operations and outcomes to inform improvements. School districts are required to report student progress throughout the district and across the curriculum. Increasingly, reports about child and family outcomes are being required by funding sources and made available to the general public. Policymakers at all levels are asking for program-level outcome data to inform funding decisions.

What is Program Evaluation and Accountability Assessment?

Program evaluation and accountability assessments are not identical but share many characteristics and practices, so they are addressed together for purposes of this manual. *Both involve systematic procedures to describe services and measure outcomes for* groups *of young children and their families.* Program evaluation and accountability assessments can be relatively small-scale endeavors focused on specific elements and operation of a classroom, comprehensive state and national assessments of student learning in a particular subject area, or anything in between. Both types of assessment may include measures of child learning and development outcomes.

Program evaluation assessments answer formative questions about the overall quality of programs, accomplished through careful descriptions of service components, participants, and resources. Program evaluations are conducted to support program quality improvement efforts. They are often required by public funding agencies to document continuous program improvement and sufficient progress in meeting the goals and objectives of the program.

Common examples of program evaluations are the annual reviews of Even Start, ECEAP, and Head Start to document child and family participation, describe program operations, and measure progress in early development and parenting skills. Similarly, child care services are reviewed at the program level for licensure and accreditation. Program evaluation results are commonly reported as indicators of quality, to answer questions about effectiveness: "Does this program work? Is it achieving the results it is supposed to?" Ideally, results of program evaluation assessments are used to document specific improvements that will increase the effectiveness of services in meeting program goals.

Accountability assessments emphasize gathering summative outcome data on child development or student learning for the purpose of providing information about the performance of an entire program. An example of accountability assessment is annual Washington Assessment of Student Learning (WASL) testing and the posting of scores in local newspapers. Comparison of local WASL mean scores to other schools and statewide averages is intended to inform the public about educational outcomes achieved school by school. Accountability measures are currently also required for federal and state funded programs, such as the:

- Federal Office of Special Education Programs (OSEP) outcome reporting requirement for infants, toddlers, and preschoolers served under Part C and Part B/Section 619 of the Individuals with Disabilities Education Act (IDEA) 2004;
- Head Start requirements for reporting preschool outcomes relative to the Child Outcomes Framework;
- Federal programs under the Elementary and Secondary Education Act; No Child Left Behind (NCLB); and
- State Accountability Measures for student achievement and performance.

Purposes of Program Evaluation and Accountability Assessment

Program evaluation is conducted primarily to support continuous improvement of programs. Information gathered during program evaluation assessments is combined and reported for a variety of purposes, often to secure additional funding, or identify components that need improvement.

For example, surveys are popular for documentation of parent/family satisfaction with early childhood programs. If the results of a preschool survey indicate that parents of children with special needs are consistently less satisfied than parents of typically developing children, the program could design interviews or focus groups with parents and use the data to show that a special education teacher should be added to the staff. Similarly, a child care program might interview parents to determine the extent of need for early morning and later evening hours of operation, and use the information to inform decisions about staffing and scheduling.

Teachers might collect group data on progress toward early literacy goals, and report the results to support the effectiveness of a newly implemented literacy-based classroom environment. Administrators often describe existing services and program operations in light of needs assessment data, to document unmet needs and request funding for additional services.

For example, a primary school principal could use data documenting reading difficulties in 15% of first graders to support a funding request for a reading specialist and/or reading coaches in the building. In each example, program evaluation data are collected at the program level to present group responses and reported internally for purposes of program improvement.

Accountability assessments are also conducted at the group level, but are used to hold whole programs and systems of services responsible for results. Accountability assessments are usually designed to inform external funding agencies, regulatory bodies, and ultimately the children and families being served, about the relative effectiveness of a program's services. Tax payers and policymakers use program accountability data to inform decisions to continue or enhance funding of programs, redirect funds or discontinue funding.

Research supports a strong connection between program quality and child outcomes, and for this reason program evaluation and accountability assessments go hand in hand. Assessment of individual children is conducted for program evaluation and accountability only if the data are to be consolidated and aggregated. In all cases, information is used to make program level decisions, rather than decisions about individual children or families.

Characteristics of Program Evaluation and Accountability Assessments

The content, administration, and interpretation of program evaluations reflect the questions at the beginning of this section, all involving program-level operations and outcomes. Program evaluations describe and measure both the quantity and the quality of services provided to young children and families.

- Quantity measures include factors such as number of service hours, education and training of personnel, amount of space, schedule of operations, and components of service.
- Quality measures tend to focus on the results of services for children and families, such as developmental and early learning outcomes, family functioning measures, ratings of the physical environment, and satisfaction with services.

A variety of qualitative and quantitative methods are used to evaluate programs: focus groups, interviews, surveys, review of records, observations, and direct testing. The methods used depend on the services offered, the particular aspects of program delivery being evaluated, and the questions being asked. Teachers use program evaluation techniques all the time to improve services to children and families.

For example, a group of early interventionists might track the rate of toddler group no-shows and cancellations for a month, and then interview parents with the best and worst attendance to compare group benefits and disadvantages. In similar fashion, movement of children throughout the center might be mapped for a short period of time to determine if furniture should be arranged to improve universal accessibility and decrease running.

Qualitative methods such as focus groups, observations, and interviews are typically used to gather information about the *processes* of program operations, such as identification of services offered, and the

actual workings of referral and intake, assessment system, curriculum design, professional development, parent support, and interagency collaboration.

For example, a few parents are asked to describe their experiences with the intake process and their responses are compared with records, program brochures, and staff descriptions. The collective information is analyzed and interpreted to determine congruence between written procedures and actual practices during intake and assessment, and the alignment between assessment and curriculum development. This type of program evaluation is used to describe what happens during a particular aspect of program operation, and how it happens, which in turn frames a context for child and family outcome data.

Direct measures of progress in meeting child and family outcomes are most often used for purposes of accountability. Accountability assessments look at *products*, or results, of early childhood programs, using direct measures of child knowledge and skills, as well as parent knowledge and behavior. A wide variety of structured observations and individual assessments are employed to collect data on child health, development, and academic status, parent-child interactions, and social interactions. *Adults collecting child outcome data should always know exactly why they are assessing youngsters, and how the data will be used.*

Three specific types of program evaluation/accountability measures have been described by Shepard, Kagan, and Wurtz (1998): 1)physical characteristics; 2)social indicators; and 3)direct measures of learning.

First, and easiest to measure, are *physical characteristics* of children, families, and services. Physical characteristics (like birth weight and average family size) are easy to count and can provide powerful comparisons over time and across programs. A family support program, for example, can use data showing clear trends toward smaller family sizes and higher birth weights as measures of success for family planning and pre-natal care services.

Social indicators are indirect measures that describe characteristics of communities, services, and families that are known to be related to early development and learning. For example, the number of families living with poverty, the availability of health insurance and publicly funded preschools, immunization rates, and access to mental health services are all variables that are known to have either a positive or adverse effect on early learning and development.

A third type of measure is *direct assessment* of learning and behavior outcomes, which requires direct testing of children and aggregation of scores into group result(s). Direct measures are most often used for accountability assessment and tracking progress of school-age children. Accountability assessment is more efficient, effective, and feasible once children reach kindergarten because schools provide ready-made populations of children from which to draw samples for testing.

Recommended Practices in Program Evaluation and Accountability Assessment

The challenge for administrators and others who are responsible for design and outcomes of early childhood programs is to collect enough data to accurately represent group outcomes within a reasonable structure and timeline. Program evaluation data can inform accountability assessments, and vice versa, but the data are not interchangeable. In this section, each recommended practice is designated as regarding program evaluation, accountability assessment, or both.

1. **Program goals provide the best starting point for both program evaluation and accountability assessment.** The goals a program is striving to achieve will determine the desired outcomes, which in term will guide selection of evaluation and accountability measures. For example, an early learning program that works to provide quality out-of-home care for young children of working parents might evaluate the extent to which the schedule of operations matches parents' work schedules. Another early learning program that claims to prepare children for kindergarten, on the other hand, would collect measures of early literacy and math achievement and compare outcomes from program exit to kindergarten entry benchmarks. Accountability measures in primary grades will be directly related to curriculum benchmarks and grade-level expectations.

2. **A good program evaluation starts with identification of all stakeholders as both respondents and as audience for the results.** At the program level, the strongest and most defensible results are obtained by collecting objective input from multiple perspectives. Those who provide and receive services, for example, have entirely different viewpoints on the same service component. Input that reflects the divergent views of parents, providers, administrators, community partners, and children themselves (when appropriate) is the best insurance that program evaluation data are objective and representative. Sharing results internally, as well as externally, and involving all stakeholders in analysis and discussion of program evaluation data provides a feedback loop that is inclusive, engaging, and meaningful for everyone involved.

3. **All program evaluation data, including intended and unintended results, should be used to inform continuous program improvements.** Early childhood professionals work very hard to provide the best services for young children, and program evaluation results that reflect success are usually anticipated and easy to present to others. A good program evaluation, however, is comprehensive in addressing all program components and services with equal attention, and often uncovers areas for improvement. The best results, negative or positive, are those that are used to maintain and improve effective service delivery. The process can be threatening at first, but after a few opportunities to interpret program evaluation data, it becomes very motivating and comforting to have systematic data to support program-level decision making. In addition, tracking program evaluation results over time becomes a powerful method for documenting ongoing program improvements.

4. **Accountability assessment data should** always **be reported as aggregated (collected and consolidated) scores, and the results should** never **be used to make decisions about individual children.** To minimize the impact of inappropriate testing for young children, systematic sampling procedures should be used to identify subsets of children for each data collection point. Matrix

sampling procedures are considered ideal for selecting subgroups of children and/or identifying sections of the instrument to be administered to each child.

5. **The use of norm-referenced, standardized tests should be avoided as the only measure of accountability for preschool programs, and steered clear of entirely for infants and toddlers.** The early childhood profession has voiced many valid concerns about administration of individual, norm-referenced test to young children. Standardized testing for accountability in early childhood programs is generally used to efficiently measure early academic knowledge, and evokes the tension that currently exists around the goals of child care and preschool programs. While there is widespread acknowledgement that academic preparation is an important aspect of preschool, there is also concern that the field's notable holistic and comprehensive view of early learning and development will be lost if accountability measures are limited to early academic outcomes. *When formal measurement tools are used, it is essential that trained professionals are responsible for assessment.* For program evaluation, it is desirable to emphasize the progress children make over time, rather than point-in-time normative comparisons.

6. **A logic model is a well-accepted method of designing program evaluations.** A logic model starts with a description of each separate program component and associated goals. A logical analysis of component goals leads to hypotheses about anticipated outcomes, and the evaluation design seeks to capture progress toward the outcomes. The logic model has the advantage of easy graphic representation and organization of the program evaluation design.

7. **Measures used in accountability assessment and program evaluation should be unbiased and proven instruments with supporting studies of reliability and validity.** Technical adequacy of the actual measures used to collect data is important to ensure objectivity and accuracy of results. There is little sense in implementing a well-designed plan for program evaluation or accountability assessment unless the measures used to generate the data are trustworthy. All measures and the methods used to collect information should be free from cultural bias and inclusive of stakeholders who speak languages other than English.

8. **Adequate support for training, technical assistance, and other professional development is necessary for successful program evaluation and accountability assessment.** Logic models, technical adequacy of measurement instruments, norm-referenced test administration, data analysis, and report writing are not commonly included in many early childhood preparation programs. These concepts and practices certainly are not intuitive! Teachers and service providers at all levels are increasingly collecting direct measures of student learning and development, but often the data are not put to good use. Classroom personnel and early intervention service providers need appropriate training to administer norm-referenced tests, and should never be asked to collect data without a clear understanding of how results will be disseminated and used. Program evaluation and accountability assessment require investments of time, energy, and resources, most of which take time away from curriculum design and instruction. Teachers, parents, service providers and specialists should ideally be educated about the goals of program evaluation and accountability assessment, and involved in decisions about methods and instruments.

General Caveats

If you are a teacher, early interventionist, or child care provider, you might be thinking that assessment for program evaluation and accountability is not all that straightforward. If you have read this far, you could also have a vague feeling of unease about possible misuse of such large amounts of data. Both are valid concerns.

Program evaluation and accountability assessments are complex and ideally conducted by objective third parties with no investment in the results. In reality, early childhood professionals are often responsible for collecting program evaluation and accountability data on their own programs. It is difficult (and probably undesirable) to be objective and neutral about your life's work, and usually classroom personnel and early intervention service providers would rather be interacting with, teaching, and otherwise supporting the development of young children.

It is worth mentioning that group data almost never look like any individual child's data, because the results combine the most skilled and the least skilled and everyone in between into a single score. No matter how accurate group data are, the information does not actually describe or match any of the children you know. Instead, the data reflect one single composite child, who of course does not exist. This is not a problem for accountability or program evaluation, but it is why we do not make individual decisions from group data.

Potential Instruments for Program Evaluation

Accountability assessment instruments for measuring child outcomes are usually designated by the funding agency or regulatory body that requires the outcome information. A wide variety of instruments is currently in use to measure child outcomes, from progress rating scales to norm-referenced tests of vocabulary knowledge. Specific instruments selected for accountability measures will depend on program goals, children's ages, sampling plans, and personnel resources. The complexity and high-stakes decisions associated with accountability assessments accounts for the fact that most successful efforts are organized and managed at the state or federal level. Therefore, there is no listing of accountability instruments here.

The instruments listed here are useful measures of specific interactional and environmental components that might be included in comprehensive program evaluations. The Additional Resources section that follows may provide more guidance for readers who are looking for direction in designing and utilizing program evaluation assessments.

A Few Useful Instruments for Program Evaluation

- *Classroom Assessment Scoring System* (CLASS), University of Virginia Press.
- *Early Childhood Environmental Rating Scale Revised Edition* (ECERS-R), Teachers College Press.
- *Early Childhood Classroom Observation Measure* (ECCOM), D. Stipek & P. Byler, Stanford University School of Education.
- *Early Language and Literacy Classroom Observation Tool* (ELLCO), Brookes Publishing.
- *Infant Toddler Environment Rating Scale Revised Edition* (ITERS-R), Teachers College Press.
- *School Age Care Environment Rating Scale* (SACERS), Teachers College Press.
- *Supports for Early Literacy Assessment* (SELA), National Institute for Early Education Research (NIEER).

Teachers and service providers at all levels are increasingly collecting direct measures of student learning and development, but often the data are not put to good use.

Characteristics of Accountability & Program Evaluation Assessments
Table 5

	Program Evaluation	Accountability
Question(s) Asked	**General Questions** • What are the major components of the program? • To what extent are children and families achieving desired outcomes? • How satisfied are parents with various components of the program?	**General Questions** • Is the program effectively serving children and families? • Is the program having a positive impact on student learning?
Results	**Descriptive Data** • Program improvement recommendations • Program descriptions • Focus group summaries • Goal discrepancy analyses • Cost/service unit	**Program Level Data** • Program improvement recommendations • Group developmental outcomes in social/behavior, adaptive and cognitive areas • Grade-level academic outcomes in content areas
Type of Measures	**Qualitative Measures** • Interviews • Observations • Focus groups **Quantitative Measures** • Satisfaction surveys • Child & family outcome measures	**Direct Outcome Measures** • Large-scale, standardized group assessments of child and family outcomes • Benchmark assessments at program entry and exit
Who Administers?	Ideally administered by external evaluators, third party consultants; often required by funding sources and regulatory agencies	Ideally administered by external evaluators, third party consultants; often required by funding sources and regulatory agencies
Administration	**Broad-based, Group Focus** • Designed to obtain input and feed-back from a broad spectrum of stake-holders (e.g. parents, administrators, staff, and collaborating agencies)	**Group Focus** • Designed to obtain composite information on performance of groups of children, families, staff
Technical Adequacy	**Work to ensure:** • Reliability and validity of surveys, outcome measures • Unbiased procedures for interviews, focus groups	**Work to ensure:** • Reliability and validity of outcome measures • Representative sampling procedures

Additional Readings and Resources

Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8. National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education (2003). Joint position statement. Washington, DC: National Association for the Education of Young Children.

Early Childhood Measures Profiles. D. Berry, L. Bridges, & M. Zaslow (2004). Child Trends. Online: http://aspe.hhs.gov/hsp/ECMeasures04/report.pdf

Considerations related to developing a system for measuring outcomes for young children with disabilities and their families. Early Childhood Outcomes Center (2004). Online: http://www.fpg.unc.edu/~eco/pdfs/considerations.pdf

Principles and Recommendations for Early Childhood Assessments. Shepard, L., Kagan, S., & Wurtz, E. (1998). Washington, D.C.: National Educational Goals Panel. Online: http://www.negp.gov

Screening and Assessment of Young English-Language Learners: Draft Recommendations. National Association for the Education of Young Children & National Association of Early Childhood Specialists in State Departments of Education (January 2005). Joint position statement in supplement to *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8.*

Taking Stock: Assessing and Improving Early Childhood Learning and Program Quality. Report of the National Early Childhood Accountability Task Force. October, 2007.

The words we use: A glossary of terms for early childhood education standards and assessments. Council of Chief State School Officers (2004). Retrieved December 14, 2006. http://www.ccsso.org/eceaglossary.

References for Part I

American Educational Research Association. (2004). *Position statement concerning high-stakes testing in preK-12 education.* Washington, DC: Author.

American Federation of Teachers. (2002, December). *At the starting line: Early childhood education programs in the 50 states.* Washington, DC: Author.

Berry, D., Bridges, L., & Zaslow, M. (2004). *Early Childhood Measures Profiles.* Online: http://aspe.hhs.gov/hsp/ECMeasures04/report.pdf.

Bowman, B., Donovan, S., & Burns, S. (Eds.). (2001). *Eager to learn: Educating our preschoolers.* Washington, DC: National Academy Press.

Bredekamp, S. & Copple, C. (1997). *Developmentally appropriate practice in early childhood programs.* Washington, DC: National Association for the Education of Young Children.

Coleman, M. R., Buysse, V., & Neitzel, J. (2006). *Recognition and response: An early intervening system for young children at risk for learning disabilities. Executive summary.* Chapel Hill: The University of North Carolina at Chapel Hill, FPG Child Development Institute. Online: http://www.recognitionandresponse.org/.

Council of Chief State School Officers. (2004). *The words we use: A glossary of terms for early childhood education standards and assessments.* Retrieved May 11, 2004: www.ccsso.org/eceaglossary.

Framework for Achieving the Essential Academic Learning Requirements in Reading, Writing, and Communication: Birth to Five. Online: http://www.k12.wa.us/CurriculumInstruct/pubdocs/birth-to-5.pdf.

Gettinger, M. (2001). Development and implementation of a performance-monitoring system for early childhood education. *Early Childhood Education Journal, 29*, 9-15.

Glascoe, F. P. (2001). Can teachers' global ratings identify children with academic problems? *Journal of Developmental & Behavioral Pediatrics, 22*, 163-168.

Goals 2000: Educate America Act, Pub. L. 103-227 (1994).

Grisham-Brown, J. (in press). Standards in early childhood education. In *Best Practices in School Psychology V.* Bethesda, MD: National Association of School Psychologists.

Head Start Bureau. (2001). *Head Start child outcomes framework.* Head Start Bulletin, no. 70. Washington, DC: Department of Health and Human Services, Administrator for Children and Families. Online: http://www.headstartinfo.org/pdf/im00_18a.pdf.

Hebbeler, K., Bailey, D., & Bruder, M. B. (October, 2006). Measuring family outcomes: A tool for program improvement. Paper presented at the 22nd Annual International Conference of the Division for Early Childhood (DEC) of the Council for Exceptional Children in Little Rock, Arkansas.

Hemmeter, M., Joseph, G., Smith, B., & Sandall, S. (Eds.). (2001). *DEC recommended practices program assessment: Improving practices for young children with special needs and their families.* Longmont, CO: Sopris West and Division of Early Childhood/Council for Exceptional Children.

Hosp, M., Hosp, J., & Howell, K. (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement.* New York: The Guilford Press.

Howell, K. W., & Nolet, V. (2000). *Curriculum-based evaluation: Teaching and decision making.* Belmont, CA: Wadsworth Thomson Learning.

Kagan, S. L., Britts, P. R., Kauerz, K., & Tarrant, K. (2005). *Washington State early learning and development benchmarks.* Olympia, WA: The State of Washington. Online: http://www.k12.wa.us/EarlyLearning/Benchmarks.aspx.

Kagan, S., Scott-Little, C., & Clifford, R. (2003). Assessing young children: What policymakers need to know and do. In *Assessing the state of state assessments: Perspectives on assessing young children,* (Eds.). Greensboro, NC: SERVE.

Macy, M. G., Bricker, D. D., & Squires, J. K. (2005). Validity and reliability of a curriculum-based assessment approach to determine eligibility for Part C services. *Journal of Early Intervention, 28*, 1-16.

McLean, M. (2005). Using curriculum-based assessment to determine eligibility: Time for a paradigm shift? *Journal of Early Intervention, 28*, 23-27.

Meisels, S. J., & Atkins-Burnett, S. (2005). *Developmental screening in early childhood: A guide.* Washington, DC: National Association for the Education of Young Children.

Meisels, S. J., & Fenichel, E. (Eds.). (1996). *New visions for the developmental assessment of infants and young children.* Washington, DC: Zero to Three/National Center for Infants, Toddlers, and Families.

National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education. (2002). *Early learning standards: Creating the conditions for success.* Joint position statement. Washington, DC: National Association for the Education of Young Children.

National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education. (2003). *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8.* Joint position statement. Washington, DC: National Association for the Education of Young Children.

National Association for the Education of Young Children, & National Association of Early Childhood Specialists in State Departments of Education. (January 2005). *Screening and Assessment of Young English-Language Learners: Draft Recommendations.* Joint position statement in supplement to *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8.* Washington, DC: National Association for the Education of Young Children.

National Child Care Information Center (2005, June). *Assessment and evaluation: Becoming an educated consumer. Part I: Child assessment.* Retrieved January 26, 2008, from http://nccic.acf.hhs.gov/pubs/goodstart/assess-eval1.pdf

National Child Care Information Center (2005, June). *Assessment and evaluation: Becoming an educated consumer. Part II: Program evaluation.* Retrieved January 26, 2008, from http://nccic.acf.hhs.gov/pubs/goodstart/ assess-eval2.pdf

National Child Care Information Center (2005, June). *Assessment and evaluation: Becoming an educated consumer. Part III: Accountability systems.* Retrieved January 26, 2008, from http://nccic.acf.hhs.gov/pubs/goodstart/ assess-eval3.pdf

National Research Council & Institute of Medicine. (2002). *From neurons to neighborhoods: The science of early childhood development.* Committee on integrating the science of early childhood development. Eds. J. Shonkoff & D. Phillips, Board of children, youth, and families, commission on behavioral and social sciences and education. Washington, DC: National Academy Press.

Neisworth, J. T., & Bagnato, S. J. (2004). The mismeasure of young children: The authentic assessment alternative. *Infants and Young Children, 17*, 198-212.

Office of Superintendent of Public Instruction. *Evaluation and assessment in early childhood special education: Children who are culturally & linguistically diverse.* Online: http://www.k12.wa.us/SpecialEd/pubdocs/CLD.doc.

Office of Superintendent of Public Instruction. *Using response to intervention for Washington's students.* Includes manual, Powerpoint presentations. Online: http://www.k12.wa.us/SpecialEd/RTI.aspx.

Rhode Island KIDS COUNT. (February 2005). *Getting ready: Findings from the national school readiness indicators initiative a 17 state partnership.* Providence, RI: Author.

Sandall, S., McLean, M., & Smith, B. (2000). *DEC recommended practices in early intervention/early childhood special education.* Longmont, CO: Sopris West.

Shepard, L., Kagan, S., & Wurtz, E. (Eds.). (1998). *Principles and recommendations for early childhood assessments.* Washington, D.C.:National Education Goals Panel.

Shonkoff, J. P., & Meisels, S. J. (Eds.). (2000). *The handbook of early childhood intervention* (2nd ed.). New York: Cambridge University Press.

Special Education: Rules for the Provision of Special Education to Special Education Students. Chapter 392-172a WAC (July, 2007). It can be accessed online: http://www.k12.wa.us/SpecialEd/pubdocs/wac/WAC_392-172a.doc

Squires, J., Nickel, R. E., & Eisert, D. (1996). Early detection of developmental problems: Strategies for monitoring young children in the practice setting. *Journal of Developmental & Behavioral Pediatrics, 17*, 420-427.

Taking Stock: Assessing and Improving Early Childhood Learning and Program Quality. Report of the National Early Childhood Accountability Task Force. October, 2007.

Washington State Infant Toddler Early Intervention Program. Online: http://www1.dshs.wa.gov/iteip/sicc1.html.

Part II: Compendium of Assessment Instruments

How do early childhood professionals make informed decisions about assessment tools and then select tools from among the thousands available? Identification of appropriate assessment instruments is an important consideration in the development of effective assessment systems for early childhood programs. Selecting an inappropriate assessment tool is analogous to using the wrong household tool, making the task at hand more difficult and producing a less desirable outcome.

Using a screening tool to inform instruction or monitor progress, for example, is something like using a screwdriver to pound a nail. The tool is simply not suited for the task, and more likely to produce confusion and frustration than to assist in building a quality structure.

This section of the guide presents summary and specific information for a large compendium of assessment instruments, and is presented as a companion to the narrative about assessment purposes and recommended practices included in Part I. Like all tools, assessment instruments are designed for specific purposes, and to be useful the tool needs to match the task. In order to make informed decisions about assessment instruments, professionals need to understand each assessment task by recognizing the exact purposes for which they are collecting information, and knowing how the resulting data will be used. For this reason, the instrument descriptions in this part of the guide will be most valuable for users who have read Part I.

Organization of the Compendium

Information in this part of the guide is organized to allow convenient access from multiple starting points: 1) a table of alphabetized listings by instrument name; 2) individual tool descriptions; and 3) tables of tools listed by age range.

Alphabetized Table of Tools (pages 75–88)

The first table lists assessment tools by name in alphabetical order, providing summary information on each tool, including:

- name
- age range of intended use
- purpose
- domains assessed by the tools
- languages in which the assessment is available

A page number on the left side of the table directs readers to the page that contains a more comprehensive description of each individual tool.

Purpose column designations

The purpose of individual tools is not as mutually exclusive as one might expect. Publishers often cite assessment instruments as serving multiple purposes, and terminology around diagnosis and monitoring of early academic problems is particularly complex and variable.

In the alphabetic table, tools are designated as both *diagnostic* and *informing instruction/monitoring progress* when they meet the following conditions:

- are norm-referenced
- are designed to be administered more than once per academic year
- provide specific information about skills that lend themselves to instruction (e.g. subtest title such as "bilateral coordination, agility, upper limb coordination", rather than simply "motor skills")
- provide concrete guidance on how to use results to develop IEPs/IFSPs and/or provide instruction

Tools for diagnosis of specific early academic problems are categorized for the purpose of informing instruction and monitoring progress. There are many useful curriculum-based measures that are not included in the tables because they are not available from commercial publishers. A large number of curriculum-based measures for early academics are available at:

1. *AIMSweb (Academic Information Management System)* web-based, curriculum-based measures and

   data management system (subscription based). Online at www.aimsweb.com
2. *Intervention Central*, RtI resources including CBM warehouse. Free online at www.intervention-central.org

Also please note the following:

- If a diagnostic tool includes a brief screening version (that is packaged with the larger diagnostic tool), it is designated in the alphabetical table under both *screening* and *diagnostic* columns.
- Tools are designated as useful for *monitoring progress* only if they can be administered at least three times per year and are brief enough to be administered regularly (e.g., < 30 minutes).
- Tools that assess approaches to learning are always designated for the purpose of *informing instruction*.

Domain column designations

- Tools that measure either *fine motor* or *gross motor* are designated with an **X** under motor domain.
- Tools that measure *self help* or *living skills* have an **X** in the column for the *adaptive* domain.
- Measures of *adaptability* will be designated as assessment in the *socio-emotional* domain.
- Program evaluation tools that measure characteristics of physical, social, and academic *environments* do not have a designation in any of the domain columns.

Individual Tool Descriptions (pages 89–164)

Each tool included in the alphabetized listing is described in more detail in this section, which makes up the bulk of Part II of this guide. Assessment tools are grouped by purpose and presented alphabetically by name within in four categories:

- screening
- informing instruction and monitoring progress
- diagnostic
- program evaluation

Tools designated in the alphabetical table as having two purposes (e.g., *screening* and *diagnostic*; *diagnostic* and *informing instruction/monitoring progress*) are described individually in the section that most closely matches their primary purpose. The left column in the alphabetical table at the beginning of Part II provides a page number that corresponds with the individual description.

Individual tool descriptions include the following information about each instrument:

- **name** of the assessment tool
- **age range** of children for whom the assessment instrument is designed
- length of **time** it takes **to administer** the assessment
- **training requirements** for the administrator, if any
- **languages** in which the assessment is available
- **source** of the assessment tool
- **administration type** (e.g., observation, direct assessment, etc.)
- **cost** of the assessment tool (most basic option)
- types of **scores available** (e.g., standard scores, referral cutpoints, etc.)
- **subscales**, or areas of assessment for which scores are available (e.g., word analysis, listening, reading comprehension, etc.)

- **norming sample** (characteristics of the group used to create the standard scores, if any)
- additional **technical information**, including reliability, concurrent validity, and other evaluation details

Assessment Tool Tables by Age Ranges (pages 165–175)

On these pages, the various assessment instruments are listed alphabetically in a grid format, based on the ages of children who can be evaluated by those instruments, as follows:

- **Infants & Toddlers:** Ages Birth to 2 Years 11 Months *(pages 158 - 159)*
- **Preschoolers:** Ages 3 Years to 4 Years 11 Months *(pages 160 - 162)*
- **Primary Grades:** Ages 5 Years and Older *(pages 163 - 165)*

Information Resources for Part II (pages 176–177)

This part of the guide contains resources that were used to develop the document and evaluate the tools included.

Appendix A (pages 179–180)

These pages include detailed information regarding the ratings system used to measure the reliability, concurrent validity, and sensitivity/specificity of the assessment tools in the individual tool descriptions.

Appendix B (pages 181–187)

Questions about early childhood curriculum, child assessment, and program evaluation are the foundation of this joint position statement, a seminal document from the National Association for the Education of Young Children (NAEYC) and the National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE).

How to Use the Compendium

A large number of assessment tools are described in this resource guide, and it is important that every user of the information is certain about the intended purpose(s) of assessment prior to selecting a particular instrument or instruments. Ideally, teams of professionals are reviewing and selecting tools. In any case, identification of appropriate tools requires careful attention to details of individual instrument descriptions.

There are at least three options for starting your search for specific assessment instruments:

- **By title** – If you already know the name of a tool and are interested in summary information, the alphabetical table is a good starting point. As noted earlier, the table gives an overview of:

  1) the purpose(s) for which the tool is intended;

  2) the ages of children for whom it is appropriate;

  3) the domains addressed by items; and

  4) publication languages.

  If this general information indicates that an instrument is potentially useful for your assessment purpose and population, the next step is to review the individual instrument description under the primary purpose of the instrument. A page number directs readers to the individual description for each tool listed.

- **By purpose** – Selecting a tool that matches the purpose of assessment is the most critical factor in finding the best instrument. Return to the narrative in Part I if you are uncertain of the purpose for which you are trying to identify an assessment instrument. Screening, for example, may be of most interest to child care providers and preschool teachers, while diagnostic tools might be more relevant for those who have responsibility for determining need for specific interventions or special services.

  Virtually all early childhood professionals will be assessing children for purposes of informing instruction and monitoring progress. Administrators will be particularly interested in tools for program evaluation and accountability.

  Use the table of alphabetized listings (pages 75–88) to identify a potential set of tools that might serve your assessment purpose by selecting those that align with your assessment purpose, the ages of the children, and the domains you are interested in assessing. Use the page number column to find the individual tool descriptions, to investigate in more detail those assessments you identify as potentially appropriate.

- **By age range tables** – For an overview of assessment tools that address the age range of children in your program, start with the age tables on pages 165–175. Identify those instruments that fit each assessment purpose, then find the page numbers for each tool by name on the alphabetical table at the beginning of Part II. The next step is to review the individual tool descriptions to determine which assessments you wish to utilize.

- **By individual tool descriptions** – The individual tool descriptions (pages 89–164) are useful when you are already familiar with a tool and want additional information, for instance checking available reliability and validity data to meet state or district guidelines. In most cases, additional information about individual tools will be necessary to make final decisions about selecting or changing instruments.

How Assessment Tools Were Selected

Assessment tools were included in this resource guide using the resources cited in Appendix A (pages 179–180). Well-known compendia of assessment instruments were reviewed and used to identify an initial set of instruments. Previous editions of tools were eliminated, as were those with a publication date of 1992 or earlier. Studies of individual instruments were also reviewed, most often to establish technical adequacy (reliability and validity information).

While we attempted to include instruments with basic reliability and validity information, often there was little or no information available. Users are cautioned about adopting assessment tools without an indication that scores obtained will be accurate and valid. In some cases (for example, Approaches to Learning), we included whatever tools we could identify, because very few tools are available.

A Note About School Readiness Assessments

We purposefully have not designated any tools specifically for school readiness assessment, despite the fact that many schools and districts engage in some type of assessment as children enter school as kindergarten students. As with all assessments, it is important that these tools be selected with care. First and foremost, the school needs to consider the reasons they are assessing entering kindergarteners. If the purpose of the assessment is to identify children who need additional supports or services, it makes most sense to choose a tool (or tools) designed for *screening* appropriate for kindergarten-aged children and covering a wide variety of domains.

At a minimum, the tool or tools selected should include cognitive, language, motor, and socio-emotional skills. If the school is assessing in-coming kindergartners in order to make instructional plans, then tools should be selected that are appropriate for *informing instruction* and are in line with the school's kindergarten curriculum.

As noted in the first section of this guide, kindergarten assessments should never be used to screen children out and delay kindergarten entry. Children who have the most to learn should have the opportunity to begin school as soon as possible, with the maximum instructional supports in place. Instead, if assessments are used as a regular part of kindergarten entrance, they should be selected to ensure that every child benefits maximally from the kindergarten experience.

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
121	Achenbach System of Empirically Based Assessment (ASEBA) - Preschool Forms	1:6 - 5 yrs			X N				X					X	75 lang. avail.
122	Achenbach System of Empirically Based Assessment (ASEBA) - School Aged Forms	6 -18 yrs			X N									X	75 lang. avail.
122	Adaptive Behavior Assessment System 2nd Ed. (ABAS-II)	0 - 89 yrs		X N	X N			X	X			X	X	X	Eng Spa
93	Ages & Stages Questionnaire (ASQ)	0:4-5:0 yrs	X N					X	X			X	X	X	Eng Spa Fr Kor
93	Ages & Stages Questionnaire: Social Emotional (ASQ:SE)	0:6 - 5:0 yrs	X N						X				X	X	Eng Spa
123	Arizona Articulation Proficiency Scale 3rd Ed.	1:6 - 18 yrs			X N				X						Eng
105	Assessment, Evaluation, and Programming System (AEPS) for Birth to Three Years, 2nd Ed.	0 - 3 yrs		X C				X	X			X	X	X	Eng
105	Assessment, Evaluation, and Programming System (AEPS) for Three to Six Years, 2nd Ed.	3 - 6 yrs		X C				X	X			X	X	X	Eng
123	Battelle Developmental Inventory 2nd Ed. (BDI-II)	0:0 - 7:11 yrs			X N			X	X			X	X	X	Eng Spa
94	Battelle Developmental Inventory -Screening Test	0:0 - 7:11 yrs	X N					X	X			X	X	X	Eng Spa
124	Bayley Scales of Infant and Toddler - Development, 3rd Ed. (Bayley-III)	0:1 - 3:6 yrs			X N			X	X			X	X	X	Eng

Age Ranges when possible are represented as Years:Months (e.g. 2:6 indicates 2 yrs, 6 mos)
C = Criterion referenced = comparison to established measure or performance
N = Norm-referenced = comparison to performance of others

Languages: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang
Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali
Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
94	Bayley Scales of Infant and Toddler Development 3rd Ed. (Bayley-III) Screening Test	0:1 - 3:6 yrs	**X** **N**					**X**	**X**			**X**			Eng
124	Behavior Assessment System for Children, 2nd Ed. (BASC-2)	2 yrs - college			**X** **N**		**X**							**X**	Eng Spa
124	Behavioral and Emotional Rating Scale 2nd Ed. (BERS-2)	5 - 18 yrs			**X** **N**									**X**	Eng
125	Bilingual Verbal Ability Test (BVAT) Normative Update	5 yrs -adult			**X** **N**				**X**						18 lang. avail.
125	Boehm Test of Basic Concepts, Preschool, 3rd Ed.	3:0 - 5:11 yrs		**X** **N**	**X** **N**			**X**	**X**	**X**	**X**				Eng Spa
126	Boehm Test of Basic Concepts 3rd Ed.	K - 2nd grade		**X** **N**	**X** **N**			**X**	**X**	**X**	**X**				Eng Spa
127	Bracken Basic Concept Scale - Revised (BBCS-R)	2:6 - 8 yrs		**X** **C/N**	**X** **C/N**			**X**	**X**		**X**				Eng Spa
95	Brief Infant Toddler Social Emotional Assessment (BITSEA) *(Screening version of the ITSEA)*	1 - 3 yrs	**X** **N**											**X**	Eng Spa
106	Brigance Comprehensive Inventory of Basic Skills - Revised (CIBS-R)	Pre-K - 6th grade		**X** **C/N**	**X** **C/N**			**X**	**X**	**X**	**X**				Eng Spa
106	Brigance Diagnostic Inventory of Early Development II (IED-II)	0 - 7 yrs		**X** **C/N**	**X** **C/N**			**X**	**X**	**X**	**X**			**X**	Eng Spa
95	Brigance Early Preschool Screen II	2:0 - 2:11 yrs	**X** **C/N**					**X**	**X**			**X**	**X**	**X**	Eng Spa

AGE RANGES when possible are represented as YEARS:MONTHS (e.g. 2:6 indicates 2 yrs, 6 mos)
C = CRITERION REFERENCED = comparison to established measure or performance
N = NORM-REFERENCED = comparison to performance of others

LANGUAGES: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
96	Brigance Infant & Toddler Screen	0 - 1:11 yrs	**X** C/N					**X**	**X**			**X**			Eng Spa
96	Brigance K & 1 Screen II	K - 1st grade	**X** C/N					**X**	**X**	**X**		**X**	**X**	**X**	Eng Spa
97	Brigance Preschool Screen II	3 - 4 yrs	**X** C/N					**X**	**X**			**X**	**X**	**X**	Eng Spa
127	Bruininks-Oseretsky Test of Motor Proficiency, 2nd Ed. (BOT-2)	4 - 21 yrs		**X** N	**X** N							**X**			Eng
107	Carey Temperament Scales (CTS)	0:1 - 12 yrs		**X** N			**X**							**X**	Eng
107	Carolina Curriculum for Infants and Toddlers with Special Needs (CCITSN) 3rd Ed.	0:0 - 3:0 yrs		**X** C				**X**	**X**			**X**	**X**	**X**	Eng
108	Carolina Curriculum for Preschoolers with Special Needs (CCPSN), 2nd Ed.	2 - 5 yrs		**X** C				**X**	**X**			**X**	**X**	**X**	Eng
161	Classroom Assessment Scoring System (CLASS)	Pre-K - 3rd grade				**X** C									Eng
128	Clinical Evaluation of Language Fundamentals, 4th Ed. (CELF-4) *(See also the CELF-P 2)*	5 - 21 yrs		**X** N	**X** N				**X**						Eng Spa
128	Clinical Evaluation of Language Fundamentals - Preschool, 2nd Ed. (CELF-P 2) *(See also the CELF-4)*	3 - 6:11 yrs		**X** N	**X** N				**X**	**X**					Eng

AGE RANGES when possible are represented as YEARS:MONTHS (e.g. 2:6 indicates 2 yrs, 6 mos)
C = CRITERION REFERENCED = comparison to established measure or performance
N = NORM-REFERENCED = comparison to performance of others

LANGUAGES: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang
Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali
Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
129	Communication & Symbolic Behavior Scale Developmental Profile, First Normed Edition (CSBS DP)	0:6 - 6:0 yrs	**X** **N**		**X** **N**				**X**					**X**	Eng
129	Comprehensive Assessment of Spoken Language (CASL)	3 - 21 yrs			**X** **N**				**X**						Eng
130	Comprehensive Test of Phonological Processing (CTOPP) (*see also TOPEL)*	5 - 24:11 yrs			**X** **N**					**X**					Eng
130	Conners, 3rd Ed. (Conners 3) (*See also Conners Comprehensive Behavior Rating Scale [CBSR])*	6 - 18 yrs			**X** **N**									**X**	Eng Spa
131	Conners Comprehensive Behavior Rating Scales (CBRS) *(See also Conners, 3rd Ed.)*	6 - 18 yrs			**X** **N**									**X**	Eng Spa
108	Creative Curriculum Developmental Continuum for Ages 3 - 5 *(see also Creative Curriculum Developmental Continuum for Infants, Toddlers, and Twos)*	3 - 5 yrs		**X** **C**				**X**	**X**			**X**		**X**	Eng Spa
109	Creative Curriculum Developmental Continuum for Infants, Toddlers and Twos *(see also Creative Curriculum Developmental Continuum for Ages 3 – 5)*	0 - 2:11 yrs		**X** **C**				**X**	**X**			**X**		**X**	Eng Spa
131	Developmental Assessment of Young Children (DAYC)	0 - 5:11 yrs			**X** **N**			**X**	**X**			**X**	**X**	**X**	Eng

AGE RANGES when possible are represented as YEARS:MONTHS (e.g. 2:6 indicates 2 yrs, 6 mos)
C = CRITERION REFERENCED = comparison to established measure or performance
N = NORM-REFERENCED = comparison to performance of others

LANGUAGES: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
97	Developmental Indicators for the Assessment of Learning, 3rd Ed. (DIAL-3) *(See also Speed DIAL)*	3 - 6:11 yrs	**X N**					X	X			**X**	X	X	Eng Spa
98	Developmental Observation Checklist System (DOCS)	0 - 6 yrs	**X N**					X	X			**X**		**X**	Eng
109	Developmental Reading Assessment 2nd Ed. (DRA 2)	K - 3rd grade		**X C**						**X**					Eng Spa
110	Devereux Early Childhood Assessment (DECA) Development Program	2 - 5 yrs		**X N**										**X**	Eng
132	Differential Ability Scales-II (DAS-II)	2:6 -17:11 yrs		**X N**	**X N**			**X**	X	X	X				Eng
110	Dynamic Indicators of Basic Early Literacy Skills, 6th Ed. (DIBELS-6) *(See also Individual Growth and Development Indicators [IGDIs] for infants and Toddlers and Get it Got It, Go! for preschoolers)*	K - 6th grade	**X C/N**	**X C/N**						**X**					Eng Spa
162	Early Childhood Environment Rating Scale – Revised (ECERS-R) *(see also FCCERS-R; ITERS-R, SACERS)*	2:6 - 5 yrs				**X C**									Eng Spa Fr Ger Hun Nor
162	Early Language and Literacy Classroom Observation Scale (ELLCO)	Pre-K- 3rd grade				**X C**			X	X					Eng

AGE RANGES when possible are represented as YEARS:MONTHS (e.g. 2:6 indicates 2 yrs, 6 mos)
C = CRITERION REFERENCED = comparison to established measure or performance
N = NORM-REFERENCED = comparison to performance of others

LANGUAGES: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
132	Early Reading Diagnostic Assessment (ERDA)	K - 3rd grade		**X** **N**	**X** **N**				**X**	**X**					Eng Spa
111	ECLS-K Approaches to Learning Sub-Scale	K - 1st grade		**X** **C**			**X**								Eng Spa Chin Lakota Hmong
133	Expressive One-Word Picture Vocabulary Test (EOWPVT) *(See also Expressive One-Word Picture Vocabulary Test [ROWPVT])*	2 - 18:11 yrs			**X** **N**				**X**						Eng
133	Expressive Vocabulary Test, 2nd Ed. (EVT-2) *(see also Peabody Picture Vocabulary Test)*	2:6 - 90 yrs			**X** **N**				**X**						Eng
98	Eyberg Child Behavior Inventory (ECBI) and Sutter-Eyberg Student Behavior Inventory-Revised (SESBI-R)	2 - 16 yrs	**X** **N**											**X**	Eng
163	Family Child Care Environment Rating Scale– Revised (FCCERS-R) *(see also ECERS-R; ITERS-R, SACERS)*	Infancy – school age				**X** **C**									Eng
99	FirstSTEp: Screening Test for Evaluating Preschoolers	2:9 - 6:2 yrs	**X** **N**					**X**	**X**			**X**	**X**	**X**	Eng
99	Fluharty Preschool Speech and Language Screening Test, 2nd Ed. (Fluharty-2)	3 - 6:11 yrs	**X** **N**						**X**						Eng

AGE RANGES when possible are represented as YEARS:MONTHS (e.g. 2:6 indicates 2 yrs, 6 mos)
C = CRITERION REFERENCED = comparison to established measure or performance
N = NORM-REFERENCED = comparison to performance of others

LANGUAGES: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
111	Get It, Got It, Go (Preschool IGDIs) *(See also Individual Growth and Development Indicators [IGDIs] for infants and Toddlers and Dynamic Indicators of Basic Early Literacy Skills [DIBELS] for elementary ages.)*	3 - 5 yrs	X C	X C				X	X						Eng Spa
134	Gilliam Autism Rating Scale - 2nd Ed. (GARS-2)	3 - 22 yrs			X N				X					X	Eng
134	Goldman-Fristoe Test of Articulation 2nd Ed. (GFTA-2) *(see also KLPA-2)*	2 - 21:11 yrs			X N				X						Eng
112	Hawaii Early Learning Profile (HELP)	0 - 3 yrs		X C				X	X			X	X	X	Eng Spa
112	Hawaii Early Learning Profile for Preschoolers (HELP for Preschoolers)	3 - 6 yrs		X C				X	X		X	X	X	X	Eng Spa
113	High/Scope Child Observation Record for Infants and Toddlers (COR-IT) *(See Also High/Scope Preschool COR)*	6 wks - 3 yrs		X C				X	X			X		X	Eng Spa
113	High/Scope Preschool Child Observation Record (COR) *(See Also High/Scope COR-IT)*	2:6 - 6 yrs		X C				X	X	X	X	X		X	Eng Spa
114	Individual Growth and Development Indicators (IGDIs) for Infants and Toddlers *(See also Get It, Got It, Go* for preschoolers and *Dynamic Indicators of Basic Early Literacy Skills [DIBELS] for elementary ages)*	0 - 3 yrs	X C	X C				X	X			X		X	Any

Age Ranges when possible are represented as Years:Months (e.g. 2:6 indicates 2 yrs, 6 mos)
C = Criterion referenced = comparison to established measure or performance
N = Norm-referenced = comparison to performance of others

Languages: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang
Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali
Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
135	Infant Toddler Developmental Assessment (IDA)	0 - 3 yrs			**X** **N**			**X**	**X**			**X**	**X**	**X**	Eng Spa
163	Infant Toddler Environment Rating Scale -Revised (ITERS-R) *(see also ECERS-R)*	0 - 2:6 yrs				**X** **C**		**X**	**X**	**X**		**X**	**X**	**X**	Eng Spa Ger Japan
135	Infant Toddler Sensory Profile *(see Sensory Profile)*	0 - 3 yrs			**X** **N**										Eng Spa
136	Infant Toddler Social Emotional Assessment (ITSEA) *(See also: BITSEA)*	1 - 3 yrs			**X** **N**									**X**	Eng Spa
100	Infant Toddler Symptom Checklist (ITSC)	0:7 - 2:6 yrs	**X** **C**					**X**	**X**			**X**	**X**	**X**	Eng
136	Iowa Test of Basic Skills (ITBS) Forms A & B	5 - 14 yrs			**X** **N**			**X**	**X**	**X**	**X**				Eng
137	Kaufman Assessment Battery for Children, 2nd Ed. (KABC-II)	3 - 18 yrs			**X** **N**			**X**							Eng Spa
137	Kaufman Brief Intelligence Test 2nd Ed. (KBIT-2)	4 - 90 yrs			**X** **N**			**X**							Eng Spa
138	Kaufman Survey of Early Academic & Language Skills (K-SEALS)	3:0 - 6:11 yrs			**X** **N**				**X**	**X**	**X**				Eng
138	Kaufman Test of Educational Achievement, 2nd Ed. (KTEA-II)	4:6 - 90+ yrs		**X** **N**	**X** **N**				**X**	**X**	**X**				Eng

AGE RANGES when possible are represented as YEARS:MONTHS (e.g. 2:6 indicates 2 yrs, 6 mos)
C = CRITERION REFERENCED = comparison to established measure or performance
N = NORM-REFERENCED = comparison to performance of others

LANGUAGES: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
139	Khan-Lewis Phonological Analysis 2nd Ed. (KLPA-2) *(see also GFTA-2)*	2 - 21 yrs			**X** **N**				**X**						Eng
139	Learning Accomplishment Profile-Diagnostic (LAP-D), 3rd Ed.	2:6 - 6:0 yrs			**X** **N**			**X**	**X**		**X**	**X**	**X**		Eng Spa
100	Learning Accomplishment Profile - Normed Screens (LAP-D Normed Screens)	3 - 5 yrs	**X** **N**					**X**	**X**			**X**			Eng Spa
140	Leiter International Performance Scale -Revised (LEITER-R)	2:0 -20:11 yrs			**X** **N**			**X**						**X**	Eng
141	Lindamood Auditory Conceptualization Test, 3rd Ed. (LAC-3)	5:0 -18:11 yrs			**X** **N**				**X**	**X**					Eng Spa
141	MacArthur-Bates Communicative Development Inventories (CDI), 2nd Ed.	0:8 - 3:1 yrs			**X** **N**				**X**						See tool descr.
142	Merrill-Palmer-Revised Scales of Development-Revised (M-P-R)	0 - 6:6 yrs			**X** **N**			**X**	**X**			**X**	**X**	**X**	Eng Spa
142	Metropolitan Readiness Test (MRT6) 6th Ed.	4 - 7 yrs			**X** **N**					**X**	**X**				Eng Spa
143	Mullen Scales of Early Learning (MSEL)	0 - 5:8 yrs			**X** **N**				**X**			**X**			Eng
114	Ounce Scale	0 - 3:6 yrs		**X** **C**				**X**	**X**			**X**		**X**	Eng Spa
143	OWLS: Listening Comprehension (LC) Scale and Oral Expression (OE) Scale *(See also OWLS WE)*	3 - 21:11 yrs			**X** **N**				**X**						Eng

AGE RANGES when possible are represented as YEARS:MONTHS (e.g. 2:6 indicates 2 yrs, 6 mos)
C = CRITERION REFERENCED = comparison to established measure or performance
N = NORM-REFERENCED = comparison to performance of others

LANGUAGES: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang
Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali
Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
144	OWLS: Written Comprehension (WE) *(See also OWLS LC/OE)*	5:0 -21:11 yrs			**X** **N**					**X**					Eng
101	Parents' Evaluation of Development Status (PEDS)	0 - 8 yrs	**X** **N**					**X**	**X**			**X**	**X**	**X**	Eng Spa Viet Hua Som Chin
144	Peabody Developmental Motor Scales 2nd Ed. (PDMS-2)	0 - 6:0 yrs		**X** **N**	**X** **N**							**X**			Eng
145	Peabody Individual Achievement Test -Revised, Normative Update (PIAT -R/NU)	5:0 -22:11 yrs			**X** **N**					**X**	**X**				Eng
145	Peabody Picture Vocabulary Test 4th Ed. (PPVT-4) *(see also Expressive Vocabulary Test)*	2:6 - 90 yrs			**X** **N**				**X**						Eng Spa
101	Pervasive Developmental Disorders Screening Test-II (PDDST-II)	1:0 - 4:0 yrs	**X** **C**						**X**					**X**	Eng Spa
115	Phonological Awareness and Literacy Screenings - Kindergarten (PALS-K) *(see also PALS PreK and PALS 1-3)*	5 yrs (K)	**X** **C**	**X** **C**						**X**					Eng
115	Phonological Awareness and Literacy Screenings - PreKindergarten (PALS -PreK)*(see also PALS K and PALS 1-3)*	4 yrs (pre-k)	**X** **C**	**X** **C**						**X**					Eng
116	Phonological Awareness and Literacy Screenings - 1-3 (PALS - 1-3) *(see also PALS K and PALS Pre-K)*	1st - 3rd grade	**X** **C**	**X** **C**						**X**					Eng

AGE RANGES when possible are represented as YEARS:MONTHS (e.g. 2:6 indicates 2 yrs, 6 mos)
C = CRITERION REFERENCED = comparison to established measure or performance
N = NORM-REFERENCED = comparison to performance of others

LANGUAGES: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
146	Phonological Awareness Test 2 (PAT-2)	5:0 - 9:11 yrs		X N	X N					X					Eng
146	Pictorial Test of Intelligence 2nd Ed. (PTI-2)	3 - 8 yrs			X N			X	X		X				Eng
147	Preschool and Kindergarten Behavior Scales, 2nd Ed. (PKBS-2)	3 - 6 yrs			X N									X	Eng Spa
147	Preschool Language Assessment Instrument, 2nd Ed. (PLAI-2)	3:0 - 5:11 yrs			X N				X						Eng
148	Preschool Language Scale, 4th Ed. (PLS-4)	0 - 6:11 yrs			X N				X						Eng Spa
148	Process Assessment of the Learner, 2nd Ed.: Diagnostic Assessment for Math [PAL-II Math] *(see also PAL-II Reading and Writing)*	K - 6th grade		X N	X N						X				Eng
149	Process Assessment of the Learner 2nd Ed.: Diagnostic Assessment for Reading and Writing [PAL-II Reading and Writing] *(see also PAL-II Math)*	K - 6th grade		X N	X N					X					Eng
116	Qualls Early Learning Inventory (QELI)	Pre-K - 1st grade		X C/N			X	x	X	X	X				Eng
149	Ready to Learn: A Dyslexia Screener	3:6 - 6:5 yrs	X N		X N			X	X	X	X	X			Eng
150	Receptive Expressive Emergent Language Scale, 3rd Ed. (REEL-3)	0 - 3 yrs			X N				X						Eng

Age Ranges when possible are represented as Years:Months (e.g. 2:6 indicates 2 yrs, 6 mos)
C = Criterion referenced = comparison to established measure or performance
N = Norm-referenced = comparison to performance of others

Languages: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang
Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali
Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
150	Receptive One-Word Picture Vocabulary Test (ROWPVT) *(See also Expressive One-Word Picture Vocabulary Test [EOWPVT])*	2 - 18:11 yrs			**X** **N**				**X**						Eng Spa
151	Scales of Independent Behavior-Revised (SIB-R)	0:3 - 80 yrs	**X** **N**		**X** **N**				**X**			**X**	**X**	**X**	Eng
164	School-Age Care Environment Rating Scale (SACERS) *(see also ECERS-R; FCCERS-R; ITERS-R)*	5 - 12 yrs				**X** **C**									Eng Fr Ger
151	Sensory Profile *(see also Infant Toddler Sensory Profile)*	3:0 - 10:0 yrs			**X** **N**										Eng Spa
152	Social Competence & Behavior Evaluation (SCBE-Preschool Edition)	2:6 - 6:4 yrs			**X** **N**									**X**	Eng Fr
102	Speed DIAL (Developmental Indicators for the Assessment of Learning) *(See also Developmental Indicators for the Assessment of Learning - DIAL)*	3 - 6.11 yrs	**X** **N**					**X**	**X**			**X**			Eng Spa
152	Stanford Achievement Test (SAT 10) 10th Ed.	K - 12			**X** **N**				**X**	**X**	**X**				Eng
153	Stanford-Binet Intelligence Scales for Early Childhood - 5th Ed. (Early SB5)	2:0 - 7:3 yrs			**X** **N**			**X**							Eng
117	Teacher Rating of Oral Language & Literacy (TROLL)	3 - 5 yrs		**X** **N**					**X**	**X**					Eng

AGE RANGES when possible are represented as YEARS:MONTHS (e.g. 2:6 indicates 2 yrs, 6 mos)
C = CRITERION REFERENCED = comparison to established measure or performance
N = NORM-REFERENCED = comparison to performance of others

LANGUAGES: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang
Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali
Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
								Health & physical development							
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Cognitive	Language	Literacy	Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
153	Temperament & Atypical Behavior Scale (TABS) Screener and Assessment Tool	0:11 -5:11 yrs	X N		X N									X	Eng
154	Test for Auditory Comprehension of Language, 3rd Ed. (TACL-3)	3:0 - 9:11 yrs			X N				X						Eng
154	Test of Early Language Development 3rd Ed. (TELD-3)	2:0 - 7:11 yrs			X N				X						Eng
155	Test of Early Mathematics Ability 3rd Ed. (TEMA-3)	3:0 - 8:11 yrs			X N						X				Eng
155	Test of Early Reading Ability 3rd Ed. (TERA-3)	3:6 - 8:6 yrs			X N					X					Eng
156	Test of Language Development - Primary, 4th Ed. (TOLD-P:4)	4:0 - 8:11 yrs			X N				X						Eng
156	Test of Phonological Awareness 2nd Ed. PLUS (TOPA-2+)	5 - 8 yrs			X N				X						Eng
156	Test of Preschool Early Literacy (TOPEL) *(see also CTOPP)*	3:0 - 5:11 yrs			X N				X	X					Eng
157	Test of Word Reading Efficiency (TOWRE)	6:0 -24:11 yrs		X N	X N					X					Eng
158	Toddler & Infant Motor Evaluation (TIME)	0 - 3:6 yrs			X N							X	X	X	Eng

Age Ranges when possible are represented as Years:Months (e.g. 2:6 indicates 2 yrs, 6 mos)
C = Criterion referenced = comparison to established measure or performance
N = Norm-referenced = comparison to performance of others

Languages: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang
Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali
Spa=Spanish • Viet=Vietnamese

Alphabetical Table of Tools with Summary Characteristics (refer to page number indicated at left for full descriptions of each tool)

	NAME	AGE	PURPOSE				DOMAINS ASSESSED IN THE TOOL								
Page Number		Age range for use of the tool	1. Screening	2. Inform instruction/ monitor progress	3. Diagnostic	4. Program Evaluation	Approaches to learning	Health & physical development			Math, numeracy	Motor	Adaptive	Social emotional	Tool available in languages indicated
								Cognitive	Language	Literacy					
117	Transdisciplinary Play Based Assessment 2nd Ed. (TPBA 2)	0 - 6 yrs		**X** **C**				X	X			**X**		**X**	Eng
158	Vineland Adaptive Behavior Scales 2nd Ed. (Vineland-II)	0 - 90 yrs			**X** **N**				**X**			X	X	X	Eng Spa
159	Wechsler Individual Achievement Test 2nd Ed. (WIAT-II)	4:0 - 85 yrs			**X** **N**				X	X	X				Eng
159	Wechsler Preschool & Primary Scale of Intelligence, 3rd Ed. (WPPSI-III)	2:6 - 7:3 yrs			**X** **N**			X	X						Eng
160	Woodcock-Johnson III NU Complete (WJ III NU)	2:0 - 90+ yrs			**X** **N**			X	X	X	X				Eng Spa
118	Work Sampling System (WSS)	3 yrs to 6th grade		**X** **C**			**X**	X	X	X	X	X		**X**	Eng Spa
118	Young Children's Achievement Test (YCAT)	4:0 - 7:11 yrs		**X** **N**					X	X	X				Eng Spa

AGE RANGES when possible are represented as YEARS:MONTHS (e.g. 2:6 indicates 2 yrs, 6 mos)
C = CRITERION REFERENCED = comparison to established measure or performance
N = NORM-REFERENCED = comparison to performance of others

LANGUAGES: Chin=Chinese • Eng=English • Fr=French • Ger=German • Hua = Huang Hun=Hungarian • Japan=Japanese • Kor=Korean • Nor=Norwegian • Som=Somali Spa=Spanish • Viet=Vietnamese

Individual Tool Descriptions

The next pages include a standard set of information about each of the tools listed on the chart (see preceding pages). The purpose of this section is to provide individuals who are selecting assessment tools some basic information about a considerable range of options. Information provided is not meant to be comprehensive. Rather, it is meant to give individuals who are selecting tools some ideas about which tools warrant further investigation.

Once potential tools have been identified from the individual descriptions, users may need to do further research on the tools of interest to learn how well each one fits their needs. For example, whenever possible we have reported reliability and validity information for the overall (or comprehensive) score provided by each tool. However, reliability and validity might be quite different for individual subscales. Users who are primarily interested in a particular subscale will need to find out about that subscale's psychometric properties before selecting it.

Likewise, early childhood assessments vary in the extent to which children with disabilities were included in the pre-testing and norming samples. Individuals who work with children with specific disabilities should be sure to consult the tools' technical manuals to determine the extent to which the tools they are considering are appropriate for the group they serve.

Descriptions of Information on the Following Tables

Name: Provides the tool's name, abbreviation, edition (if any) and year of publication. New editions are being published all the time, but to the extent possible we have included the most updated version of each tool.

Age Range: Indicates the ages of individuals that can be assessed with this tool. Whenever possible, the information is presented in years, with a colon separating years and months. For example: 3:6 indicates 3 years and 6 months.

Time to Administer: Indicates how long the entire battery typically takes to administer.

Administrator Required/Training Needed: Many assessment tools require specialized training to administer correctly. This section indicates the type of educational background and specialized training required.

Cost: Indicates the least expensive option for getting started with each tool. Generally, the price includes needed manuals and scoring forms. It doest not include the cost of computer scoring software or manipulatives, unless they are required for use.

Available Languages: If any part of a tool is available in a language other than English, it is listed here. However, often only one component is available in another language. For instance, a tool that has a parent-report component, might have a Spanish version of the parent form, but not of the entire measure. In that case Spanish would be listed. Note that all other information (e.g., norming sample, reliability, validity) refers to the English-language version of the tool. Readers should not assume that versions of the tool in other languages have these same psychometric properties. Also, it should be noted that sometimes versions in languages other than English are based on an earlier version of the tool.

Source: Provides contact information for learning more about the tool and purchasing it. Generally, this is also the tool's publisher.

Administration: Indicates if the tool is individually administered, group administered, or designed for classroom-level administration. It also indicates if the tool is completed by the parent, teacher, or examiner. In addition, when available it includes information about the number of items in the scale.

Scores Available: Indicates the types of scores obtained from each measure; for example, standard score percentile rank, stanines, age-equivalent, comparisons to cutpoints.

Subscales: Indicates all the subscales, subtests or domains included in the tool. Note, not all subscales are appropriate for all ages of children and not all available scores are available for each subtest.

Norming Sample: For norm-referenced tests, this section indicates how many individuals were included in the norming sample, along with information about the population which the sample represents. When available, information about the inclusion or exclusion of special populations is indicated here.

Reliability: Reliability refers to the extent to which the scores provided by the tool are stable; that is, the extent to which the scores consistently measure the same thing across items within the test, from testing to testing, and across examiners. For this document, reliability has been ranked as High, Adequate, or Low, based on the most comprehensive and general score available from the test. In general, users should seek tools with high reliability; although those with adequate reliability should be considered if the tool meets other needs of the testing situation. Appendix A contains specific information about how reliability (High, Adequate, and Low) was determined.

Concurrent Validity: Validity is the extent to which the tool measures what it is intended to measure. There are many different ways of measuring validity and readers should consult each tool's technical manual to find out how the assessment's authors validated their tools.

This document provides information on only one type of validity: concurrent. Concurrent validity refers to the extent to which the scores from this tool are related to scores from tools designed to measure the same skill, ability, or aptitude. *Concurrent validity is a useful indicator only if there are other tools that do a good job of measuring the construct of interest.* Even when such measures exist, each measure is designed to evaluate something slightly different, so we do not anticipate perfect associations between measures.

For this document, Concurrent validity has been designated has High, Adequate, or Low. In general, readers should select tools with higher concurrent validity; keeping in mind that concurrent validity will generally be low or not reported for tools that are designed to assess skills, aptitudes, or abilities that are not generally assessed by other tools. In addition, validity is generally lower for younger children. *Readers are reminded that concurrent validity only indicates how well this tool relates to other well-known tools.*

Appendix A contains specific information about how concurrent validity (High, Adequate, and Low) was determined. When available, concurrent validity is reported for tools designed to *inform instruction/ monitor progress*, *diagnostic*, and *program evaluation*. Concurrent validity is only reported for *screening tools* when *sensitivity/specificity* information is unavailable.

Sensitivity/Specificity (reported for *screening tools* only)**:** When considering tools that are designed for screening, the central question about their utility is: How well does the tool correctly differentiate between children who do and do not need further evaluation? For this reason, each screening tool's specificity and selectivity are rated (High, Adequate, or Low) when the information is available. Specificity (also called *true negative*) refers to the percentage of children who do not have a disability who are correctly not referred for further evaluation. Selectivity (also called *true positive*) is the percentage of children who do have a disability who are correctly identified by the screening tool. Appendix E contains specific information about how sensitivity/specificity (High, Adequate, and Low) was determined.

Note: *This section includes any other information that the authors of this guide felt would be useful in selecting measures for further review.*

Individuals who work with children with specific disabilities should be sure to consult the tools' technical manuals to determine the extent to which the tools they are considering are appropriate for the group they serve.

Collection of Screening Tools

Page	Name
93	Ages & Stages Questionnaire, 2nd Edition (ASQ)
93	Ages & Stages Questionnaire: Social Emotional, 2nd Edition (ASQ:SE)
94	Battelle Developmental Inventory, 2nd Edition, Screening Test
94	Bayley Scales of Infant and Toddler Development, 3rd Edition (Bayley-III) Screening Test
95	Brief Infant Toddler Social Emotional Assessment (BITSEA)
95	Brigance Early Preschool Screen II
96	Brigance Infant & Toddler Screen
96	Brigance K & 1 Screen II
97	Brigance Preschool Screen II
129	Communication and Symbolic Behavior Scale Developmental Profile, First Normed Edition (CSBS DP)*
97	Developmental Indicators for the Assessment of Learning, 3rd Edition (DIAL-3)
98	Developmental Observation Checklist System (DOCS)
110	Dynamic Indicators of Basic Early Literacy Skills, 6th Edition (DIBELS-6) #
98	Eyberg Child Behavior Inventory (ECBI) and Sutter-Eyberg Student Behavior Inventory-Revised (SESBI-R)
99	First STEp: Screening Test for Evaluating Preschoolers
99	Fluharty Preschool Speech and Language Screening Test, 2nd Edition (Fluharty-2)
111	Get It, Got It, Go (Preschool IGDIs) #
114	Individual Growth and Development Indicators (IGDIs) for Infants and Toddler #
100	Infant Toddler Symptom Checklist (ITSC)
100	Learning Accomplishment Profile - Normed Screens (LAP-D Normed Screens)
101	Parents' Evaluation of Development Status (PEDS)
101	Pervasive Developmental Disorders Screening Test – II (PDDST-II)
115	Phonological Awareness and Literacy Screenings- Kindergarten (PALS-K) #
115	Phonological Awareness and Literacy Screenings-PreKindergarten (PALS-PreK) #
116	Phonological Awareness and Literacy Screenings-1-3 (PALS-1-3) #
149	Ready to Learn: A Dyslexia Screener *
151	Scales of Independent Behavior - Revised (SIB-R) *
102	Speed DIAL (Developmental Indicators for the Assessment of Learning)
153	Temperament and Atypical Behavior Scale (TABS) Screener and Assessment Tool*

- Tools marked with a pound sign have two purposes: *screening* and *inform instruction/monitor progress*.

Their descriptions appear in the Inform Instruction/Monitor Progress section, because those are their primary purposes.

* - Tools marked with an asterisk have two purposes: *screening* and *diagnostic*.

Their descriptions appear in the Diagnostic section, because that is their primary purpose.

Ages & Stages Questionnaire, Second Edition [ASQ] (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:4 – 5:0 yrs	10 – 20 minutes	Parents/caregiver completes questionnaire. Scored by paraprofessional.	$199

Available Languages: English, Spanish, French, Korean

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624 (800) 638-3775; www.brookespublishing.com

Administration: The ASQ is composed of 19 questionnaires given in 2-6 month intervals starting at 4 months of age and continuing through 60 months of age. Parent/primary care-giver completes 30-item questionnaire about individual child or a professional completes questionnaire after a home visit. Questionnaires are scored by program-staff members.

Scores Available: Scores are compared to cut-points for referral.

Subscales: Communication, Gross Motor, Fine Motor, Problem Solving, Personal-Social.

Norming Sample: 2,328 children from both risk and non-risk populations whose families were educationally, economically, and ethnically diverse. More details about the sample may be available in the Technical Manual.

Reliability: High (.80 or higher)

Sensitivity/Specificity: Adequate sensitivity (65% - 79%); High specificity (80% or higher)

Ages & Stages Questionnaire: Social Emotional, Second Edition [ASQ:SE] (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:6 – 5:0 years	10 – 15 minutes	Parents/caregiver completes questionnaire. Scored by paraprofessional.	$149

Available Languages: English, Spanish

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624 (800) 638-3775; www.brookespublishing.com

Administration: Parents complete 19 to 33 item questionnaires at eight designated intervals between 6 and 60 months. Questionnaires are scored by program-staff members.

Scores Available: Scores are compared to cut-points for referral.

Subscales: Self-Regulation, Compliance, Communication, Adaptive Functioning, Autonomy, Affect, Interaction with People

Norming Sample: 3,014 children between 6 and 36 months, closely approximating the U.S. population with regard to income, parental education, and ethnicity (2000 U.S. Census).

Reliability: High (over .80)

Sensitivity/Specificity: Adequate sensitivity (65% - 79%); High specificity (80% or higher)

Battelle Developmental Inventory, Second Edition, Screening Test (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:0 – 7:11 years	10 – 30 minutes	Early childhood teachers, early interventionists, special educators, psychologists, heath professionals, and other professionals familiar with psychometric procedures.	$299

Available Languages: English, Spanish

Source: Riverside Publishing Company, 8420 Bryn Mawr Avenues, Chicago, IL 60631
(800) 767-8420; www.riverpub.com

Administration: Direct assessment and observation of individual child, plus interview with child's parents conducted by early childhood professional.

Scoring: Raw scores are calculated for each domain and compared to cut-points for -1.0, -1.5, and -2.0 standard deviations.

Subscales: Personal-social, adaptive, motor, communication, and cognitive ability

Norming Sample: 2,500 children closely matching the U.S. population with regard to age, sex, race/ethnicity, region and socioeconomic level (2001 U.S. Census). Children with disabilities were not included in the norming sample, but were included in reliability and validity studies.

Reliability: High (over .80)

Sensitivity/Specificity: Both High (over 80%)

Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III) Screening Test (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:1 – 3:6 years	15 – 25 minutes	Trained technicians or paraprofessionals can administer and score. Professional with training in educational or psychological assessment required for interpretation.	$299

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child

Scoring: For each subscale, scores are compared to cut-points indicating "at risk," "emerging" or "proficient."

Subscales: Cognitive, Receptive Communication, Expressive Communication, Fine Motor, Gross Motor

Norming Sample: 1,675 children representative of U.S. population in terms of gender, race-ethnicity, region, and parental education (2000 Census).

Reliability: High (over .80)

Sensitivity/Specificity: Information may be available in Technical Manual. Percent of misdiagnosed children was low, ranging from 2 to 8%.

Brief Infant Toddler Social Emotional Assessment [BITSEA] (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
1:0 – 3:0 years	7 – 10 minutes	Professional with training in testing needed for interpretation.	$99

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: 42 items completed by the parent and/or child care provider; items were drawn from the pool of ITSEA items.

Subscales: Problems, Competence**,** combined Problem and/or Competence scale

Scores Available: Cut-points indicating need for further assessment are provided with separate age bands for boys and girls.

Norming Sample: 600 children from 42 states, similar to U.S. population in terms of ethnicity, parent education, and region (2002 Census).

Reliability: High (.80 of higher)

Sensitivity/Specificity: Both High (80% of higher) (in a sample of children with and without autism)

Notes: Designed as a social-emotional development screening test.

Brigance Early Preschool Screen II (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:0 – 2:11 years	15 minutes	For use by teachers, paraprofessionals, therapists, nurses, physicians. No special training required.	$148

Available Languages: English, Spanish

Source: Curriculum Associates, Inc., 153 Rangeway Road, North Billerica, MA 01862
(800) 225-0248; www.curriculumassociates.com

Administration: Direct assessment of individual child. Includes separate forms for children 2:0 to 2:5 years and for children 2:6 to 2:11 years.

Scores Available: Total scores are compared to cut-points for suspected developmental delay and possible advanced development in 3-month increments. Separate cut-points are provided for children at risk because of psychosocial disadvantage. Scores can be converted to quotients, age equivalents, and percentiles.

Subscales: General Knowledge and Comprehension; Speech and Language, Gross-Motor Skills; Fine Motor Skills, Preacademic, Social-Emotion and Self-Help.

Norming Sample: Based on a subset of the data collected for the Brigance IED-II norm creation (no separate administration); about 75 children per form. More detailed information about the sample's demographic characteristics.

Reliability: High (.80 or higher)

Sensitivity/Specificity: Both High (80% or above) (with small samples of fewer than 30 children)

Note: Designed to identify children who need additional testing because they might have developmental problems or intellectual giftedness. Some of this norming, reliability, and sensitivity/specificity information may come from earlier versions.

Brigance Infant & Toddler Screen (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:0 – 1:11 years	10 – 20 minutes	For use by teachers, paraprofessionals, therapists, nurses, physicians. No special training required.	$148

Available Languages: English, Spanish

Source: Curriculum Associates, Inc., 153 Rangeway Road, North Billerica, MA 01862
(800) 225-0248; www.curriculumassociates.com

Administration: Can be administered by direct observation (eliciting skills) or parent report. Separate forms for infants and toddlers.

Scores Available: Total scores are compared with cut-points indicating need for further evaluation. Scores can be converted to quotients, age equivalents, and percentiles.

Subscales: General Knowledge and Comprehension; Speech and Language, Gross-Motor Skills; Fine Motor Skills, Preacademic.

Norming Sample: 203 infants and 179 toddlers from 29 sites. Further information about the sample's demographic characteristics is not available.

Reliability: High (.80 or higher)

Sensitivity/Specificity: Adequate (between 65% and 79%)

Notes: Some of the norming, reliability, and sensitivity/specificity information may come from earlier versions of the screen.

Brigance K & 1 Screen II (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
K – 1st grade	10 – 20 minutes	For use by teachers, paraprofessionals, therapists, nurses, physicians. No special training required.	$148

Available Languages: English, Spanish

Source: Curriculum Associates, Inc., 153 Rangeway Road, North Billerica, MA 01862
(800) 225-0248; www.curriculumassociates.com

Administration: Direct assessment of individual child. Separate forms for K and 1st grade. Social-emotional, self-help, and reading readiness are assessed through supplemental parent or teacher rating scales.

Scores Available: Totals are compared to cut-points. Raw scores are converted to standard scores, percentiles, quotients, age equivalents, percentages of delay, and deviation scores.

Subscales: General Knowledge and Comprehension; Speech and Language, Gross-Motor Skills; Fine Motor Skills, Preacademic, Social-Emotion, Self-Help, Reading Skills and Manuscript Writing.

Norming Sample: 1,366 children, reflective of U.S. population in terms of region, gender, and parent education (although no children from major metropolitan areas were included).

Reliability: High (.80 or higher)

Sensitivity/Specificity: Sensitivity High (80% or higher); specificity may be available in the technical manual.

Notes: Designed to screen children for developmental delays or for giftedness. Some of the norming, reliability and sensitivity/specificity information may come from earlier versions.

Brigance Preschool Screen II (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 4 years	10 – 15 minutes	For use by teachers, paraprofessionals, therapists, nurses, physicians. No special training required.	$148

Available Languages: English, Spanish

Source: Curriculum Associates, Inc., 153 Rangeway Road, North Billerica, MA 01862
(800) 225-0248; www.curriculumassociates.com

Administration: Direct assessment of individual child. Also includes observation, teacher rating, and parent rating forms.

Scores Available: Cut-points are provided for identifying children needing additional testing. Scores can be converted to quotients, age equivalents, and percentiles.

Subscales: General Knowledge and Comprehension, Speech and Language, Gross-Motor Skills; Fine Motor Skills, Preacademic, Social-Emotion and Self-Help.

Norming Sample: Roughly 90 children per age group; stratified for gender, race/ethnicity, and parent education (2005 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Sensitivity/Specificity: Not reported

Notes: Designed to screen young children for developmental delays or for giftedness. Some of the norming, reliability, and validity information may come from earlier versions.

Developmental Indicators for the Assessment of Learning, Third Edition [DIAL-3] (1998)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 6:11 years	20 – 30 minutes	Professional trained in special education, early childhood education, psychology, or other related area.	$501

Available Languages: English, Spanish

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child. Can be set up as stations to test large numbers of children at once. Parent report used for Self-Help and Social subscale.

Scores Available: Standard deviation, percentile ranks, standard scores, and percentile cut-points indicating "potential delay" or "OK" by chronological age at two-month intervals.

Subscales: Motor, Concepts, Language, Self-Help, and Social.

Norming Sample: 1,560 English-speaking children representative of U.S. population in terms of gender, race-ethnicity, region, and parent education (1994 U.S. Census); 605 Spanish-speaking children from the U.S. Mainland, Puerto Rico, and Panama.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Sensitivity/Specificity: Not reported

Note: The Speed Dial is a shortened version of the DIAL.

Developmental Observation Checklist System [DOCS] (1994)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 6 years	30 minutes	Parent needs 4th grade reading level. Some background in testing required for proper scoring.	$165

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Three part three-part inventory/checklist system completed by the parent. Includes 540 items, but only a portion are completed depending on child's age.

Scores Available: Quotients, NCE scores, age equivalents, and percentiles.

Subscales: 3 Components: General development (DC), adjustment behavior (ABC), and parent stress and support (PSSC). General development (DC) is further broken down into 4 subscales: Language, Motor, Social, and Cognitive Development.

Norming Sample: More than 1,400 children, representative of the U.S. population with regard to gender, geographic region, race/ethnicity, and urban/rural residence (1990 Census).

Reliability: High (.80 or higher)

Sensitivity/Specificity: May be available in the technical manual.

Concurrent Validity: High (.70 or higher) (n = 20-35)

Eyberg Child Behavior Inventory [ECBI] and Sutter-Eyberg Student Behavior Inventory-Revised [SESBI-R] (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2 – 16 years	10 minutes	Sixth grade reading level needed to complete forms. Interpretation requires graduate training.	$170

Available Languages: English

Source: Psychological Assessment Resources; 16204 North Florida Avenue, Lutz FL 33549
(800) 331-8378; www.parinc.com

Administration: Parent (ECBI; 36 items) or Teacher (SESBI-R; 38 items) reports on child's behavior by indicating how often listed behaviors occur and whether or not the behavior is problematic for the responder.

Scores Available: Intensity and problem raw scores can be converted to T-scores. Cut-points for clinical significance are provided.

Subscales: Intensity and Problem

Norming Sample: *ECBI*: 798 children (aged 2 to 16 years), from six outpatient pediatric settings in the southeastern United States; *SESBI-R*: 415 elementary school children from 11 schools in Gainesville, Florida. Neither sample was nationally representative.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Sensitivity/Specificity: May be available in technical manual.

Note: Designed to assess the current frequency and severity of disruptive behaviors in the home and school settings.

FirstSTEp: Screening Test for Evaluating Preschoolers (1993)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:9 – 6:2 years	15 – 20 minutes	Psychology, counseling, or related degree, plus coursework in testing.	$250

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Cognitive, Communication and Motor domains assessed directly. Social-Emotional subscale is a checklist completed by the examiner based on observations. Adaptive Behavior subscale is a checklist is completed by the parent or caregiver.

Scores Available: Cut-points for "Within Acceptable Limits," "Caution," or "At-Risk".

Subscales: Cognitive (Quantitative Reasoning, Picture Completion, Visual Position in Space, Problem Solving); Communication (Auditory Discrimination, Word Retrieval, Association, Sentence and Digit Repetition); Motor (Visual-Motor Integration, Fine Motor Planning, Balance, Gross Motor Planning); plus Social-Emotional and Adaptive.

Norming Sample: 1,433 children similar to U.S. population with regard to region, race/ethnicity, and parent education (1988 U.S. Census).

Reliability: High (.80 or higher)

Sensitivity/Specificity: Information may be available in technical manual. False positive rate (i.e., percent of children who were referred but did not have a disability) was Adequate (15% to 21%) and false negative rate (i.e., percent of children who were not referred by did have a disability) was Low (1% to 3%).

Note: Used to identify children who may have mild to severe developmental delays.

Fluharty Preschool Speech and Language Screening Test, Second Edition [Fluharty-2] (2001)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 6:11 years	10 minutes	Graduate degree with coursework in testing and training in speech and language.	$168

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child.

Scores Available: Standard scores, percentiles, age equivalents, and quotient scores (RLQ, ELQ, and GLQ).

Subscales: 5 subtests: Articulation, Repeating Sentences, Following Directives and Answering Questions, Describing Actions, Sequencing Events; 3 quotients: Receptive Language, Expressive Language, General Language.

Norming sample: 705 children from 21 states, generally representative of the U.S. population with regard to age, gender, race, residence (urban versus rural), ethnicity, and disability status (1998 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher) (n=23)

Sensitivity/Specificity: May be available in the Technical Manual.

Note: Designed to identify children who need a more comprehensive diagnostic evaluation of their speech and language skills.

Infant Toddler Symptom Checklist [ITSC] (1995)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:7 – 2:6 years	10 minutes	Completed by parent or paraprofessional with no training.	$80

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Parent or caregiver completes a checklist about child's symptoms, using a three point scale: *never or sometimes*, *most times*, or *past*. There are six versions: a single short version for general screening purposes and five age-specific screens for both diagnostic and screening purposes: 7 to 9 months, 10 to 12 months, 13 to 18 months, 19 to 24 months, and 25 to 30 months. The versions range from 18 to 31 items.

Scores Available: Total score is compared to cut-points for at risk of having a regulatory disorder warranting further diagnosis.

Subscales: 9 domains: (1) self-regulation, (2) attention, (3) sleep, (4) eating or feeding, (5) dressing, bathing, and touch, (6) movement, (7) listening and language, (8) looking and sight, and (9) attachment/emotional functioning.

Norming Sample: Not normed

Reliability: Not reported

Sensitivity/Specificity: Information may be available in technical manual. False positive rate (i.e., percent of children who were referred but did not have a disability) and false negative rates (i.e., percent of children who were not referred by did have a disability) were both low (0% to 14%).

Note: Designed to screen for regulatory and sensory disorders.

Learning Accomplishment Profile – Normed Screens [LAP-D Normed Screens] (1997)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 5 years	15 minutes	Designed to be administered by professionals including psychologists, occupational and physical therapists, physicians, nurses and social workers. Can also be administered by kindergarten and child care teachers with training.	$400 Kit (Ages 3, 4, & 5) $140 Individual (Ages 3, 4 or 5)

Available Languages: English, Spanish

Source: Kaplan Early Learning Company, 1310 Lewisville-Clemmons Road, Lewisville, NC 27023
(800) 334-2014; www.kaplanco.com

Administration: Direct assessment of individual child.

Scores Available: Z-scores, T-scores, Developmental Scores, Percentile Ranks; cut-points are provided for identifying children who should be referred for further assessment.

Subscales: None

Norming Sample: 907 children in 15 states; African American children were somewhat over-represented and White and Latino children were somewhat underrepresented, as compared to the 1990 Census. No other comparisons with the Census provided.

Reliability: High (.80 or higher)

Sensitivity/Specificity: Sensitivity: High (80% or higher); Specificity: not reported

Note: Assesses fine motor, gross motor, cognitive and language domains.

Parents' Evaluation of Development Status [PEDS] (1997)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 8 years	5 minutes	Can be administered by a range of professionals and paraprofessional, including office staff, after reading brief scoring and administration guide.	$30

Available Languages: English, Spanish, Vietnamese, Huang, Somali and Chinese

Source: Forepath, Ltd., P.O. Box 23186, Washington D.C., 20026
(615) 776-4121; www.forepath.org

Administration: Interview with parent or parent completes brief questionnaire.

Norming Sample: 2,823 children representative of the U.S. population (year unknown) in terms of ethnicity, parental education, income, urbanicity, developmental disability.

Scores Available: Categorizes responses into low, medium or high risk for developmental and behavioral/mental health problems. A longitudinal score and form remain in child's medical record.

Reliability: High (.80 or higher)

Sensitivity/Specificity: Adequate (between 65% and 79%)

Note: Designed to screen for developmental and behavioral problems needing further evaluation. Collects information about: Global/Cognitive; Expressive Language and Articulation; Receptive Language; Fine-Motor; Gross-Motor; Behavior; Social-emotional; Self-Help; School; and Other. Parents can also complete this screener on-line for $9.95.

Pervasive Developmental Disorders Screening Test–II [PDDST–II] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
1:0 – 4:0 years	10 – 20 minutes	Parents complete screening form(s), clinician needed to interpret results.	$145

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Includes three forms completed by parents or a person familiar with the child and interpreted by a clinician. *Stage 1*, the Primary Care Screener (PCS, 22 items), can be used in primary care settings for children. *Stage 2*, the Developmental Clinic Screener (DCS, 14 items), can be used in clinics where children are being screened for possible developmental delays. Finally, *Stage 3*, the Autism Clinic Severity Screener (ACSS, 12 items), can be used in clinics that are conducting a complete diagnostic assessment on children with Autism Spectrum Disorders.

Scores Available: Raw scores are compared to cut-points for positive versus negative screen.

Subscales: None (designed to screen for autism spectrum disorders)

Norming Sample: Not normed

Reliability: Not reported

Sensitivity/Specificity: *Stage 1*: High (80% or higher); *Stage 2:* Adequate Sensitivity (65% to 79%), Low (below 65%) Specificity; *Stage 3:* Low (below 65%)

Note: Designed to screen for several autistic spectrum disorders, including autistic disorder, pervasive developmental delay, and Asperger's disorder.

Speed DIAL [Developmental Indicators for the Assessment of Learning] (1998)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 6:11 years	15 – 20 minutes	Professional trained in special education, early childhood education, psychology, or other related area.	$192

Available Languages: English, Spanish

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child. Can be set up as stations to test large numbers of children at one time.

Scores Available: Scores are compared with cut-points to determine potential delay.

Subscales: None

Norming Sample: 1,560 English-speaking children representative of the U.S. population with regard to gender, race-ethnicity, region, and parent education (1994 U.S. Census); 605 Spanish-speaking children from the U.S. Mainland, Puerto Rico, and Panama.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Sensitivity/Specificity: Not reported

Note: Speed DIAL is a shortened version of the DIAL-3, measuring only three domains (Motor, Concepts and Language).

Collection of Assessment Tools for Informing Instruction & Monitoring Progress

Page	Name
122	Adaptive Behavior Assessment System, 2nd Edition [ABAS-II] *
105	Assessment, Evaluation, and Programming System (AEPS) for Birth to Three Years, 2nd Edition
105	Assessment, Evaluation, and Programming System (AEPS) for Three to Six Years, 2nd Edition
124	Behavior Assessment System for Children, 2nd Ed. (BASC-2) *
125	Boehm Test of Basic Concepts, Preschool, 3rd Edition *
126	Boehm Test of Basic Concepts, 3rd Edition*
127	Bracken Basic Concept Scale – Revised (BBCS-R) *
106	Brigance Comprehensive Inventory of Basic Skills-Revised (CIBS-R)
106	Brigance Diagnostic Inventory of Early Development II (IED-II)
127	Bruininks-Oseretsky Test of Motor Proficiency, 2nd Edition (BOT-2) *
107	Carey Temperament Scales (CTS)
107	Carolina Curriculum for Infants and Toddlers with Special Needs (CCITSN), 3rd Edition
108	Carolina Curriculum for Preschoolers with Special Needs (CCPSN), 2nd Edition
128	Clinical Evaluation of Language Fundamentals, 4th Edition (CELF-4) *
128	Clinical Evaluation of Language Fundamentals – Preschool, 2nd Edition (CELF-P 2) *
108	Creative Curriculum Developmental Continuum for Ages Three to Five
109	Creative Curriculum Developmental Continuum for Infants, Toddlers & Twos
109	Developmental Reading Assessment, 2nd Edition (DRA 2)
110	Devereux Early Childhood Assessment (DECA)
132	Differential Ability Scales – II (DAS-II) *
121	Dynamic Indicators of Basic Early Literacy Skills, 6th Edition (DIBELS-6)
110	Early Reading Diagnostic Assessment, 2nd Edition (ERDA 2) *
132	ECLS-K Approaches to Learning Sub-Scale
111	Get It, Got It, Go (Preschool IGDIs)
112	Hawaii Early Learning Profile (HELP) (0-3 years)
113	Hawaii Early Learning Profile for Preschoolers (HELP for Preschoolers) (3-6 years)
113	High/Scope Child Observation Record for Infants and Toddlers (COR-IT)
113	High/Scope Preschool Child Observation Record (COR)
114	Individual Growth Development Indicators (IGDIs) for Infants and Toddlers
138	Kaufman Test of Educational Achievement, 2nd Edition (KTEA-II) *
114	Ounce Scale
144	Peabody Developmental Motor Scales, 2nd Edition (PDMS-2) *

Page	Name
115	Phonological Awareness and Literacy Screenings- Kindergarten (PALS-K)
115	Phonological Awareness and Literacy Screenings-PreKindergarten (PALS-PreK)
116	Phonological Awareness and Literacy Screenings-1-3 (PALS-1-3)
146	Phonological Awareness Test 2 (PAT-2) *
148	Process Assessment of the Learner, 2nd Edition: Diagnostic Assessment for Math (PAL-II Math) *
149	Process Assessment of the Learner, 2nd Edition: Diagnostic Assessment for Reading and Writing (PAL-II Reading and Writing) *
116	Qualls Early Learning Inventory (QELI)
117	Teacher Rating of Oral Language & Literacy (TROLL)
157	Test of Word Reading Efficiency (TOWRE) *
117	Transdisciplinary Play Based Assessment, 2nd Edition (TPBA 2)
118	Work Sampling System (WSS)
118	Young Children's Achievement Test (YCAT)

* Tools marked with an asterisk have two purposes: i*nform instruction/monitor progress* and *diagnostic*.

Their descriptions appear in the Diagnostic section, because that is their primary purpose.

Assessment, Evaluation, and Programming System [AEPS] for Birth to Three Years, Second Edition (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 3 years	1 – 2 hours	Service providers, interventionists, home visitors, and specialists. Little training required.	$179

Available Languages: English

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624 (800) 638-3775; www.brookespublishing.com

Administration: Provider or specialist reports about individual child's skills, based on observation and family input.

Scores Available: Raw scores and frequencies are calculated to track each child's progress over time. Raw scores can be compared to cut-points.

Subscales: Fine motor, gross motor, cognitive, adaptive, social-communication, and social.

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Notes: Curriculum-based assessment for use with children who have disabilities or are at risk of developmental delays.

Assessment, Evaluation, and Programming System [AEPS] for Three to Six Years, Second Edition (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 6 years	1 – 2 hours	Service providers, interventionists, home visitors, and specialists Little training required.	$179

Available Languages: English

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624 (800) 638-3775; www.brookespublishing.com

Administration: Provider or specialist reports on individual child's skills, based on observation and family input.

Scores Available: Raw scores and frequencies are calculated to track each child's progress over time. Raw scores can be compared to cut-points.

Subscales: Fine motor, gross motor, cognitive, adaptive, social-communication, and social

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: Not reported

Notes: Curriculum-based assessment for use with children who have disabilities or are at risk of developmental delays.

Brigance Comprehensive Inventory of Basic Skills –Revised (CIBS-R) (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
PreK– 6th grade	45 – 75 minutes	No specialized training required.	$185

Available Languages: English, Spanish (called ABS-R)

Source: Curriculum Associates, Inc., 153 Rangeway Road, North Billerica, MA 08162 (800) 225-0248; www.curriculumassociates.com

Administration: Direct assessment of individual child. The total program consists of 154 assessments in eight areas, covering Readiness, Speech, Listening, Reading, Spelling, Writing, Research and Study Skills, and Mathematics.

Scores Available: *Criterion-referenced:* listing of skills mastered; *Norm-referenced*: percentiles, age and grade equivalent scores, instructional ranges, and standard scores.

Subscales: Readiness, Speech, Listening, Study Skills, Reading, Spelling, Writing, Math.

Norming Sample: 1,121 children from six U.S. cities in four geographical regions; demographic characteristics roughly comparable to U.S. population in terms of gender, race-ethnicity, region, community type, and SES (1997 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Brigance Diagnostic Inventory of Early Development II [IED-II] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 7 years	25 – 30 minutes	Criterion referenced: teachers and other educators. Norm-referenced: training in psychological assessment.	$185

Available Languages: English, Spanish

Source: Curriculum Associates, Inc., 153 Rangeway Road, North Billerica, MA 08162 (800) 225-0248; www.curriculumassociates.com

Administration: Both criterion and norm referenced components are administered via parent and teacher interviews, observations, and direct assessment of individual child. Components are overlapping, but not identical, and require separate scoring.

Subscales: *Criterion-referenced*: Preambulatory Motor, Gross Motor, Fine Motor, Self-Help, Speech and Language, General Knowledge and Comprehension, Social and Emotional Development, Readiness, Basic Reading, Manuscript Writing, and Basic Math. *Norm-referenced*: Motor, Language, Academic-Cognitive, Daily Living, and Social-Emotional. These five skills are used to create an Adaptive Behavior composite.

Scores Available: *Criterion-referenced* Includes a system of identifying areas of developmental delay and obtaining instructional objectives. There is also a comprehensive skill sequence for assessing children with suspected developmental delays and obtaining more detailed information about developmental skills. *Norm-referenced*: derived quotient scores, confidence intervals, percentiles, age equivalents, and instructional ranges.

Norming Sample: 1,171 in 24 states; demographic characteristics fairly similar to U.S. population with regard to race-ethnicity, parental education, and free-reduced lunch program participation, with a somewhat more highly educated and urban-suburban-residing population than the U.S. population (2003 U.S. Census).

Reliability: *Criterion referenced*: Not reported; *Norm referenced*: Adequate (.65 to .79)

Concurrent Validity: High (.70 or higher)

Carey Temperament Scales [CTS] (2000)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:1 – 12 years	20 minutes	Early high school reading level required for administration. Professional needed for scoring and interpretation.	$70

Available Languages: English

Source: Behavioral-Developmental Initiatives, 14636 North 55th Street, Scottsdale, AZ 85254 (800) 405-2313; www.b-di.com

Administration: Parent report about individual child; 75-100 items; 5 different questionnaires for children of different ages.

Scores Available: Category score for each of the nine areas, which are then compared to the norms for the category. Professional Report includes the temperament profile, raw and standardized scores, individualized interpretive report and validity. Caregiver Report contains the temperament profile and an interpretive report of scores.

Subscales: Activity level, rhythmicity, approach-withdrawal, adaptability, intensity, mood, attention span and persistence, distractibility, and sensory threshold.

Norming sample: 200-500 children. Sample is not nationally representative. Normed on the East Coast and primarily included middle class, Caucasian children.

Reliability: Adequate (.65 to .79)

Concurrent Validity: Not reported

Notes: The CTS can help caregivers understand a child's temperament and behavioral style. Temperament is not considered amenable to intervention.

Carolina Curriculum for Infants and Toddlers with Special Needs [CCITSN], Third Edition (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:0 – 3:0 years	60 – 90 minutes	Paraprofessional or professional	$45

Available Languages: English

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624 (800) 638-3775; www.brookespublishing.com

Administration: Combined assessment and curriculum approach. The volume first takes users through the assessment process. Once the initial assessment is complete, professionals select curricular items that correspond to each child's special needs. The curriculum itself is divided into 24 teaching sequences covering five developmental domains: cognition, communication, social adaptation, fine motor, and gross motor.

Scores Available: Assessment log and developmental progress chart. Developmental Progress Reports can be used to create a profile of skills.

Subscales: Cognition, communication, personal-social, fine motor skills, and gross motor skills, with more than 20 developmental subdomains

Norming Sample: Not normed

Reliability: Not reported

Concurrent Validity: Not reported

Note: Website indicates that Spanish versions of the forms will be available shortly.

Carolina Curriculum for Preschoolers with Special Needs [CCPSN], Second Edition (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2 – 5 years	60 – 120 minutes	Paraprofessional or professional	$45

Available Languages: English

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624
(800) 638-3775; www.brookespublishing.com

Administration: Combined assessment and curriculum approach. The volume first takes users through the assessment process. Once the initial assessment is complete, professionals select curricular items that correspond to each child's special needs. The curriculum itself is divided into 22 teaching sequences covering five developmental domains: cognition, communication, social adaptation, fine motor, and gross motor.

Scores Available: Assessment log and developmental progress chart. Developmental Progress Reports can be used to create a profile of skills.

Subscales: Cognition, communication, personal-social, fine motor skills, and gross motor skills, with more than 20 developmental subdomains.

Norming Sample: Not normed

Reliability: Not reported

Concurrent Validity: Not reported

Note: Website indicates that Spanish versions of the forms will be available shortly.

Creative Curriculum Developmental Continuum for Ages Three to Five (2000)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 5 years	Ongoing	Teacher administers, scores and interprets.	$76

Available Languages: English, Spanish

Source: Teaching Strategies, Inc., PO Box 42243 Washington, DC 20015
(800) 637-3652; www.teachingstrategies.com

Administration: Teachers collect data throughout the school year through multiple methods of assessment such as checklists and anecdotal notes of growth. The teacher observes the child's learning in relation to the goals set by the Creative Curriculum framework. This recorded information is then used to rate child's development on the Developmental Continuum (ratings used are Forerunner, Level I, Level II, or Level III). Information about an individual child can be rated up to three times a year (fall, winter, and spring), allowing the user to assess change over time.

Scores Available: Information may be available in Technical Manual.

Subscales: 4 main constructs: Social/Emotional Development, Physical Development, Cognitive Development, and Language Development.

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: Not reported

Note: Curriculum-based assessment based on The Creative Curriculum for Ages Three to Five.

Creative Curriculum Developmental Continuum for Infants, Toddlers & Twos (2006)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 2:11 years	Ongoing	Teacher administers, scores and interprets.	$106

Available Languages: English, Spanish

Source: Teaching Strategies, Inc., PO Box 42243 Washington, DC 20015
(800) 637-3652; www.teachingstrategies.com

Administration: Teachers use 21 objectives to monitor child progress, based on on-going observations.

Scores Available: Information may be available in Technical Manual.

Subscales: 4 main constructs: Social/Emotional Development, Physical Development, Cognitive Development, and Language Development.

Norming Sample: Not normed

Reliability: May be available in technical manual.

Concurrent Validity: Not reported

Note: Curriculum-based assessment based on The Creative Curriculum for Infants, Toddlers & Twos.

Developmental Reading Assessment, Second Edition [DRA 2] (2006)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
K – 3rd grade	10 – 20 minutes	Administered by the child's teacher, who complete training DVD.	$311

Available Languages: English, Spanish (Evaluación del Desarrollo de la Lectura 2 [EDL2])

Source: Pearson Learning Group, 145 South Mount Zion Road, P.O. Box 2500, Lebanon, IN 46052
(800) 321-3106; www.pearsonlearning.com

Administration: Direct assessment of individual child by teacher.

Scores Available: Independent reading level; stage of reading development; accuracy rate (percent of correctly read words); comprehension and fluency as scored on a rubric; benchmarks.

Subscales: Fluency, comprehension, accuracy; also includes separate word analysis assessment (Phonological Awareness, Metalanguage, Letter/Word Recognition, Phonics, and Structural Analysis)

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Devereux Early Childhood Assessment [DECA] (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2 – 5 years	10 minutes	Parents/teachers need a 6th grade reading level. Professional needed for scoring and interpretation.	$200

Available Languages: English, Spanish

Source: Kaplan Early Learning Company, 1310 Lewisville-Clemmons Road, Lewisville, NC 27023 (800) 334-2014; www.kaplanco.com

Administration: Raters, such as parents or teachers, complete a 37-item rating scale about the individual child.

Scores Available: Raw scores, percentile scores, T-scores, normal curve equivalent scores, and individual profiles. Cut-points provided for "below average," "average," and "above average"

Subscales: Two main scales: Total Protective Factors and Behavioral Concern. Total Protective Factor scale has three subscales (Imitative, Self-control, Attachment).

Norming Sample: One sample of 2,000 children was used to norm the protective scale and another sample of 1,108 children was used to norm the behavioral problem scale. Children in the samples closely represent the U.S. population with regard age, gender, geographical region, race/ethnicity, and socioeconomic status (1995 Census).

Reliability: High (.80 or higher)

Concurrent Validity: Not reported. (See Technical Manual for information on Criterion Validity)

Note: Tool designed to designed to evaluate self-protecting factors and behavioral concerns among preschool children ages.

Dynamic Indicators of Basic Early Literacy Skills, Sixth Edition [DIBELS-6] (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
K – 6th grade	1 minute each; 10 minutes total	Education personnel can administer and interpret with little training.	Free at website $1 per year per student for data system

Available Languages: English, Spanish (Indicadores Dinámicos del Éxito en la Lectura)

Source: Institute for the Development of Educational Achievement, 1211 University of Oregon, Eugene, Oregon 97403-1211, dibels.uoregon.edu

Administration: Direct assessment of individual child administered 3 to 4 times per year.

Scores Available: Percentiles, benchmark goals, cut-points (low risk, some risk, at risk)

Subscales: Initial Sounds Fluency, Phonemic Segmentation Fluency, Nonsense Word Fluency, Oral Reading Fluency, Retell Fluency, Letter Naming Fluency, Word Use Fluency

Norming Sample: Percentile scores were developed using data from all schools that use the DIBELS system, including about 40,000 kindergartners, 40,000 first graders, 15,000 second graders, and 10,000 third graders. Very large, but not nationally representative.

Reliability: High (.80 or higher). Not reported for Word Use Fluency, Oral Reading Fluency, Retell Fluency.

Concurrent Validity: Adequate (.50 to .69). Not reported for Word Use Fluency or Retell Fluency.

ECLS-K Approaches to Learning Sub-Scale (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
K – 1st grade	Less than 5 minutes	No training required.	None

Available Languages: English, Spanish, Chinese, Lakota, and Hmong

Source: National Center for Education Statistics. See http://nces.ed.gov/pubs96/9618.pdf (p. 34) and http://nces.ed.gov/pubs2001/2001029rev_1_4.pdf (section 2.3.2). Individual items are included in note (below).

Administration: Parent or teacher respond to six items about the child.

Scores Available: Raw scores only

Subscales: None

Norming Sample: Not normed

Reliability: High (.80 or higher) for teacher report, Adequate (.65 to .79) for parent report.

Concurrent Validity: Adequate (.50 to .69), with teacher report of academic performance.

Note: Tool has only been used for research purposes. The full scale appears below.

Please rate this child on the following characteristics:

	Never	Sometimes	Often	Very Often
1. attentiveness	1	2	3	4
2. task persistence	1	2	3	4
3. eagerness to learn	1	2	3	4
4. learning independence	1	2	3	4
5. flexibility	1	2	3	4
6. organization	1	2	3	4

Get It, Got It, Go (Preschool IGDIs) (2000)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 5 years	Less than 10 minutes	Can be administered by psychologists, teachers, paraprofessionals, volunteers, or others.	Free at website

Available Languages: English, Spanish

Source: Center for Early Education and Development (CEED),University of Minnesota, Education Sciences Building, Suite 40, 56 East River Road, Minneapolis, MN 55455, 612-625-3058; 612-625-2093; ggg.umn.edu

Administration: Direct assessment of individual child, repeated regularly throughout the year (4-8 week intervals).

Scores Available: Child data are entered into an on-line system that generates a child report., which includes: (a) a table and a graph of IGDI scores, (b) a trend line, and (c) an aim line. The trend and aim lines offer a visual comparison between the child's actual rate of progress and a desired rate of progress. The aim line is based on a study group of English-speaking preschool children without identified disabilities. Programs can create their own aims if desired.

Subscales: Picture Naming, Alliteration, Rhyming

Norming Sample: Not normed

Reliability: Adequate (.65 to .79)

Concurrent Validity: Adequate (.50 to .69)

Note: A "Movement IGDI" for infants/toddlers and preschoolers is available at http://cehd.umn.edu/ceed/projects/movement/default.html

Hawaii Early Learning Profile [HELP] (0–3 years) (2006)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 3 years	Ongoing observation	Physical, speech, and occupational therapists, early childhood educators, infant specialists, psychologists, social workers, and nurses.	$60

Available Languages: English, Spanish

Source: Vort Corporation, P.O. Box 60132-W, Palo Alto, CA 94306
(650) 322-8282; www.vort.com

Administration: Checklist (685 items in 58 strands) is completed by examiner based on observation and parent report.

Scores Available: Skills within each strand are developmentally sequenced. Each strand includes HELP skills which focus upon a specific underlying key concept and are hierarchical in nature (i.e., one skill leads to or builds the foundation for the next skill). Activity Guide includes thousands of task-analyzed, practical curriculum activities and intervention strategies indexed by the 685 HELP skills

Subscales: Cognitive, Language, Gross Motor, Fine Motor, Social, Self-Help

Norming Sample: Not normed

Reliability: Not reported

Concurrent Validity: Not reported

Note: Curriculum-based assessment.

Hawaii Early Learning Profile for Preschoolers [HELP for Preschoolers] (3–6 years) (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 6 years	Ongoing observation	Physical, speech, and occupational therapists, early childhood educators, infant specialists, psychologists, social workers, and nurses.	$65

Available Languages: English, Spanish

Source: Vort Corporation, P.O. Box 60132-W, Palo Alto, CA 94306
(650) 322-8282; www.vort.com

Administration: Checklist (622 items in 46 strands) is completed by examiner based on observation and parent report.

Scores Available: Skills within each strand are developmentally sequenced. Each strand includes HELP skills which focus upon a specific underlying key concept and are hierarchical in nature (i.e., one skill leads to or builds the foundation for the next skill). Activity Guide includes thousands of task-analyzed, practical curriculum activities and intervention strategies indexed by the 685 HELP skills

Subscales: Cognitive, Language, Gross Motor, Fine Motor, Social, Self-Help

Norming Sample: Not normed

Reliability: Not reported

Concurrent Validity: Not reported

Note: Curriculum-based assessment.

High/Scope Child Observation Record for Infants and Toddlers [COR-IT] (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
6 weeks – 3 years	Observation over several weeks/months	Conducted by teacher or other caregiver. Two day training by High/Scope is recommended.	$175

Available Languages: English, Spanish

Source: High/Scope Press, 600 North River Street, Ypsilanti, MI 48198-2898
(800) 407-7377; www.highscope.org

Administration: Teachers, parents, or other caregivers record observations (anecdotes) as they care for and play with the child. For each of the 28 items, the caregiver compares examples with the anecdotes to rank the child's typical behavior on a five-point scale on the development summary form and enters the highest level of behavior the child achieved.

Scores Available: Raw scores for each time period; growth profile; category reports (summarizes skills with corresponding activities to support development); group summary; family report.

Subscales: Sense of Self, Social relations, Creative Representation, Movement, Communication and Language, Exploration and Early Logic

Norming sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher) ($n = 30$)

Note: Curriculum-based assessment.

High/Scope Preschool Child Observation Record [COR] (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:6 – 6:0 years	Observation over several weeks/months	Conducted by teacher or other caregiver. Two day training by High/Scope is recommended.	$175

Available Languages: English, Spanish

Source: High/Scope Press, 600 North River Street, Ypsilanti, MI 48198-2898
(800) 407-7377; www.highscope.org

Administration: Teachers, parents, or other caregivers record observations (anecdotes) as they care for and play with the child. For each of the 32 items, the caregiver compares examples with the anecdotes to rank the child's typical behavior on a five-point scale on the development summary form and enters the highest level of behavior the child achieved.

Scores Available: Raw scores for each time period; growth profile; category reports (summarizes skills with corresponding activities to support development); group summary; family report.

Subscales: Initiative, Social Relations, Creative Representations, Music and Movement, Language and Literacy, Logic and Mathematics.

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: Low (below .50) (Note: The correlation between COR Total and the CSAB were non-significant due to very small sample size [n=10]. Larger, significant correlations were attained for some sub-scales (such as language/literacy and social relations), using somewhat larger sample sizes ($n = 22$-26).

Note: Curriculum-based assessment

Individual Growth and Development Indicators [IGDIs] for Infants and Toddlers (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 3 years	40 minutes	Early childhood practitioners and interventionists who have completed the training.	Forms free at website; $1 per child for processing

Available Languages: Can be conducted in any language, as long as the assessor speaks the child's primary language

Source: Juniper Gardens Children's Project, University of Kansas, 650 Minnesota, Suite 2, Kansas City, KS 66101-2800
(913) 321-3143; www.igdi.ku.edu

Administration: Structured interaction with individual child and parent; repeated often to track progress.

Scores Available: Child data are entered into an on-line system that generates both child and program level reports. Child report includes Graphs that provide visual comparison between individual child and normative growth trajectories for individual behaviors.

Subscales: Communication, movement, social, problem solving, parent-child interaction

Norming Sample: Not normed

Reliability: High (.80 or higher) for communication, movement, social, problem solving; reliability for parent-child interactions not reported.

Concurrent Validity: Adequate (.50 to .69). Not reported for parent-child interactions.

Ounce Scale (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 3:6 years	Ongoing observation	Can be administered by early intervention specialists or child care staff.	$123

Available Languages: English, Spanish

Source: Pearson Learning Group, 145 South Mount Zion Road, P.O. Box 2500, Lebanon, IN 46052
(800) 321-3106; www.pearsonlearning.com

Administration: Consists of three elements: (1) *Observation Record* provides a focus for observing and documenting child's everyday behaviors, (2) *Family Album* provides a structure for parents to learn about and record their child's development, and (3) *Developmental Profile* uses Observation Record and Family Album to evaluate each child's development over time, comparing observations to specific performance standards.

Scores Available: Behaviors are marked as "Developing as Expected" or "Needs Development" and space is available to record comments.

Subscales: Personal Connections, Feelings about Self, Relationships with Other Children, Understanding and Communication, Exploration and Problem Solving, Movement and Coordination

Norming sample: Not normed

Reliability: Not reported

Concurrent Validity: Not reported

Phonological Awareness and Literacy Screenings– Kindergarten [PALS-K] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5 years (kindergarten)	20 – 25 minutes	Teachers who have read the manual and scoring guide.	$95

Available Languages: English

Source: PALS – University of Virginia, PO Box 800785, Charlottesville, VA 22908
(866) 372-7257; www.pals.virginia.edu

Administration: Direct assessment of individual child.

Scores Available: Summed scores compared to cut-points for determining which students need instruction in addition to the regular classroom literacy instruction.

Subscales: Rhyme Awareness, Beginning Sound Awareness, Alphabet Knowledge, Letter Sounds, Spelling, Concept of Word, Word Recognition in Isolation

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Phonological Awareness and Literacy Screenings–PreK [PALS-PreK] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
4 years (pre-kindergarten)	20 – 25 minutes	Teachers who have read the manual and scoring guide.	$75

Available Languages: English

Source: PALS – University of Virginia, PO Box 800785, Charlottesville, VA 22908
(866) 372-7257; www.pals.virginia.edu

Administration: Direct assessment of individual child.

Scores Available: Developmental ranges, expectations (including cut-points)

Subscales: Name Writing, Alphabet Knowledge, Beginning Sound Awareness, Print and Word Awareness, Rhyme Awareness, Nursery Rhyme Awareness

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Phonological Awareness and Literacy Screenings–1–3 [PALS 1–3] (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
1st – 3rd grade	20 – 25 minutes	Teachers who have read the manual and scoring guide	$95

Available Languages: English

Source: PALS – University of Virginia, PO Box 800785, Charlottesville, VA 22908
(866) 372-7257; www.pals.virginia.edu

Administration: Direct leveled assessment of individual child.

Scores Available: Summed scores compared to cut-points for determining which students need instruction in addition to the regular classroom literacy instruction.

Subscales: *Entry Level*: Spelling; Word Recognition in Isolation

Level A Oral Reading in Context: Passage Reading, Timing Passage Readings, Comprehension, Fluency

Level B: Alphabetics (for students whose Entry Level scores do not meet grade-level criteria):

Alphabet Recognition, Letter Sounds, Concept of Word

Level C (for students whose do not met Level B benchmarks): Phonemic Awareness: Blending, Sound-to-Letter

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Qualls Early Learning Inventory [QELI] (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
PreK – 1st grade	5 – 10 minutes	Teacher completed questionnaire. Minimal training required.	$32 for PreK or K-1st

Available Languages: English

Source: Riverside Publishing Company, 8420 Bryn Mawr Avenues, Chicago, IL 60631
(800) 767-8420; www.riverpub.com

Administration: Questionnaire completed by teacher based on observations of student made over time in a naturalistic setting. Separate editions for PreK and K-1st.

Scores Available: Diagnostic Report indicates the extent to which each child has mastered each skill and has and indicates *delayed*, *developing* or *developed* for each of the 6 subscales. Student Profile Narrative for Communicating with Parents. Class, Building or System Summaries provide information about percent of children *delayed*, *developing* and *developed* in each of the six domains and compares those numbers to national averages.

Subscales: General Knowledge, Oral Communication, Written Language, Math Concepts, Work Habits, Attentive Behavior

Norming Sample: 2,108 Kindergarten children in 47 states and 2,939 pre-kindergarten children in 19 states. Demographic information about the children is not reported.

Reliability: High (.80 or higher)

Concurrent Validity: Low (below .50) (reported for kindergarten only)

Teacher Rating of Oral Language & Literacy [TROLL] (2001)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 5 years	5 - 10 minutes	Can be complete by teachers, reading specialists and parents. No training required.	Free**

**www.ciera.org/library reports/inquiry-3/0-016/3-016.pdf

Available Languages: English

Source: Center for the Improvement of Early Reading Achievement, University of Michigan School of Education, Rm. 2002 SEB, 610 E. University Ave., Ann Arbor, MI 48109-1259
(734) 647-6940; www.ciera.org

Administration: Teacher or parent reports on child's skills based on observations. Includes 25 items, each rated on a 4-point scale.

Scores Available: Raw scores are converted to percentile groups. Recommendations and meaning for each percentile group are provided.

Subscales: Language Use, Reading, Writing

Norming Sample: Over 900 children in the Northeastern U.S.; all low-income, high-risk. Not nationally representative.

Reliability: High (.80 or higher)

Validity: Low (below .50) (significant, but low [below .50], correlations are reported with direct measures of children's early literacy skills)

Note: Intended as a system for teachers to quickly and consistently monitoring children's literacy progress.

Transdisciplinary Play Based Assessment, Second Edition [TPBA 2] (2008)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 6 years	May be available in Administration Guide	May be available in Administration Guide.	$55

Available Languages: English

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624
(800) 638-3775; www.brookespublishing.com

Administration: While the child engages in natural play with a parent, professional watches and gathers qualitative information about what the child can do.

Scores Available: May be available in Administration Guide

Subscales: May be available in Administration Guide

Norming Sample: Not normed

Reliability: May be available in Administration Guide

Concurrent Validity: May be available in Administration Guide

Note: Curriculum-based assessment, tied to the Transdisciplinary Play-Based Intervention, *Second Edition* (TPBI2).

Work Sampling System, [WSS] (1998)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 years – 6th grade	About 15 minutes per checklist	Completed by teachers. Workshops available & instructions in manual.	$90

Available Languages: English, Spanish

Source: Pearson Learning Group, 145 South Mount Zion Road, P.O. Box 2500, Lebanon, IN 46052 (800) 321-3106; www.pearsonlearning.com

Administration: Includes three elements: Portfolios, Developmental Guidelines and Checklists, and Summary Reports. Portfolios are used to track a child's efforts, achievements, and progress throughout the year by collecting relevant student work. Teachers rate each construct on the Developmental Checklists as: *Not Yet*, *In Progress*, or *Proficient*. A Summary Report is to be prepared three times a year (replacing conventional report cards), rating each Functional Component on Performance (*Developing as Expected* or *Needs Development*) and Progress (*As Expected* or *Other Than Expected*). Teachers can also add comments to the ratings.

Scores Available: Total scores on in each domain are tracked over time to monitor progress. Electronic reports available on-line (www.worksamplingonline.com)

Subscales: Includes 7 domains: Personal and Social Development, Language and Literacy, Mathematical Thinking, Scientific Thinking, Social Studies, the Arts, and Physical Development and Health. Each domain is further broken down into Functional Components.

Norming sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Young Children's Achievement Test [YCAT] (2000)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
4:0 – 7:11 years	25 – 45 minutes	Bachelor's degree in related field and/or training in testing.	$210

Available Languages: English, Spanish (Prueba de Habilidades Adacémicas Inciciales (PHAI)

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897 (800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child.

Scores Available: Standard scores, percentiles, age equivalents, normal curve equivalents (NCEs), z-scores, T-scores, and stanines.

Subscales: General Information, Reading, Mathematics, Writing, Spoken Language, Early Achievement Composite

Norming Sample: 1,224 children from 32 states, representative of the U.S. population with regard to geographic region, gender, race, residence, ethnicity, family income, parents' educational attainment, and disability status (1997 Census).

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69) (with the Slosson Intelligence Test for Children and Adults-Revised)

Collection of Diagnostic Tools

Page	Name
122	Achenbach System of Empirically Based Assessment (ASEBA) – School Age Forms
122	Adaptive Behavior Assessment System, 2nd Edition (ABAS-II)
123	Arizona Articulation Proficiency Scale, 3rd Edition
123	Battelle Developmental Inventory, 2nd Edition (BDI-II)
124	Bayley Scales of Infant and Toddler Development, 3rd Edition (Bayley-III)
124	Behavior Assessment System for Children, 2nd Edition (BASC-2)
124	Behavioral and Emotional Rating Scale, 2nd Edition (BERS-2)
125	Bilingual Verbal Ability Test (BVAT), Normative Update
125	Boehm Test of Basic Concepts, Preschool, 3rd Edition
126	Boehm Test of Basic Concepts, 3rd Edition
127	Bracken Basic Concept Scale – Revised (BBCS-R)
106	Brigance Comprehensive Inventory of Basic Skills-Revised (CIBS-R) *
106	Brigance Diagnostic Inventory of Early Development II (IED-II) *
127	Bruininks-Oseretsky Test of Motor Proficiency, 2nd Edition (BOT-2)
128	Clinical Evaluation of Language Fundamentals, 4th Edition (CELF-4)
128	Clinical Evaluation of Language Fundamentals – Preschool, 2nd Edition (CELF-P 2)
129	Communication and Symbolic Behavior Scale Developmental Profile, First Normed Edition (CSBS DP)
129	Comprehensive Assessment of Spoken Language (CASL)
130	Comprehensive Test of Phonological Processing (CTOPP)
130	Conners 3rd Edition (Conners 3)
131	Conners Comprehensive Behavior Rating Scales (CBRS)
131	Developmental Assessment of Young Children (DAYC)
132	Differential Ability Scales – II (DAS-II)
132	Early Reading Diagnostic Assessment, 2nd Edition (ERDA 2)
133	Expressive One-Word Picture Vocabulary Test (EOWPVT)
133	Expressive Vocabulary Test, 2nd Edition (EVT-2)
134	Gilliam Autism Rating Scale, 2nd Edition (GARS-2)
134	Goldman Fristoe Test of Articulation, 2nd Edition (GFTA-2)
135	Infant Toddler Developmental Assessment (IDA)
135	Infant Toddler Sensory Profile
136	Infant Toddler Social & Emotional Assessment (ITSEA)
136	Iowa Test of Basic Skills (ITBS), Forms A & B
137	Kaufman Assessment Battery for Children, 2nd Edition (KABC-II)
137	Kaufman Brief Intelligence Test, 2nd Edition (KBIT-2)
138	Kaufman Survey of Early Academic and Language Skills (K-SEALS)

Page	Name
138	Kaufman Test of Educational Achievement, 2nd Edition (KTEA-II)
139	Khan-Lewis Phonological Analysis, 2nd Edition (KLPA-2)
139	Learning Accomplishment Profile-Diagnostic (LAP-D), 3rd Edition
140	Leiter International Performance Scale-Revised (LEITER-R)
141	Lindamood Auditory Conceptualization Test, 3rd Edition (LAC-3)
141	MacArthur-Bates Communicative Development Inventories (CDI), 2nd Edition
142	Merrill-Palmer-Revised Scales of Development (M-P-R)
142	Metropolitan Readiness Test (MRT6), 6th Edition
143	Mullen Scales of Early Learning (MSEL)
143	OWLS: Listening Comprehension (LC) Scale & Oral Expression (OE) Scale
144	OWLS: Written Expression (WE) Scale
144	Peabody Developmental Motor Scales, 2nd Edition (PDMS-2)
145	Peabody Individual Achievement Test-Revised, Normative Update (PIAT-R/NU)
145	Peabody Picture Vocabulary Test, 4th Edition (PPVT-4)
146	Phonological Awareness Test 2 (PAT-2)
146	Pictorial Test of Intelligence, 2nd Edition (PTI-2)
147	Preschool and Kindergarten Behavior Scales, 2nd Edition (PKBS-2)
147	Preschool Language Assessment Instrument, 2nd Edition (PLAI-2)
148	Preschool Language Scale, 4th Edition (PLS-4)
148	Process Assessment of the Learner, 2nd Edition: Diagnostic Assessment for Math (PAL-II Math)
149	Process Assessment of the Learner, 2nd Edition: Diagnostic Assessment for Reading and Writing (PAL-II Reading and Writing)
149	Ready to Learn: A Dyslexia Screener
150	Receptive Expressive Emergent Language Scale, 3rd Edition (REEL-3)
150	Receptive One-Word Picture Vocabulary Test (ROWPVT)
151	Scales of Independent Behavior-Revised (SIB-R)
151	Sensory Profile
152	Social Competence and Behavior Evaluation (SCBE-Preschool Edition)
152	Stanford-Binet Achievement Test, 10th Edition (SAT 10)
153	Stanford Binet Intelligence Scales for Early Childhood, 5th Edition (Early SB5)
153	Temperament and Atypical Behavior Scale (TABS) Screener & Assessment Tool
154	Test for Auditory Comprehension of Language, 3rd Edition (TACL-3)
154	Test of Early Language Development, 3rd Edition (TELD-3)
155	Test of Early Mathematics Ability, 3rd Edition (TEMA-3)
155	Test of Early Reading Ability, 3rd Edition (TERA-3)
156	Test of Language Development - Primary, 4th Edition (TOLD-P:4)
156	Test of Phonological Awareness, 2nd Edition PLUS (TOPA-2+)

Page	Name
152	Test of Preschool Early Literacy (TOPEL)
157	Test of Word Reading Efficiency (TOWRE)
158	Toddler and Infant Motor Evaluation (TIME)
158	Vineland Adaptive Behavior Scales, 2nd Edition (Vineland-II)
159	Wechsler Individual Achievement Test, 2nd Edition (WIAT-II)
159	Wechsler Preschool and Primary Scale of Intelligence, 3rd Edition (WPPSI-III)
160	Woodcock-Johnson III Normative Update Complete (WJ III NU)

* Tools marked with an asterisk have two purposes: *inform instruction/monitor progress* and *diagnostic*.

Their descriptions appear in the Inform Instruction/Monitory Progress section, because that is their primary purpose.

Achenbach System of Empirically Based Assessment [ASEBA], Preschool Forms (2000)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
1:6 – 5 years	10 – 15 minutes	5th grade reading level for checklist. Graduate training for interpretation.	$150

Available Languages: Forms available in more than 75 languages

Source: Achenbach System of Empirically Based Assessment, 1 South Prospect St., Burlington, VT 05401-35456 (802) 264-6432; www.aseba.org

Administration: Parent completes a 99-item checklist about child's behavior (CBCL/1 ½-5) and/or teacher or caregiver completes the 99-item Caregiver-Teacher Report (C-TRF/1 ½-5). There is also an optional additional Language Development Survey (LDS).

Scores Available: T-scores, percentiles, cut-points for normal, borderline, clinical

Subscales: Internalizing, externalizing, and total problems; 7 syndromes (e.g., emotionally reactive, anxious/ depressed, withdrawn, aggressive) and 5 DSM profiles (e.g., affective, anxiety, attention deficit/ hyperactivity). LDS provides language scores.

Norming Sample: 1,728 (CBCL/1 ½ -5) and 1,113 (C-TRF) children. The LDS sample consisted of 278 parents from the CBCL sample. Technical Manual may include additional information about sample characteristics. Norming sample did not include children who had received mental health or special education services (such children were included in scale development).

Reliability: High (.80 or greater)

Concurrent Validity: Adequate (.50 to .69)

Achenbach System of Empirically Based Assessment [ASEBA], School Aged Forms (2001)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
6 – 18 years	15 minutes	5th grade reading level for checklist. Graduate training for interpretation.	$57

Available Languages: Forms available in more than 75 languages

Source: Achenbach System of Empirically Based Assessment, 1 South Prospect St., Burlington, VT 05401-35456 (802) 264-6432; www.aseba.org

Administration: Parent completes the Child Behavior Checklist/6-18 (CBCL/6-18) and/or teachers complete the Teacher's Report Form/6-18 (TRF).

Scores Available: T-scores, percentiles, cut-points for normal, borderline, clinical

Subscales: Total competence; total problems; internalizing; externalizing; 3 competence scales (activities, social, and school); 8 Syndromes (e.g., aggressive, anxious/ depressed, attention problems, rule-breaking, social problems, withdrawn/depressed); 6 DM-oriented scales (e.g., affective, anxiety, attention deficit/hyperactivity, conduct).

Norming Sample: 1,753 (CBCL/6-18) and 2,319 (TRF) nonreferred children representative of the 48 contiguous states with regard to SES, ethnicity, region, and urban-suburban-rural residence (census year unknown). Children who had been referred for mental health or special education services within the past year were excluded (such children were included in scale development).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Notes: Multicultural Options (2007) module provides multicultural norms for ages 6-18.

Adaptive Behavior Assessment System, Second Edition [ABAS-II] (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 89 years	15 – 20 minutes	No training needed to complete checklist; interpretation requires advanced training/experience in testing	$183

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259 (800) 211-8378; harcourtassessment.com

Administration: Teacher/provider or parent rates behavioral frequency of various skills using a 4-point rubric for individual child. Each skill area contains at least 20 items. Forms include: Parent (Birth-5); Parent (5 -21 years); Teacher/Day Care (2-5 years); Teacher (5-21 years). (Adult form also available).

Scores Available: Age-based percentile ranks, test-age equivalents, standard scores, and scaled scores. A supplemental analysis of strengths and weaknesses can be performed based on tables of confidence intervals provided in the manual.

Subscales: Provides an overall composite score (General Adaptive Composite-GAC) and scores (if age appropriate), on each of the 10 DSM-TR adaptive skill areas (Communication, Community Use, Functional Academics, Home-School Living, Health and Safety, Leisure, Self-Care, Self-Direction, Social, and Work) and for the three adaptive domains identified by AAMR (Conceptual, Social, and Practical), plus motor skills.

Norming Sample : Ranged from 750 (Teacher-Daycare Provider Form, 2-5 years) to 1,690 (Teacher Form, 5 to 21 years); closely matched to U.S. population with regard to age, gender, race-ethnicity, education, geographic region (2000 U.S. Census).

Reliability: High (.80 or greater).

Concurrent Validity: Adequate (.50 to 69)

Arizona Articulation Proficiency Scale, Third Edition (Arizona-3) (2001)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
1:6 – 18 years	2 – 10 minutes	Administration and interpretation require a graduate degree in speech pathology.	$150

Available Languages: English

Source: Pro Ed, Inc., 8700 Shoal Creek Boulevard, Austin, Texas 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child by a speech-language professional.

Scores Available: Percentiles, standard scores, & age equivalents, articulatory impairment ratings (normal, mildly impaired, moderately impaired, or severely impaired)

Subscales: None

Norming Sample: Sample included 5,500 individuals between 18 months and 18 years of age, closely matching U.S. population with regard to geographic region, ethnicity, and socioeconomic status (1998 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Note: Clinical tool used to assess difficulties in articulatory proficiency.

Battelle Developmental Inventory, Second Edition [BDI-2] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:0 – 7:11 years	1 – 2 hours	Early childhood teachers, early interventionists, special educators, psychologists, heath professionals, etc. Familiarity with psychometric procedures recommended.	$902

Available Languages: English, Spanish

Source: Riverside Publishing Company, 8420 Bryn Mawr Avenues, Chicago, IL 60631
(800) 767-8420; www.riverpub.com

Administration: Direct assessment and observation of individual child, plus interview with child's parents conducted by early childhood professional.

Scoring: Age equivalents, scaled scores, developmental quotient, z-scores, T-scores, normal curve equivalents, percentile ranks, confidence intervals.

Subscales: Adaptive (Self-Care and Personal Responsibility); Personal-Social (Adult Interaction, Peer Interaction, and Self-Concept and Social Role); Communication (Receptive and Expressive); Motor (Gross, Fine, and Perceptual); Cognitive (Attention and Memory, Reasoning and Academic Skills, and Perception and Concepts).

Norming Sample: 2,500 children closely matching the U.S. population with regard to age, sex, race/ethnicity, region, SES (2001 U.S. Census). Children with disabilities were not included in the norming sample, but were included in reliability and validity studies.

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Bayley Scales of Infant and Toddler Development, Third Edition [Bayley-III] (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:1 – 3:6 years	30 – 90 minutes	Appropriately trained, practitioners, including early intervention and child development specialists, school psychologists, assessment specialists.	$895

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child by trained professional; primary caregiver responds to questionnaire about child's social-emotional and adaptive functioning.

Scores Available: Scaled scores, composite scores, and percentile ranks for each subscale. Growth scores and developmental age scores for the cognitive, language, and motor scales.

Subscales: Cognitive, language (receptive, expressive, total), motor (fine, gross, total), social-emotional, adaptive (e.g., community use, functional pre-academics, home living, health and safety, self-care, self-direction, total), language composite; motor composite.

Norming Sample: For the cognitive, language and motor scales: 1,700 children representative U.S. population in terms of race-ethnicity, age, sex, parent educational level, and geographic location (2000 U.S. Census). For the social-emotional scale, 456 children; for the adaptive behavior scale, 1,350 children.

Reliability: High (.80 or above)

Concurrent Validity: High (.70 or higher)

Behavior Assessment System for Children, Second Edition [BASC-2] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2 years – college age	10 – 30 minutes	Completed by parent or teacher. Graduate training for interpretation.	$385

Available Languages: English, Spanish

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Teacher Rating Scale (TRS, 100-139 items), Parent Rating Scale (PRS, 134-160 items), Structured Developmental History (SDH), and a Student Observation System (SOS) for use with observed classroom behavior. (Self-Report of Personality (SRP) form available for children 8 and older). Three separate age-levels of each form available.

Scores Available: T-scores and percentiles for general clinical populations

Subscales: Aggression, Conduct Problems, Atypicality, Locus of Control, Social Stress, Anxiety, Depression, Somatization, Sense of Inadequacy, Self-Esteem, Self-Reliance, Attention Problems, Learning Problems, Attitude to School, Attitude to Teachers, Sensation Seeking, School Problems, Inattention-Hyperactivity, Adaptability, Social Skills, Leadership, Study Skills, Relations with Parents, Interpersonal Relations, Withdrawal, Anger Control, Bullying, Developmental Social Disorders, Emotional Self-Control, Executive Functioning, Negative Emotionality, Resiliency, Social Stress, Ego Strength, Mania, Test Anxiety.

Norming Sample: General norm samples (TRS = 4,650; PRS = 4,800; SRP = 3,400) were matched to U.S. population with regard to SES, race-ethnicity, region, and special education classification (2001 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Behavioral and Emotional Rating Scale, Second Edition [BERS-2] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5:0 – 18:11 years	10 minutes	Forms completed by parent and/or teacher; graduate training required for interpretation.	$165

Available Languages: English

Source: PRO-ED 8700 Shoal Creek Blvd Austin TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Multi-modal assessment system including three forms: parent (Parent Rating Scale, 57 items) and teacher or other professional (Teacher Rating Scale, 52 items) (Also includes a Youth Rating Scale for ages 11 through 18; 57 items).

Subscales: Strength Index (Composite Score), plus Interpersonal Strength, Family Involvement, Intrapersonal Strength, School Functioning, Affective Strength.

Scores Available: Raw scores, standard score, scaled scores and percentile ranks. Separate norms for children who have and have not been diagnosed an emotional or behavioral disorder (EBD).

Norming Sample: TRS**:** 2,176 w/o EBD diagnosis; 816 with EBD diagnosis; PRS: 927 children. The demographic characteristics of the standardization sample are generally equivalent to the most recent 2001 U.S. Census data.

Reliability: High (.80 or higher)

Concurrent Validity: May be available in Technical Manual

Bilingual Verbal Ability Test [BVAT] Normative Update (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5 years – adult	30 minutes	Must be bilingual and have training in testing or be supervised by someone with training in testing.	$503

Available Languages: Kits available in English & Spanish. Test records available in Arabic, Chinese, English, German, Hindi, Korean, Polish, Turkish, French, Russian, Haitian-Creole, Italian, Japanese, Portuguese, Spanish, and Vietnamese

Source: Riverside Publishing, 3800 Golf Road, Suite 100, Rolling Meadows, IL 60008
(800) 323-9540; www.riverpub.com

Administration: Direct assessment of individual child. Examiner administers the three subtests in English first; any item that was answered incorrectly or skipped is then administered in the native language.

Scores Available: Age equivalents, grade equivalents, instructional zones, relative proficiency index (RPI), percentile ranks, standard scores, W score, T score, NCE, z score, stanine, and CALP levels.

Subscales: Overall verbal ability (bilingual), English language proficiency;

Norming Sample: 8,818 subjects in more than 100 geographically diverse U.S. communities, representative of the U.S. population with regard to community size, gender, race/ethnicity, schooling, and adults' education, occupational status and specific occupation (2000 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Note: Measures bilingual verbal ability in English and another language. Also yields an aptitude measure that can be used in conjunction with the WJIII Tests of Achievement.

Boehm Test of Basic Concepts, Preschool, Third Edition (2001)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3:0 – 5:11 years	20 – 30 minutes	Speech-language pathologists, school psychologists, educators, diagnosticians familiar with standardized testing procedures.	$166

Available Languages: English, Spanish

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202 www.proedinc.com

Administration: Direct assessment of individual child; 76 items (with basal/ceiling rules)

Scores Available: Raw Scores, percentile, performance range (upper, middle or lower third), percent correct

Subscales: None

Norming Sample: 660 children representative of the U.S. population with regard to gender, race-ethnicity, region, and parental education (census year unknown).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Notes: Designed to assess young children's understanding of the basic relational concepts (size, direction, position in space, time, quantity, classification, and general) that are important for language and cognitive development. Includes a parent form to use in reporting results to parents; provides suggestions for home-activities.

Boehm Test of Basic Concepts, Third Edition (2000)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
K – 2nd grade	30 – 45 minutes	Speech-language pathologists, school psychologists, educators, and diagnosticians; requires familiarity with standardized testing procedures.	$99

Available languages: English and Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Group-administered in a classroom setting or individual assessment; 50 items.

Scores Available: Raw scores, percent correct, performance range, and percentiles by grade, with separate norms for fall and spring testing for grades K, 1, and 2

Subscales: More than 6,000 students in the fall testing session and 4,000 in the spring session, representative of the U.S. population. Census year and specific sample characteristics may be available in the Examiner's Manual.

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Notes: Designed to assess student understanding of basic concepts, such as those that describe qualities of people-objects, spatial relationships, time, and quantity.

Bracken Basic Concept Scale–Revised [BBCS–R] (1998)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:6 – 8 years	30 minutes	Individuals knowledgeable in the administration and interpretation of assessments (e.g., school psychologists, special education teachers).	$375

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child; 308 items.

Scores Available: Subtest and Composite, Percentile Ranks, Scaled scores, Standard Scores, Confidence Intervals and Concept Age Equivalents. Manual includes the data needed to conduct ipsative interpretations of a child's performance.

Subscales: Colors, Letters, Numbers/Counting, Sizes, Comparisons, Shapes, Direction/Position, Self/Social Awareness, Texture/Materials, Quantity, Time/Sequence

Norming Sample: 1,100 children, representative of the U.S. population of children with regard to age, gender, race/ethnicity, region and parent education (1995 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Bruininks–Oseretsky Test of Motor Proficiency, Second Edition [BOT–2] (2006)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
4 – 21 years	45 – 60 minutes	Physical therapists, occupational therapists, adaptive physical education teachers, special education professionals	$749

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child, 53 items.

Scores Available: For each subtest and composite: scale scores, stand scores, confidence intervals, percentile rank, age equivalent, descriptive categories (e.g., average, above average, below average). Profile analysis to evaluate strengths and weaknesses.

Subscales: 8 subtests (Fine Motor Precision, Manual Dexterity, Bilateral Coordination, Balance, Running Speed and Agility, Upper-Limb Coordination, Strength); 5 composites: (Fine Motor Control, Manual Coordination, Body Coordination, Strength & Agility, Total)

Norming Sample: 1,520 children in 239 sites in 38 states, closely matched to US population in terms of gender, race/ethnicity, socio-economic status, and mother's education (2001 U.S. Census) and special education status (per the U.S. Department of Education).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Clinical Evaluation of Language Fundamentals, Fourth Edition [CELF–4] (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5 – 21 years	30 –60 minutes	Professional with training and experience in test administration and interpretation.	$469

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child.

Scores Available: Standard scores, percentile ranks, age equivalents, criterion-referenced cut-points, item analysis

Subscales: *Composites:* Core Language, Receptive Language, Expressive Language, Language Content, Language Structure, and Working Memory. *Subtests*: Concepts and Following Directions, Core Language, Expressive Language, Expressive Vocabulary, Familiar Sequences, Formulated Sentences, Language Content, Language Memory, Language Structure, Number Repetition, Phonological Awareness, Rapid Automatic Naming, Recalling Sentences, Receptive Language, Semantic Relationships, Sentence Assembly, Sentence Structure, Understanding Spoken Paragraphs, Word Associations, Word Classes-Receptive, Word Classes-Expressive, Word Definitions, Word Structure

Norming Sample: 2,650 students, representative of the U.S. population with respect to age, gender, race/ethnicity, region, and parent education (2000 U.S. Census). Sample included children receiving special education services.

Reliability: High (.80 or above)

Concurrent Validity: Adequate (.50 to .69)

Clinical Evaluation of Language Fundamentals – Preschool, Second Edition [CELF–P 2] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 6:11 years	30 – 45 minutes	Professional with training and experience in test administration and interpretation.	$329

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child.

Scores Available: Scaled scores, standard scores, percentile ranks, confidence intervals, criterion-referenced cut-points, item analysis.

Subscales: *Composites*: Core Language, Receptive Language, Expressive Language, Language Content, Language Structure. *Subtests*: Sentence Structure, Word Structure, Expressive Vocabulary, Concepts and Following Directions, Recalling Sentences, Basic Concepts, Word Classes, Recalling Sentences in Context, Phonological Awareness, Pre-Literacy Rating Scale, Descriptive Pragmatics Profile

Norming Sample: 800 children, representative of U.S. population in terms of gender, race/ ethnicity, region; and primary-caregiver education (2000 U.S. Census). Sample included children with children receiving special education services.

Reliability: High (.80 of higher)

Concurrent Validity: High (.70 or higher)

Communication and Symbolic Behavior Scale Developmental Profile, First Normed Edition [CSBS DP] (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:6 – 2:0 years * (up to 6:0 years for atypical development)	1 hour	Certified speech-language pathologist, early interventionist, psychologist, pediatrician, or other professional.	$399

Available Languages: English

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624 (800) 638-3775 www.brookespublishing.com

Administration: Parent or primary caregiver completes the Infant Toddler Checklist (24 multiple choice items). If the checklist indicates concern, caregiver completes the Caregiver Questionnaires and the professional administers a Behavior Sample.

Scores Available: Standard scores and percentiles. Cut-points provided for the Infant Toddler Checklist (concern vs. no concern).

Subscales: Social Composite (Emotion and Eye Gaze, Communication, Gestures); Speech Composite (Sounds, Words); Symbolic Composite (Understanding, Object Use)

Norming Sample: Infant Toddler Checklist: 2,188 children; Caregiver Questionnaire: 790 children; Behavior Sample: 337 children. Children were primarily from the Tallahassee, FL area. Not nationally representative.

Reliability: High (over .80)

Concurrent Validity: May be available in Technical Manual.

Note: For assessing children at risk for communication delays and impairments.

Comprehensive Assessment of Spoken Language [CASL] (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 21 years	30 – 45 minutes	Speech/language pathologists, psychologists, educational diagnosticians, early childhood specialists, and other professionals.	$364

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437 (877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child.

Scores Available: Age-based and grade-based standard scores, grade and test-age equivalents, percentiles, normal curve equivalents (NCEs), and stanines

Subscales: Basic Concepts, Antonyms, Sentence Completion, Syntax Construction, Paragraph Comprehension, Pragmatic Judgment, Synonyms, Sentence Completion, Idiomatic Language, Grammatical Morphemes, Sentence Comprehension, Grammaticality Judgment, Nonliteral Language, Meaning from Context, Inference, Ambiguous Sentence, Lexical-Semantic Index, Syntactic Index, Supralinguistic Index, Receptive Index, Expressive Index, Core Composite.

Norming Sample: 1,700 examinees, stratified to match 1994 U.S. Census data on gender, race-ethnicity, region, and mother's education level

Reliability: High (.80 or above)

Concurrent Validity: May be available in the Technical Manual.

Note: Tests require no reading or writing ability.

Comprehensive Test of Phonological Processing [CTOPP] (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5 – 24:11 years	30 minutes	Professional with training in testing.	$254

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437 (877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child. Two versions: one for 5 and 6 year olds, and one for ages 7 through 25. (See also Preschool Comprehensive Test of Phonological Processing.)

Scores Available: Percentiles, standard scores, and age and grade equivalents

Subscales: 3 quotients: Phonological Awareness, Phonological Memory, and Rapid Naming; 14 subtests: Elision, Blending Words, Sound Matching, Memory for Digits, Nonword Repetition, Rapid Color Naming, Rapid Digit Naming, Rapid Letter Naming, Rapid Object Naming, Blending Nonwords, Phoneme Reversal, Segmenting Words, and Segmenting Nonwords.

Norming Sample: Over 1,600 people, representative of the U.S. population as a whole with regard to gender, race, ethnicity, residence, family income, educational attainment of parents, and geographic regions, as reported in the 1997 Statistical Abstract of the United States.

Reliability: High (.80 or above)

Concurrent Validity: Adequate (.50 to .69)

Conners Third Edition [Conners 3] (2008)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
6 – 18 years	20 minutes	Interpretation requires graduate-level courses in tests/measurement or equivalent.	$280

Available Languages: English, Spanish

Source: Multi-Health Systems Inc., P.O. Box 950, North Tonawanda, NY 14102-0950; (800) 456-3003; www.mhs.com

Administration: Multi-assessment tool that gathers information from parent, teacher and youth (8-18 years). Short (about 40 items) and long (110 items) forms available for each reporter. Ratings are made on a 4-point scale and are based on the past month

Scores Available: T-Scores, Percentiles, and Standard Error of Measurement

Subscales: See Technical Manual for full list. Examples include: Executive Functioning, Learning Problems, Aggression, Peer Relations, Family Relations, Inattention Hypteractivity/Impulsivity, Oppositional Defiant Disorder, Conduct Disorder

Norming sample: Approximately 2,900 parent or teacher forms from the general population and over 1,000 parent or teacher forms from the clinical population; representative of the general U.S. population in terms of ethnicity/race, gender, and parent education level (U.S. Bureau of the Census, 2000).

Reliability: High (.80 or above)

Concurrent Validity: High (.70 or higher)

Note: Assessment of ADHD, comorbid disorders, and associated features; co-normed with the CBRS; preschool version scheduled for publication in 2009.

Conners Comprehensive Behavior Rating Scales [CBRS] (2008)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
6 – 18 years	20 minutes	Interpretation requires graduate-level courses in tests/measurement or equivalent.	$150

Available Languages: English, Spanish

Source: Multi-Health Systems Inc., P.O. Box 950, North Tonawanda, NY 14102-0950
(800) 456-3003; www.mhs.com

Administration: Multi-assessment tool that gathers information from parent (198 items), teacher (204 items), and youth (8-18 years; 181 items). Ratings are made on a 4-point scale and are based on the past month

Scores Available: T-Scores, Percentiles, and Standard Error of Measurement

Subscales: See Technical Manual for full list. Examples include: Emotional Distress, Aggressive Behaviors, Academic Difficulties, Hyperactivity/Impulsivity, Social Problems, Violence Potential, Oppositional Defiant Disorder, Major Depressive Episode, Generalized Anxiety Disorder, Bullying (Perpetration, Victimization), Post Traumatic Stress, Phobia.

Norming sample: Approximately 2,900 parent or teacher forms from the general population and over 1,000 parent or teacher forms from the clinical population; representative of the general U.S. population in terms of ethnicity/race, gender, and parent education level (U.S. Bureau of the Census, 2000).

Reliability: High (.80 or above)

Concurrent Validity: High (.70 or higher)

Note: Assesses a wide range of behavioral, emotional, social, and academic issues in school-aged youth; co-normed with the Conners 3; preschool version scheduled for publication in 2009.

Developmental Assessment of Young Children [DAYC] (1998)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 5:11 years	10 – 20 minutes per subtest	Four-year degree psychology, counseling, or a closely related field, plus satisfactory completion of course-work in testing.	$235

Available Languages: English

Source: Psychological Assessment Resources; 16204 North Florida Avenue, Lutz FL 33549
(800) 331-8378; www.parinc.com

Administration: Information gathered through observation, interview of caregivers, and direct assessment. Each of the 5 subscale consists of between 58 and 87 items.

Scores Available: Subtest raw scores, subtest age equivalents, subtest standard scores, a quotient standard score, and percentiles.

Subscales: Cognition, Communication, Social/Emotional Development, Physical Development, and Adaptive Behavior

Norming Sample: National sample of 1269 individuals, residing in 27 states, comparable 1996 U.S. Census in terms of geographic region, gender, race, rural or urban residence, ethnicity, family income, educational attainment of parents, and disability status.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Note: Subscales can be administered separately or as a group.

Differential Ability Scales – II [DAS-II] (2007)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:6 – 17:11 years	Core Battery: 45 – 60 minutes	Administered by a psychologist.	$1,080

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259 (800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child.

Scores Available: Ability scores, *T* scores, cluster scores, composite scores and percentile ranks. In addition, confidence intervals are available for the cluster and composite scores; standard error of measurement information and age equivalents are provided for the subtest ability scores. For children ages 5:0-8:11 years, there is also a School Readiness cluster that measures three sets of abilities related to early school success and failure.

Subscales: *Core battery*: Verbal, Nonverbal, and Spatial reasoning; *Diagnostic subtests*: Copying, Early Number Concepts, Matching Letter-Like Forms, Matrices, Naming Vocabulary, Pattern Construction, Phonological Processing, Picture Similarities, Rapid Naming, Recall of Designs, Recall of Digits, Recall of Objects, Recall of Sequential Order, Recognition of Pictures, Sequential and Quantitative, Verbal Similarities, Speed of Information Processing, Verbal Comprehension, Word Definitions

Norming sample: May be available in Technical Manual.

Reliability: May be available in Technical Manual.

Concurrent Validity: May be available in Technical Manual.

Note: Primary purpose is diagnostic, but reported to be useful in informing instruction.

Early Reading Diagnostic Assessment, Second Edition [ERDA 2] (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
Kindergarten – 3rd grade	45 – 60 minutes	No specific training or qualifications required.	$320 per grade

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259; (800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child.

Scores Available: Grade-based percentiles for fall, winter and spring; also cut-points for emerging/below basic, basic, proficient.

Subscales: Concepts of Print, Phonological and Phonemic Awareness (phonemes, rhyming, syllables, rimes), Phonics (letter recognition, pseudoword decoding), Fluency (passage fluency, target words in context, word reading, RAN), Vocabulary (receptive/expressive, word opposites, synonyms, word definitions), Comprehension (story retelling, reading comprehension, listening comprehension). Six composite scores: Brief Vocabulary, Full Vocabulary, RAN-Automaticity, Phonological Awareness, Narrative Passage, and Information Passage Fluency.

Norming Sample: 800 students, representative of the U.S. population in terms of grade, gender, race-ethnicity, geographic region, and parent education (U.S. Census 2000). Norming sample excluded students who were receiving special education and those identified as Limited English Proficient.

Reliability: Adequate (.65 to .79)

Concurrent Validity: Not reported (See Technical Manual for other types of validity.)

Expressive One-Word Picture Vocabulary Test [EOWPVT] (2000)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2 – 18:11 years	15 – 20 minutes	Can be administered by non-professional with training. Requires graduate training and experience for interpretation.	$159

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child; includes 170 test items with basal and ceiling rules.

Scores Available: Raw scores can be converted into a standard score, percentile rank, age equivalent, normal curve equivalent (NCEs), scaled scores, T-scores, and stanines.

Subscales: None

Norming Sample: 2,327 children closely matched to the 1998 US Census data with regard to region, race/ethnicity, gender, parental education level, urban versus rural residence, and disability status.

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Note: Measures expressive vocabulary only. See ROWPVT for a measure of receptive vocabulary. Spanish version is for ages 4 to 12:11 years.

Expressive Vocabulary Test, Second Edition [EVT-2] (2007)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:6 – 90 years	10 – 20 minutes	Bachelor's degree, including coursework in principles of measurement and in the administration and interpretation of tests, and formal training speech and language.	$379

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child; 2 forms (A & B), 190 test items per form; with basal and ceiling rules

Scores Available: Age- and grade-based standard scores, Growth Scale Values (GSVs), percentiles, stanines, normal curve equivalents (NCEs), age and grade equivalents

Subscales: None (test of expressive vocabulary only)

Norming Sample: 3,450 individuals, closely matched to the 2004 Census data with regard to age, sex, race/ethnicity, geographic region, parent education. The sample of children ages 2 to 18 included representative proportions of children with special needs.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Note: Co-normed with the PPVT-4.

Gilliam Autism Rating Scale, Second Edition [GARS-2] (2006)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 22 years	5 – 10 minutes	School psychologist, educational diagnostician, autism specialist, speech-language pathologist, or a similarly trained professional who knows how to interpret test information to diagnose autism.	$131

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Examiner completes a total of 42 items (14 per subscale) based on the frequency of occurrence from 0 (never observed) to 3 (frequently observed). The examiner can collect information for use in completing tool using a 25-item parent interview form and from direct observation of the child.

Scores Available: Subscale standard scores, percentile ranks, Autism Index, Probability of Autism classification (very likely, possibly, unlikely).

Subscales: Stereotyped Behaviors, Communication, Social Interaction

Norming Sample: 1,107 individuals with autism in 48 states, representative of the U.S. 2000 census with regard to race and geographic region.

Reliability: High (.80 or higher)

Concurrent Validity: May be available in the Technical Manual.

Note: Designed to help in the identification and diagnosis of autism and estimate the severity of the disorder.

Goldman-Fristoe Test of Articulation, Second Edition [GFTA-2] (2000)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2 – 21:11 years	15 minutes for Words-in-Sounds; other two subtests vary	Graduate training & clinical experience in speech/language pathology.	$245

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child.

Scores Available: Raw scores, age-based standard scores, test-age equivalents

Subscales: Sounds-in-Words, Sounds-in-Sentences, Stimulability.

Norming Sample: 2,350 examinees stratified to match the 1998 U.S. Census data on gender, race/ethnicity, region, and SES as determined by mother's education level.

Reliability: High (.80 or higher)

Concurrent Validity: Not reported

Note: Designed to measure of articulation of consonant sounds. Designed as a companion tool to the Khan-Lewis Phonological Analysis, Second Edition [KLPA-2]

Infant – Toddler Developmental Assessment [IDA] (1995)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 3:0 years	Varies	Team of trained clinicians (2 or more). Training tapes available.	$619

Available Languages: English, Spanish

Source: Riverside Publishing, 3800 Golf Road, Suite 100, Rolling Meadows, IL 60008
(800) 323-9540; www.riverpub.com

Administration: There are six IDA phases designed to be conducted by a team of two or more professionals: (1) Referral & Pre-interview Data Gathering, (2) Initial Parent Interview, (3) Health Review,

(4) Developmental Observation and Assessment, (5) Integration and Synthesis, and (6) Share Findings, Completion, and Report.

Scores Available: Percentage delay in each domain; areas of competency and concern ARE be identified.

Subscales: 8 domains: Gross Motor, Fine Motor, Relationship to Inanimate Objects (Cognitive), Language/ Communication, Self-Help, Relationship to Persons, Emotions and Feeling States (affects), and Coping Behavior

Norming Sample: Sample of 100 children, not nationally representative, using the fourth phase of the assessment.

Reliability: High (.80 or higher)

Concurrent Validity: Not reported (84-100 percent agreement on developmental age with Bayley and Vineland)

Infant Toddler Sensory Profile (2002)

Age Range	Time to Administer	Administrator Required Training Needed	Cost
0 – 3:0 years	15 minutes	No qualifications/training required.	$162

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Caregivers complete a questionnaire reporting the frequency (*almost always* to *almost never*) with which infants respond to various sensory experiences; separate forms for birth to 6 months (36 items) and 7 to 36 months (48 items).

Scores Available: Scores calculated for each "quadrant" (low registration, sensation seeking, sensory sensitivity, sensation avoiding, low threshold). For birth to 6 months, cut-points are provided for "typical performance" or "consult and follow-up." For 7 to 36 months, cut-points are provided for "typical performance," "probable difference" or "definite difference." A quadrant grid summarizes scores.

Subscales: General processing, auditory processing, visual processing, tactile processing, vestibular processing, oral sensory processing.

Norming Sample: 809 children without disabilities for creation of the scoring structure. Technical Manual may include details about sample's demographic characteristics. Children with disabilities were included in other aspects of the standardization process.

Reliability: Adequate (.65 to .79)

Concurrent Validity: Adequate (.50 to .69)

Notes: Designed to evaluate sensory-related difficulties and understand how sensory processing affects the child's daily functioning performance.

Infant Toddler Social and Emotional Assessment [ITSEA] (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
1:0 – 3:0 years	25 - 30 minutes	Self-administration requires 4th – 6th grade reading level. Interpretation requires training and experience in standardized testing.	$199

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Parent (PF) or child care (CPF) provider report (self-administered or interview); PF includes 170 items; CPF includes an additional 7 items specific to child care.

Subscales: Three problem domains (Externalizing, Internalizing, Dysregulation) and one competence domain with three to six subscales per domain. Maladaptive, Social Relatedness, and Atypical Item Clusters are also scored.

Scores Available: T-scores and percentile ranks divided by 6-month age bands and gender. Cut-points provided for "of concern."

Norming Sample: 600 children from 42 states, similar to U.S. population with regard to ethnicity, parent education, and region (2002 U.S. Census).

Reliability: High (.80 of higher)

Concurrent Validity: Adequate (.50 to .69)

Notes: Designed to identify children and caregivers who may benefit from additional dialogue about children's development to determine the presence of abnormal behaviors, psychopathology, or delays. By itself, it is not sufficient to make such a determination.

Iowa Test of Basic Skills [ITBS], Form A (2001), Form B (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5 – 14 years	30 minutes per test	Professionals such as teachers, administrators, and counselors, with appropriate training.	$151 per level

Available Languages: English

Source: Riverside Publishing Company, 8420 Bryn Mawr Avenues, Chicago, IL 60631
(800) 767-8420; www.riverpub.com

Administration: Group administered test of individual child skills. Tests are ordered by levels (5 to 14) corresponding to target ages. Total number of items in the Complete battery varies from 146 (Level 5) to 515 (Level 14); Core and Survey Batteries have fewer items.

Scores Available: Raw scores, percent-correct scores, grade equivalents, developmental standard scores, percentile ranks, stanines, and normal curve equivalents are available for most content areas. Item-by-item normative data are also provided.

Subscales: For ages 5 - 8: Vocabulary, Word Analysis, Listening, Language, Reading Words, Reading Comprehension, Spelling, Mathematics, Math Concepts, Math Problems, Math Computation, Social Studies, Science, Sources of Information, Composite, Reading Total, Math Total, Reading Profile Total, Survey Battery Total, Core Total.

Norming Sample: Approximately 170,000 students in spring; 76,000 students in fall. Samples were designed to represent the national population of children in grades K to 8 with regard to public-private schools, geographic regions, SES, district/diocese size, grade level, and race/ethnicity (census year may be available in the Technical Manual).

Reliability: High (.80 or higher)

Concurrent Validity: Not reported. (Guide to Research and Development provides evidence of predictive validity for grades 8-12 and discussion of other types of validity.)

Kaufman Assessment Battery for Children, Second Edition [KABC-II] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 18 years	25 – 55 minutes	Graduate training, plus experience in administering and interpreting individual assessments.	$795

Available Languages: English, Spanish

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child.

Scores Available: Age-based standard scores, age equivalents, and percentile ranks**,** Score Summary Table, Graphic Profile, Narrative Report, Planned Clinical Comparisons, Ability/Achievement Discrepancy

Subscales: Short Term Memory, Visual Processing, Long-Term Storage and Retrieval, Fluid Reasoning, and Crystallized Ability, scales, yielding a Fluid-Crystallized Index composite. The Nonverbal scale yields the Nonverbal Index.

Norming Sample: 3,025 children, closely representing the U.S. population with regard to gender, ethnicity, parental education, region, and educational and psychological classifications (i.e., gifted and talented and specific learning disabled) (2001 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: May be available in the Technical Manual.

Note: Co-normed with the KTEA-II to allow for comparison of ability and performance.

Kaufman Brief Intelligence Test, Second Edition [KBIT-2] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
4:0 – 90 years	20 minutes	Can be administered by trained technicians or para-professionals; properly qualified professional needed for interpretation of results.	$219

Available Languages: English, Spanish

Source: Pearson Learning Group, 145 South Mount Zion Road, P.O. Box 2500, Lebanon, IN 46052
(800) 321-3106; www.pearsonlearning.com

Administration: Direct assessment of individual child.

Scores Available: Standard scores, 90% confidence intervals, percentile ranks, descriptive categories, and age-equivalents.

Subscales: Crystallized (verbal), Fluid (nonverbal), IQ composite

Norming Sample: 2,120 individuals, closely matched to the U.S. population with regard to race/ethnicity and education (2001 U.S. Census). Southern region of the U.S. was over-represented, and the Northeastern region was under-represented. Special education and gifted-talented students were included at school age.

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Kaufman Survey of Early Academic and Language Skills [K-SEALS] (1993)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3:0 – 6:11 years	15 - 25 minutes	Teachers, speech and language professionals, medical professionals.	$262

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child.

Scores Available: Age-based standard scores, percentile ranks, descriptive categories, and age equivalents. Performance on the Articulation Survey subtest can be interpreted using descriptive categories (Normal, Below Average, Mild Difficulty, or Moderate to Severe Difficulty) and item error analysis procedures.

Subscales: Expressive Skills, Receptive Skills, Number Skills, Letter & Word Skills, plus an Early Academic and Language Skills composite.

Norming Sample: 1,000 individuals similar to the U.S. population with regard to gender, race, region, and parent education (1990 U.S. census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Kaufman Test of Educational Achievement, Second Edition [KTEA-II] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
4:6 – 90+ years	15 – 80 minutes	Graduate training, plus experience in administering and interpreting assessments.	$332

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child.

Scores Available: Age- and grade-based standard scores, age and grade equivalents, percentile ranks, normal curve equivalents (NCEs), stanines, error analysis, prescriptive information for remediation planning.

Subscales: *Reading Composite*: Letter and Word Recognition, Reading Comprehension; *Other Reading Related subtests*: Phonological Awareness, Nonsense Word Decoding, Word Recognition Fluency, Decoding Fluency, Associational Fluency, Naming Facility; *Math Composite:* Math Concepts and Applications, Math Computation; *Oral Language Composite:* Listening Comprehension, Oral Expression; *Written Language Composite:* Written Expression, Spelling; *Other Composites*: Comprehensive Achievement, Decoding, Oral Fluency, Reading Fluency, Sound-Symbol

Norming Sample: 3,025 children, closely representing U.S. population with respect to gender, ethnicity, parent education, region, and educational and psychological classifications (i.e., gifted and talented and specific learning disabled) (2001 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher) (reading, writing, oral composites)

Note: Co-normed with the KABC-II to allow for comparison of ability and performance.

Khan-Lewis Phonological Analysis, Second Edition [KLPA-2] (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2 – 21 years	10 - 30 minutes	Graduate training and clinical experience in speech/ language pathology.	$144

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437 (877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child.

Scores Available: Standard Score, Percentiles, Test-Age Equivalents, and percent-of-occurrence.

Subscales: None

Norming sample: 2,350 examinees stratified to match the U.S. population with regard to gender, race/ethnicity, region, and mother's education level (1998 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: Not reported.

Note: Provides analysis of overall phonological process usage. Designed as a companion tool to the Goldman-Fristoe Test of Articulation-Second Edition (GFTA-2).

Learning Accomplishment Profile – Diagnostic [LAP-D], Third Edition (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:6 – 6:0 years	60 – 90 minutes	Trained teachers, paraprofessionals, clinicians, special educators, speech-language pathologists, and others familiar with child development.	$800

Available Languages: English, Spanish

Source: Kaplan Early Learning Company, 1310 Lewisville-Clemmons Road, Lewisville, NC 27023 (800) 334-2014; www.kaplanco.com

Administration: Direct assessment of individual child. Can be administered one-to-one or in a station format. Includes 226 items, with basal and ceiling rules.

Scores Available: Raw scores, percentile ranks, Z-scores, T-Scores, Normal Curve Equivalents (NCE), and age equivalent scores, specific developmental skill data.

Subscales: Fine Motor (Writing, Manipulation), Cognitive (Counting, Matching), Language (Naming, Comprehension), Gross Motor (Body Movement, Object Movement)

Norming Sample: 2,099 children (1,124 English-speaking; 975 Spanish-speaking) from five areas throughout the U.S.; sample was representative of the U.S. population with regard to language, region, age, race, gender (2000 US Census); 77 had a diagnosed disability.

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Leiter International Performance Scale–Revised [LEITER–R] (1997)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:0 – 20:11 years	75 minutes	Degree in psychology, counseling, or related field, plus courses in testing.	$939

Available Languages: English

Source: Psychological Assessment Resources; 16204 North Florida Avenue, Lutz FL 33549 (800) 331-8378; www.parinc.com

Administration: Direct assessment of individual child. Optional examiner, parent, teacher, and self (ages 9 and older) rating scales.

Scores Available: Raw scores, scaled scores, growth scores, percentile ranks, NCEs.

Subscales: 20 subtests that comprise two batteries, Visualization and Reasoning (VR) and Attention and Memory (AM). See Technical Manual for full list of subtests.

Examples of VR subtests: Figure Ground, Design Analogies, Form Completion, Matching, Sequential Order, Repeated Patterns, Classification, Paper Folding, Figure Rotation

Examples of AM subtests: Immediate Recognition, Forward Memory, Attention Sustained, Reverse Memory, Spatial Memory, Delayed Pairs, Delayed Recognition, Attention Divided

Also provides VR Composite Score, AM Composite Score, Full IQ, Brief IQ, Memory Screen.

Examples of Examiner Rating subscales: Attention, Activity Level, Sociability, Anxiety

Examples of Parent Rating subscales: Attention, Activity Level, Impulsivity, Adaptation

Norming Sample: 1,719 individuals for the VR battery and 763 individuals for the AM battery; closely matched to the U.S. population with regard to gender, ethnicity, parental education, and region (1993 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Note: Designed as a non-verbal test of intelligence and cognitive ability.

Lindamood Auditory Conceptualization Test, Third Edition [LAC-3] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5:0 – 18.11 years	20 – 30 minutes	Must have intact phoneme awareness, be competent in the administration of educational or psychological tests, and be knowledgeable of the procedures for administering and scoring the LAC-3.	$196

Available Languages: English, Spanish

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child.

Scores Available: Raw scores, standard scores, descriptive ratings and percentages based on standard scores, percentile ranks, age equivalents, and grade equivalents.

Subscales: Isolating Phoneme Patterns, Tracking Phonemes, Counting Syllables, Tracking Syllables, Tracking Syllables and Phonemes.

Norming sample: 1,003 individuals in 8 states, similar to the U.S. population with regard to geographic area, age, gender, ethnicity, family income, parental education, and disability status (2001 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: May be available in the Technical Manual.

Note: Assesses an individual's ability to perceive and conceptualize speech sounds in isolation and within and across syllables.

MacArthur-Bates Communicative Development Inventories, Second Edition [CDI] (2006)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:8 – 3:1 years	20 – 40 minutes	Parent can complete with no training.	$100

Available Languages: English, Spanish, plus numerous other languages, some with norms

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624
(800) 638-3775; www.brookespublishing.com

Administration: Parent completes one of three questionnaires about the child. The Words and Gestures form is for children from 8 to 18 months. The Words and Sentences form is for children 16 to 30 months. The CDI-III is for children 30 to 37 months.

Scores Available: May be available in Technical Manual.

Subscales: May be available in Technical Manual.

Norming Sample: May be available in Technical Manual.

Reliability: May be available in Technical Manual.

Concurrent Validity: May be available in Technical Manual.

Merrill-Palmer-Revised Scales of Development [M-P-R] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 6:6 years	45 minutes	Degree in psychology, counseling, or a closely related field, plus coursework in testing.	$1,079

Available Languages: English, Spanish

Source: Psychological Assessment Resources; 16204 North Florida Avenue, Lutz FL 33549 (800) 331-8378; www.parinc.com

Administration: Direct assessment of individual child by examiner, plus parent report.

Scores Available: Standard scores, percentiles, age equivalents, growth scores, and growth score profile. Derived scores are recorded on the Summary Report and may be plotted to indicate patterns of strengths and weaknesses.

Subscales: Expressive Language, Social-Emotional Development, Social-Emotional Temperament, Self-Help-Adaptive, Cognitive Battery, and Gross Motor Skills.

Norming Sample: 1,400 children, representative of U.S. population with regard to gender, ethnicity, socioeconomic level, and geographic region (2000 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher) (small sample size)

Metropolitan Readiness Test [MRT6], Sixth Edition (1994)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
4 – 7 years	85 – 100 minutes (4 sessions)	A master's degree in psychology, education, occupational therapy, speech-language pathology, social work, or related field and formal training assessment.	$163 (Level 1) $225 (Level 2)

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259 (800) 211-8378; harcourtassessment.com

Administration: *Level 1* (pre-kindergarten and beginning K): direct assessment of individual child; *Level 2* (middle K to beginning 1st grade): group administered test of individual child skills.

Scores Available: Percentile ranks, stanines, normal curve equivalents, scaled scores, and standard scores . Content-Referenced Performance Ratings (proficiency, acquisition, and needed instruction) based on authors' judgment and Stanines Classifications (above average, average, or below average) based on normative data are also provided.

Subscales: Beginning Reading Skill Area (Visual Discrimination, Beginning Consonants, Sound-Letter Correspondence, Aural Cloze with Letter), Story Comprehension, Quantitative Concepts and Reasoning, Prereading Composite

Norming Sample: May be available in the Technical Manual.

Reliability: Level 1: High (.80 or higher); Level 2: High (.80 or higher)

Concurrent Validity: Level 1: Not reported, Level 2: Adequate (.50 to .69)

Mullen Scales of Early Learning [MSEL] (1995)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 5:8 years	1 year: 15 minutes 3 years: 25-35 minutes 5 years: 40-60 minutes	Administrators should have graduate training and experience in infant assessment. Training video available.	$728

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child.

Scores Available: T score; confidence intervals, percentile rank, age equivalent, developmental stage, descriptive category, profile analysis; an early learning composite can be derived

Subscales: Gross Motor, Visual Reception, Fine Motor, Receptive Language, Expressive Language.

Norming Sample: Sample included 1,849 children, representative of the U.S. Population with regard to race, socioeconomic status, region, and community size (1990 U.S. Census). Excluded children with special needs

Reliability: High (.80 or higher

Concurrent Validity: High (.70 or higher)

OWLS: Listening Comprehension [LC] Scale, Oral Expression [OE] Scale (1995)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 21:11 years	LCS: 5-15 minutes OES: 10-25 minutes	School psychologists, speech pathologists, and educational diagnosticians.	$581

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child. LC consists of 111 items; OE consists of 96 items.

Scores Available: Age-based standard scores, percentiles, NCEs, stanines, test-age equivalents

Subscales: Listening Comprehension (LC), Oral Expression (OE), Oral Composite

Norming Sample: 1,795 individuals 74 different sites, representative of the U.S. population with regard to gender, region, race/ethnicity, and socioeconomic status (1991 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Note: Co-normed with OWLS Written Expression (WE) scales

OWLS: Written Expression Scale [WE] Scale (1996)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5:0 – 21:11 years	15 – 25 minutes	School psychologists, speech pathologists, educational diagnosticians, etc.	$149

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Can be administered individually or in small groups; about 15 items.

Scores Available: Age-based and grade-based standard scores, grade and age equivalents, percentiles, normal curve equivalents (NCEs), and stanines

Subscales: Written Expression, Language Composite (if LC and OE are also administered)

Norming Sample: 1,795 individuals 74 different sites, representative of the U.S. population with regard to gender, region, race/ethnicity, and socioeconomic status (1991 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Note: Co-normed with OWLS Listening Comprehension (LC) and Oral Expression (OE) scales.

Peabody Developmental Motor Scales, Second Edition [PDMS-2] (2000)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 6:0 years	45 – 60 minutes	Occupational therapists, physical therapists, diagnosticians, early interventionist, adapted physical education teachers, psychologists, etc.	$365

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202 www.proedinc.com

Administration: Direct assessment of individual child.

Scores Available: Raw scores, percentiles, age equivalents, and standard scores. Each item on which the child met the criterion for mastery is marked on the Profile of Item Mastery section, enabling examiner to compare the child's performance on the items he or she has mastered with that of the normative sample.

Subscales: Reflexes (8 items), Stationary (30 items), Locomotion (89 items), Object Manipulation (24 items), Grasping (26 items), Visual-Motor Integration (72 items); plus Fine Motor, Gross Motor and Total Motor Quotients.

Norming Sample: 2,003 children in 46 states; representative of the U.S. population with regard to geography, gender, and race (1997 Census).

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69) (sample included only 2-year olds)

Note: The Peabody Motor Activities Program (P-MAP) (sold separately or with the PDMS) is the instruction/treatment program for the PDMS-2. It contains units organized developmentally by skill area. After a child's motor skills have been assessed and the examiner has completed all sections of the Profile/ Summary Form, the examiner selects units from the P-MAP to use to facilitate the child's development in specific skill areas.

Peabody Individual Achievement Test – Revised, Normative Update [PIAT-R/NU] (1998)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5:0 – 22:11 years (grades K-12)	1 hour	Master's degree in psychology, education, occupational therapy, speech/language pathology or related field, plus formal training in assessment.	$415

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child.

Scores Available: Standard scores, percentiles, stanines, and NCEs by grade (fall, winter, and spring norms) or by age; age equivalents and grade equivalents. Record forms include a Developmental Score Profile for profiling age and grade equivalents and a Standard Score Profile for profiling for age- or grade-based standard scores.

Subscales: General Information, Reading Recognition, Reading Comprehension, Mathematics, Spelling, Written Expression. Also, can create Total Reading, Written Language, and Total Test Scores.

Norming Sample: Varied by subtest from low of 1,285 for Written Expression to high of 2,809 for Mathematics Application. Generally representative of the U.S. population with regard to gender, parent[s] education, and race/ethnicity (1994 U.S. Census). Gifted students and those with special needs were included.

Reliability: High (.80 or higher) (based on data collected for the 1989 revision)

Concurrent Validity: Adequate (.50 to .69) (based on data collected for the 1989 revision from 5 and 6 year olds)

Note: This tool's content was last updated in 1989, but its norms were updated in 1998.

Peabody Picture Vocabulary Test, Fourth Edition [PPVT-4] (2007)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:6 – 90 years	10 – 15 minutes	Bachelor's degree, including coursework in measurement, testing, and formal training speech and language.	$379

Available Languages: English, Spanish (Test de Vocabulario en Imagenes Peabody [TVIP], last updated in 1986)

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child; 2 forms, 228 items per form, with basal and ceiling rules.

Scores Available: Age- and grade-based standard scores, Growth Scale Values (GSVs), percentiles, stanines, normal curve equivalents (NCEs), age and grade equivalents

Subscales: None (test of receptive vocabulary only)

Norming Sample: 3,450 individuals, closely matched to the U.S. population with regard to age, sex, race/ethnicity, region, parent education (2004 U.S. Census). The sample of children ages 2 to 18 included representative proportions of children with special needs.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Note: Co-normed with the EVT-2.

Phonological Awareness Test 2 [PAT-2] (2007)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5:0 – 9:11 years	40 minutes	Administrator needs to have coursework and experience with standardized testing and phonolgical awareness.	$170

Available Languages: English

Source: LinguiSystems, 3100 4th Avenue, East Moine, IL 61244
(800) 776-4338; www.linguisystems.com

Administration: Direct assessment of individual child.

Scores Available: Raw scores are converted to age equivalencies, percentile ranks, and standard scores.

Subscales: Rhyming, segmentation, isolation, deletion, substitution, blending, graphemes, decoding, invented spelling

Norming Sample: 1,582 reflecting the national school population with regard to race, gender, age, and educational placement (2004 Census).

Reliability: High (.80 or higher)

Concurrent Validity: Not reported.

Note: Manual provides information on using test scores for instructional planning.

Pictorial Test of Intelligence, Second Edition [PTI-2] (2001)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 8 years	15 – 30 minutes	Examiners should have formal training in assessment. Specific qualified professionals generally include psychologists, psychological associates, educational diagnosticians, SLPs, teachers, and counselors.	$143

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child; 98 items.

Scores Available: Standard scores, age equivalents, and percentile ranks, Pictorial Intelligence Quotient

Subscales: Verbal abstractions, form discrimination, and quantitative concepts

Norming Sample: 970 children in 15 states, intended to be representative of the U.S. population in terms of region, sex, race, rural-urban, parent income level, parent education, and disability status; however rural children were somewhat underrepresented (1997 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69) (n = 15 nondisabled 3-year-olds)

Note: Designed to measure general intelligence. All items use a multiple-choice format, allowing examinees to indicate their choice via pointing or eye gaze; no verbal expressive skill required.

Preschool and Kindergarten Behavior Scales, Second Edition [PKBS-2] (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3 – 6 years	12 minutes	Completed by parent or caregiver. Basic understanding of the principles of education and psychological testing needed for interpretation.	$110

Available Languages: English, Spanish

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Parent, teacher or other caregiver responds to 76 items about the child, using 4-point scales (0 = *never*, 1 = *rarely*, 2 = *sometimes*, and 3 = *often*).

Scores Available: Raw scores are converted to standard scores, percentile ranks, and risk levels.

Subscales: Social Skills (34 items) and Problem Behaviors (42 items). Social Skills section is further broken down into 3 subscales: Social Cooperation, Social Interaction, and Social Independence. Problem Behaviors section is broken into two subscales: Externalizing Problems and Internalizing Problems. In addition, 5 supplementary problem behavior subscales are available for optional use, including Self-Control-Explosive, Attention Problems-Overactive, Antisocial-Aggressive, Social Withdrawal, Anxiety-Somatic Problems).

Norming Sample: 3,317 children; similar to the U.S. population with regard to ethnicity, socioeconomic status, and special education classification (2000 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: May be available in Technical Manual.

Preschool Language Assessment Instrument, Second Edition [PLAI-2] (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3:0 – 5:11 years	3:0 – 5:11 years	Formal training in assessment and familiarity with the specific area of the test content. Qualified professionals generally include psychologists, educational diagnosticians, SLP's, teachers, and counselors.	$197

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child.

Scores Available: Standard scores, percentile ranks, and age equivalencies

Subscales: *Norm referenced*: Matching, Analysis, Reordering, Reasoning, Receptive Mode, Expressive Mode, plus a Discourse Ability Score gives an overall estimate of performance. *Nonstandardized*: Adequacy of Response, Interfering Behaviors.

Norming Sample: 463 children in 16 states, similar to U.S. population with regard to socioeconomic status, gender, and disability (1999 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69) ($n = 38$, in 3 Northeastern states)

Note: Designed to assess the child's ability to meet the demands of classroom discourse.

Preschool Language Scale, Fourth Edition [PLS-4] (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 6:11 years	20 – 45 minutes	Master's degree in psychology, education, occupational therapy, speech/language pathology, or related field and formal training assessment.	$235

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child, plus optional caregiver questionnaire.

Scores Available: Standard scores, percentile ranks, and age equivalents are available for birth to 11 months (3-month intervals) and 1 year through 6 years, 11 months (6-month intervals). The *Language Sample Checklist* provides mean length of utterance (MLU) and summary profile. The *Articulation Screener* provides age-appropriate cut-points that help a clinician determine if further articulation testing is advisable.

Subscales: Two core subscales: Auditory Comprehension, Expressive Communication. Three supplemental assessments: Language Sample Checklist, Articulation Screener, and Caregiver Questionnaire.

Norming Sample: 2,400 children at 357 sites in 48 states, representative of the U.S. population with regard to race/ethnicity, parental education, and region (2000 Census). Included children with disabilities.

Reliability: High (.80 or higher)

Concurrent Validity: Not reported (Technical Manual includes evidence of other types of validity.)

Process Assessment of the Learner, Second Edition: Diagnostic Assessment for Math [PAL-II Math] (2007)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
Kindergarten – 6th grade	30 – 60 minutes	Master's degree in psychology, education, occupational therapy, speech/language pathology, or related field and formal training assessment.	$299

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child.

Scores Available: Scale scores, base rates, cumulative percentages.

Subscales: May be available in the Technical Manual.

Norming sample: 700 children, representative of the U.S. population with regard to age, race/ethnicity, sex, parent education, and region (2003 Census); children with diagnosed delays or who had been referred due to a suspected delay were not included.

Reliability: High (.80 or higher)

Concurrent Validity: May be available in the Technical Manual.

Note: Designed to measure the development of cognitive processes that are critical to learning math skills and actual math performance. Co-normed with the PAL-II Reading and Writing.

Process Assessment of the Learner, Second Edition: Diagnostic Assessment for Reading and Writing [PAL-II Reading and Writing] (2007)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
Kindergarten – 6th grade	30 – 60 minutes	Master's degree in psychology, education, occupational therapy, speech/language pathology, or related field and formal training assessment.	$399

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child.

Scores Available: Scale scores, base rates, cumulative percentages, fluency scores, RAN/RAS change scores.

Subscales: Full list available in Technical Manual. Examples include: Alphabet Writing, Copying, Compositional Fluency; Expository Note-Taking, Expository Report Writing, Verbal Working Memory, Written Sentences, Pseudoword Decoding, Rapid Automatized Naming-Letters, Rhyming.

Norming sample: 700 children, representative of the U.S. population with regard to age, race/ethnicity, sex, parent education, and region (2003 Census); children with diagnosed delays or who had been referred due to a suspected delay were not included.

Reliability: High (.80 or higher)

Concurrent Validity: May be available in Technical Manual.

Note: Developed to facilitate the creation of assessment driven interventions in the areas of reading and writing. Conormed with the PAL-II Math.

Ready to Learn: A Dyslexia Screener (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3:6 – 6:5 years	35 minutes	School psychologists, qualified teachers, and other professionals with training in formal assessment.	$250

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child.

Scores Available: Raw scores are converted into Risk Index Scores. Subtest scores are coded by risk category: Green (low risk), Yellow (moderate risk), and Red (high risk). Record form provides a graph for plotting the raw scores and manual provides scaled growth scores for assessing growth over time (minimum of 6-month interval).

Subscales: Includes Rapid Naming, Phonological Discrimination, First Letter Sounds, Rhyming, Sound Order, Bead Threading, Shape and Letter Copying, Corsi Frog (working memory), Balance, Postural Stability, Digit Span, Repetition, Teddy and Form Matching, Receptive Vocabulary, Digit Naming, and Letter Naming. Also provides a Risk Index.

Norming sample: 510 children, representative of U.S. population with regard to sex, parent education, race/ethnicity and region (2000 Census). About 4% of the children were diagnosed with language impairment, developmental delay, or risk for delay.

Reliability: Adequate (.65 to .79)

Concurrent Validity: Adequate (.50 to .69)

Receptive Expressive Emergent Language Scale, Third Edition [REEL–3] (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 3 years	20 minutes	Professional with training in interpretation of speech/language assessments.	$100

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Caregiver interview.

Scores Available: Standard scores, percentile ranks, and age equivalents

Subscales: Receptive Language, Expressive Language, Inventory of Vocabulary Words.

Norming Sample: Sample included 1,112 individuals in 32 states, matched to the U.S. Population on the basis of age, gender, race, ethnicity, geographic location (2000 Census); 2% had language disabilities; 7% had other disabilities.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Receptive One-Word Picture Vocabulary Test [ROWPVT] (2000)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2 – 18:11 years	15 – 20 minutes	Can be administered by non-professional with training. Requires graduate training and experience for interpretation.	$159

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child; includes 170 test items with basal and ceiling rules.

Scores Available: Raw scores can be converted into a standard score, percentile rank, age equivalent, normal curve equivalent (NCEs), scaled scores, T-scores, and stanines.

Subscales: None

Norming Sample: 2,327 children closely matched to the U.S. population with regard to region, race/ethnicity, gender, parent education, urban versus rural residence, and disability status (1998 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Note: Measures receptive vocabulary only. See EOWPVT for a measure of expressive vocabulary. Spanish version is for ages 4 to 12:11 years.

Scales of Independent Behavior – Revised [SIB-R] (1996)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:3 – 80 years	full: 45 – 60 min short/early: 15 – 20 min.	Administration requires no specialized training, but a high level of training is required for proper interpretation.	$248

Available Languages: English

Source: Riverside Publishing, 3800 Golf Road, Suite 100, Rolling Meadows, IL 60008
(800) 323-9540; www.riverpub.com

Administration: Structured interview or checklist. Three forms: *full scale* (259 items), *short form* (40 items), and *early development form* (40 items, 0:3 to 6:11 years).

Scores Available: Age equivalent scores, cluster W scores, standard scores, percentiles, Relative Mastery Indexes, Adaptive Behavior Skill Levels, Support Score, Instructional and Developmental Ranges, Functional Limitations Index, four Maladaptive Behavior Indexes

Subscales: See Technical Manual for full list. Examples include *Adaptive Behavior* (Gross Motor, Fine Motor, Social Interaction, Language Comprehension, Language Expression, Eating and Meal Preparation, Toileting); *Problem Behavior* (Hurtful to Self, Unusual/Repetitive Habits, Withdrawal/Inattentive, Disruptive).

Norming Sample: 2,182 individuals in 15 states and more than 60 communities, similar to the U.S. population with regard to gender and race (1990 Census).

Reliability: High (.80 or higher)

Concurrent Validity: Low (below .50) (somewhat higher for individuals with disabilities than those without disabilities).

Sensory Profile (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3:0 – 10:0 years	30 minutes	No training required for administration. Interpretation requires understanding of sensory processing.	$162

Available Languages: English, Spanish

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Caregivers complete a 125-item questionnaire reporting the frequency with which infants respond to various sensory experiences (5-point scale: always to never).

Scores Available: Cut-points for each section classify child's sensory processing abilities into: Typical Performance, Probable Difference, or Definite Difference.

Subscales: See Technical Manual for full list. Examples include *Sensory Processing* (Auditory, Visual, Multisensory, Oral Sensory Processing); *Modulation* (Modulation Related to Body Position and Movement, Modulation of Movement Affecting Activity Level); *Behavioral and Emotional Responses* (Emotional-Social Responses, Thresholds for Response).

Norming Sample: Over 1,000 children without disabilities for creation of the scoring structure. Technical Manual may include demographic information about the sample. Children with disabilities were included in other aspects of the standardization process.

Reliability: Adequate (.65 to .79)

Concurrent Validity: May be available in Technical Manual.

Notes: Designed to measure sensory processing abilities and profile the effects of sensory processing on functional performance in the child's daily life.

Social Competence and Behavior Evaluation [SCBE– Preschool Edition] (1995)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:6 – 6:4 years	15 minutes	Completed by the child's teacher with no special training.	$93

Available Languages: English, French

Source: Western Psychological Services, 12031 Wilshire Boulevard, Los Angeles, CA 90025 (800) 642-8857; www.wpspublish.com

Administration: Teacher response to 80 items (10 per subscale) of questions about the child using a six- point scale (*sometimes* to *always*).

Scores Available: The raw scores are converted to T scores and percentile ranks. Cut-points are provided for good adjustment and adjustment problems.

Subscales: 8 subscales: Depressive-Joyful, Anxious-Secure, Angry-Tolerant, Isolated-Integrated, Aggressive-Calm, Egotistical-Prosocial, Oppositional-Cooperative, Dependent-Autonomous. 4 Summary Scales: Social Competence, Externalizing Problems, Internalizing Problems, General Adaptation.

Norming Sample: 1,263 children at six sites in Indiana and Colorado. African American children and children with low SES parents were overrepresented as compared with the 1991 Census.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69) in a sample of French-speaking Canadian children.

Stanford Achievement Test, Tenth Edition [SAT 10] (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
Kindergarten – 12th grade	1.75 to 5.5 hours	No special training.	$50 per level

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259 (800) 211-8378; harcourtassessment.com

Administration: Group administration; 13 different test levels

Scores Available: Scaled Scores, National and Local Percentile Ranks and Stanines, Grade Equivalents, and Normal Curve Equivalents

Subscales: All of the levels of the Stanford 10 have subtests in the disciplines of Reading, Mathematics, and Language. Specific subtests vary by level and include Sounds and Letters, Word Study Skills, Word Reading, Sentence Reading, Reading Comprehension, Mathematics, Mathematics Problem-Solving, Mathematics Procedures, Language, Spelling, Listening to Words and Stories, Listening, Environment, Science, Social Science. Total Reading and Total Mathematics scores are available.

Norming Sample: The spring standardization involved 250,000 students, and the fall standardization involved 110,000, selected to match the U.S. population in terms of region, SES, urbanicity, and ethnicity (2000 Census). Special education students who would routinely be tested were included in the standardization samples.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Note: The first two levels of the SAT 10 are called the Stanford Early School Achievement Test.

Stanford-Binet Intelligence Scales for Early Childhood, Fifth Edition [Early SB5] (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:0 – 7:3 years	30 – 50 minutes	High level of training in child development, and assessment methods.	$355

Available Languages: English

Source: Riverside Publishing, 3800 Golf Road, Suite 100, Rolling Meadows, IL 60008
(800) 323-9540; www.riverpub.com

Administration: Direct assessments of individual child.

Scores Available: Standard scores, confidence intervals, percentile ranks, change-sensitive scores (CSSs), and age-equivalents. Trained examiners can also conduct qualitative analyses on domain and subtest score differences. Four types of IQ scores are available: Abbreviated Battery IQ, Nonverbal IQ, Verbal IQ, and Full Scale IQ.

Subscales: 10 subscales assessing *Nonverbal* and *Verbal*: Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Processing, Working Memory

Norming Sample: 1,660 children, largely similar to the U.S. population for age, gender, ethnicity, locale, and SES (2001 Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Note: This is a specialized version of the SB5. The SB5 covers ages 2 to 80+ years. The two tests were normed together.

Temperament and Atypical Behavior Scale [TABS] Screener and Assessment Tool (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0:11 – 5:11 years	15 minutes	Completed by a parent or professional familiar with child. Professional required for interpretation.	$85

Available Languages: English

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624
(800) 638-3775; www.brookespublishing.com

Administration: Parent or a professional familiar with the child completes the 15 item (yes/no) TABS Screener. If the Screener identifies an area of concern, the parent or professional completes the 55-item (yes/no/need help) TABS Assessment Tool.

Scores Available: Screener provides cut-points for further evaluation.
Assessment Tool provides standard scores, percentiles, and standard deviations.

Subscales: Detached, Hypersensitive-Active, Underreative, Dysregulated

Norming Sample: 621 typically developing children in 33 states and 3 Canadian provinces. Additional information about the norming sample was not collected because the authors believe that the atypical behaviors being assessed are atypical in any social class, geographic area, and cultural group.

Reliability: High (.80 or higher)

Concurrent Validity: Not reported; there are no other measures designed to address these behaviors with this age group. Manual provides evidence of other types of validity.

Note: Designed to detect emerging problems in temperament and self-regulatory behavior in infants, toddlers, and preschoolers. Validation studies included children with disabilities.

Test for Auditory Comprehension of Language, Third Edition [TACL-3] (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3:0 – 9:11 years	15 – 25 minutes	Requires a trained speech/language clinician for administration.	$275

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child; 142 items, with ceiling rules.

Scores Available: Standard scores, percentile ranks, and age equivalents

Subscales: Vocabulary, Grammatical Morphemes, Elaborated Phrases

Norming Sample: 1,102 children, representative of the U.S. population with regard to SES, ethnicity, gender, and disability (2000 estimates U.S. Census). Children with speech-language disorders and children with learning disabilities were included in the sample.

Reliability: High (.80 or higher)

Concurrent Validity: May be available in Technical Manual.

Test of Early Language Development, Third Edition [TELD-3] (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:0 – 7:11 years	15 – 45 minutes	Requires a trained speech/language clinician for administration.	$293

Available Languages: English

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: Direct assessment of individual child; 76 items.

Scores Available: Standard scores, percentiles, age-equivalent scores

Subscales: Receptive Language and Expressive Language

Norming Sample: 2,217 children from 35 states, representative of the U.S. population in terms of geography, gender, race, ethnicity, education and socioeconomic status (1997 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Test of Early Mathematics Ability, Third Edition [TEMA-3] (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3:0 – 8:11 years	40 minutes	Bachelors in psychology, counseling, or related field, plus coursework in testing	$265

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child; 72 items with basal and ceiling rules.

Scores Available: Standard scores, percentile ranks, and age and grade equivalents

Subscales: None

Norming Sample: 1,219 children in 16 states, representative of the U.S. population in terms of geographic region, gender, and ethnicity (2001 Census).

Reliability: High (.80 of higher)

Concurrent Validity: Adequate (.50 to .69) (n = 43 to 62)

Note: Test skills in the following domains: numbering skills, number-comparison facility, numeral literacy, mastery of number facts, calculation skills, and understanding of concepts.

Test of Early Reading Ability, Third Edition [TERA-3] (2001)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3:6 – 8:6 years	30 minutes	Bachelors in psychology, counseling, or related field, plus coursework in testing.	$265

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child; 80 items with basal and ceiling rules.

Scores Available: Standard scores, percentiles, age and grade equivalents

Subscales: Alphabet, Conventions, Meaning

Norming Sample: 875 children from 22, representative of the U.S. population with regard to SES, gender and disability (2000 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69) (n = 64 to 70 2nd and 3rd graders in South Dakota and Texas)

Note: Assesses mastery of early developing reading skills (as opposed to readiness for reading).

Test of Language Development – Primary, Fourth Edition [TOLD-P:4] (2008)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
4:0 – 8:11 years	1 hour	Formal training in assessment and familiarity with the specific area of the test content. Qualified professionals generally include psychologists, psychological associates, educational diagnosticians, SLPs, teachers, and counselors.	$299

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child.

Scores Available: May be available in Technical Manual.

Subscales: 9 subtests**:** Picture Vocabulary, Relational Vocabulary, Oral Vocabulary, Syntactic Understanding, Sentence Imitation, Morphological Completion, Word Discrimination, Word Analysis, Word Articulation. Subtests can be combined to create 3 composites: semantics and grammar; listening, organizing, and speaking; and overall language ability.

Norming Sample: May be available in Technical Manual.

Reliability: May be available in Technical Manual.

Concurrent Validity: May be available in Technical Manual.

Test of Phonological Awareness, Second Edition PLUS [TOPA-2+] (2004)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5 – 8 years	Kindergarten: 30 – 45 minutes Early elementary: 15 – 30 minutes	No special training required; must carefully read manual.	$210

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Group or individual administration; separate versions for Kindergarten (35 items) and early elementary (20 items)

Scores Available: Standard scores, NCEs, percentile ranks

Subscales: Kindergarten: Initial Sound-Same, Initial Sound-Different, Letter-Sounds;
Early Elementary: Ending Sound-Same, Ending Sound Different

Norming Sample: 1,035 students for the Kindergarten version and 1,050 students for the Early Elementary version. Both samples were in 26 states and were representative of the U.S. population with regard to geographic region, gender, race, ethnicity, family income, and educational attainment (2001 U.S. Census).

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Test of Preschool Early Literacy (TOPEL) (2007)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
3:0 – 5:11 years	25 – 30 minutes	Early childhood educators, special educators, psychologists, diagnosticians, and other professionals.	$207

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child; 98 items

Scores Available: Raw scores, standard scores, and percentiles

Subscales: Print Knowledge, Definitional Vocabulary, Phonological Awareness

Norming Sample: 842 children in 12 states, closely matched to the U.S. population with regard to geographic area, gender, race/ethnicity, parental education (U.S. Census 2001). Norming sample had slightly lower family incomes and slightly fewer children with disabilities than the U.S. population.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Note: Prior to publication, an earlier version of this tool was called the "Preschool CTOPP."

Test of Word Reading Efficiency [TOWRE] (1999)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
6:0 – 24:11 years	5 – 10 minutes	Can be administered by anyone who can read and understand the test manual and has knowledge of standard assessment procedures.	$178

Available Languages: English

Source: PRO-ED, Inc., 8700 Shoal Creek Boulevard, Austin, TX 78757-6897
(800) 897-3202; www.proedinc.com

Administration: Direct assessment of individual child

Scores Available: Percentiles, standard scores, and age and grade equivalents

Subscales: Sight word efficiency and phonetic decoding efficiency

Norming Sample: 1,500 individuals from 30 states, reflecting U.S. population demographics (1997 Census). The Technical Manual may include additional details about the norming sample.

Reliability: High (.80 or higher)

Concurrent Validity: May be available in the Technical Manual.

Toddler and Infant Motor Evaluation [TIME] (1994)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 3:6 years		Occupational or physical therapists; or appropriately trained educators, medical or mental health professionals.	$417

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Examiner observes a parent or caregiver interacting/playing with the child and prompts the parent to elicit specific motor abilities.

Scores Available: Standard Scores, Percentile Ranks, cut-points for "motor impairment"

Subscales: 5 primary subtests: Mobility, Stability, Motor Organization, Functional Performance, Social-Emotional Abilities; 3 Clinical Subtests: Quality Rating, Component Analysis Rating, Atypical Positions.

Norming sample: 731 children in 10 states; representative of the U.S. population in terms of race/ethnicity, gender, socioeconomic status, and age (1990 U.S. Census). Sample included children with and without motor delays.

Reliability: High (.80 or higher) ($n = 33$)

Concurrent Validity: Not reported. See Technical Manual for other types of validity.

Note: Designed to evaluate the overall quality of infant and toddler movement, rather than isolated skills, in children who have atypical motor development.

Vineland Adaptive Behavior Scales, Second Edition [Vineland-II] (2005)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 90 years	20 – 60 minutes	Graduate-level training in psychology or other social work, as well as experience in individual assessment and interpretation.	$230

Available Languages: English, Spanish

Source: Pearson Assessments, 5601 Green Valley Drive, Bloomington, MN 55437
(877) 242-6767; www.pearsonassessments.com

Administration: *Survey Interview:* administered by a professional to a parent or caregiver using a semi-structured interview format. *Parent/Caregiver Rating* uses a rating scale format to cover the same content as the Survey Interview. *Expanded Interview* uses the semi-structured interview to yield a more comprehensive assessment of adaptive behavior. *Teacher Rating*: questionnaire format, completed by teacher or child care provider.

Scores Available: Standard Scores, Percentile Ranks, Adaptive Levels, Age Equivalents, Maladaptive Levels

Subscales: Communication (receptive, expressive, written); Daily Living Skills (personal, domestic, community); Socialization (interpersonal relationships, play and leisure time, coping skills); Motor (fine, gross); Maladaptive Behavior (internalizing, externalizing, other)

Norming Sample: 3,687 individuals in 242 sites in 44 states, representative of the U.S. population with regard to race/ethnicity, socioeconomic status, and region (U.S. Census, 2001). Children with educational, psychological, or physical classifications were included in numbers proportional to their numbers in the general school-age population.

Reliability: High (.80 or higher)

Concurrent Validity: Adequate (.50 to .69)

Note: Designed to assess personal and social skills among individuals with special needs.

Wechsler Individual Achievement Test, Second Edition [WIAT-II] (2001)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
4:0 – 85 years	45 minutes – 2 hours, depending on age	Master's degree in psychology, education, occupational therapy, speech-language pathology, social work, or closely related field, plus training in assessment.	$417

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child.

Scores Available: Standard Score, Percentile Ranks, Stanines, NCEs; Age and Grade Equivalents, quartile scores, and decile score; detailed skills analysis specifying an individual's strengths and appropriate intervention targets. Norms provided for fall, winter, and spring.

Subscales: Oral Language, Listening Comprehension, Written Expression, Spelling, Pseudoword Decoding, Word Reading, Reading Comprehension, Numerical Operations, Mathematics Reasoning

Norming Sample: 2,950 school-aged children, representative of the U.S. population with regard to gender, race/ethnicity, region, and parental education (1998 U.S. Census). Students with disabilities were included in the standardization sample in proportion to their representation in public school programs.

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Wechsler Preschool and Primary Scale of Intelligence, Third Edition [WPPSI-III] (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:6 – 7:3 years	30 – 60 minutes	A doctorate degree in psychology, education, or closely related field with formal training in testing or a state licensure or certification.	$888

Available Languages: English

Source: Harcourt Assessment, 19500 Bulverde Road, San Antonio, TX 78259
(800) 211-8378; harcourtassessment.com

Administration: Direct assessment of individual child

Scores Available: Scaled scores by age, IQ

Subscales: *Verbal IQ* (Information, Receptive Vocabulary, Vocabulary, Word Reasoning, Picture Naming, Comprehension, Similarities); *Performance IQ* (Block Design, Object Assembly, Matrix Reasoning, Picture Concepts, Object Assembly, Picture Completion); *Full Scale IQ* is a composite of Verbal and Performance IQ. Additional supplemental scales: General Language Quotient, Processing Speed.

Norming Sample: 1,700 children, representative of the U.S. population with regard to parental education, race/ethnicity and region (2000 Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Woodcock-Johnson III Normative Update Complete [WJ III NU] (2006)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
20 – 90+ years	Cognitive: 35 – 45 minutes Achievement: 55 – 65 minutes	Requires graduate training in testing, plus extensive training with this battery.	$1,553 for complete battery

Available Languages: English, Spanish (Batería III Woodcock-Muñoz)

Source: Riverside Publishing, 3800 Golf Road, Suite 100, Rolling Meadows, IL 60008 (800) 323-9540; www.riverpub.com

Administration: Direct assessment of individual child. Two separate batteries: Cognitive (COG) and Achievement (ACH).

Scores Available: Standard scores, percentile ranks, W scores, T scores, NCEs, Z scores, stanines, grade and age equivalents, relative proficiency indexes, cognitive-academic language proficiency levels, instructional zones, developmental zones, relative performance index

Subscales: The *WJ III COG* consists of 20 tests tapping seven cognitive factors: Comprehension-Knowledge, Long-Term Retrieval, Visual-Spatial Thinking, Auditory Processing, Fluid Reasoning, Processing Speed, and Short-Term Memory. The *WJ III ACH* contains 22 tests tapping five curricular areas: Reading, Oral Language, Mathematics, Written Language, and Academic Knowledge (e.g., science, social studies).

Norming Sample: Over 8,782 individuals in more than 100 geographically diverse communities; representative of the U.S. population on 10 community and individual variables and 13 socio-economic status variables (2005 Census).

Reliability: High (.80 or higher)

Concurrent Validity: High (.70 or higher)

Note: Two distinct batteries (Cognitive and Achievement), normed together. Can be used to create ability and achievement discrepancy scores.

Children learn by exploring, thinking about, and inquiring about all sorts of phenomenal material. These experiences help children investigate "big ideas."

Collection of Program Evaluation Tools

Page	Name
161	Classroom Assessment Scoring System (CLASS)
162	Early Childhood Environment Rating Scale – Revised (ECERS-R)
162	Early Language and Literacy Classroom Observation Scale (ELLCO)
163	Family Child Care Environment Rating Scale – Revised
163	Infant Toddler Environment Rating Scale – Revised (ITERS)
164	School-Age Care Environment Rating Scale (SACERS)

Classroom Assessment Scoring System [CLASS] (2006)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
PreK – 3rd grade	2 hours	Observer must attend a training session and pass a reliability test.	$600 per person for training $20 manual

Available Languages: English

Source: Center for Advanced Study of Teaching and Learning, University of Virginia, 350 Old Ivy Way, Suite 100, Charlottesville, VA 22903
(866) 301-8278; classobservation.com

Administration: Trained observer scores classroom on 11 dimensions, using 7-point scales.

Scores Available: Average score (1 -7) on each subscale

Subscales: Emotional Support, Classroom Organization, and Instructional Support

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: Low (below .50) Significant correlations were found with other measures of classroom quality, but they were generally low, possibly because this tool measures different aspects of the classroom than other quality measures.

Note: There are separate pre-kindergarten and K – 3rd grade versions of this tool, as well as versions for older grades.

Early Childhood Environment Rating Scale – Revised Edition [ECERS-R] (1998)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
2:6 – 5 years	2.5 – 5 hours	Administrator must read and practice the scale & have knowledge of child development.	$18

Available Languages: English, Spanish, French, German, Hungarian, Norwegian

Source: Teachers College Press, Teachers College, Columbia University, New York, NY 10027 (800) 575-6566; www.teacherscollegepress.com

Administration: Observer watches early childhood classroom and asks questions of the teacher. Environment is rated on 46 7-point scales (*inadequate* to *excellent*).

Scores Available: Total score, plus subscale scores, each ranging from 1 to 7

Subscales: Space and Furnishings, Personal Care Routines, Language-Reasoning, Activities, Interaction, Program Structure, Parents and Staff

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: Not reported

Early Language and Literacy Classroom Observation Scale [ELLCO] (2002)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
PreK – 3rd grade	60 – 90 minutes	Can be administered by teachers, principals, administrators, supervisors, program directors, or researchers.	$50

Available Languages: English

Source: Paul H. Brookes Publishing Co., Inc., P.O. Box 10624, Baltimore, MD 21285-0624 (800) 638-3775; www.brookespublishing.com

Administration: Observational tool for researchers, supervisors, program directors, principals, administrators, and/or teachers to use in measuring classroom-level language and literacy environment. Has three components: 1) literacy environment checklist, 2) classroom observations and teacher interview, and 3) literacy activities rating scale.

Scores Available: Cut-points provided for *exemplary*, *basic*, and *deficient*.

Subscales: 1) Literacy environment checklist provides a total score, books subscale, and writing subscale; 2)classroom observation and teacher interview provides a total score, general classroom environment subtotal, and language, literacy and curriculum subtotal; 3)literacy activities rating scale provides a total score, full-group book reading subtotal, and writing subtotal.

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: Not reported, but this tool is highly associated with vocabulary and early literacy scores, accounting for 67-80% of between-classroom variance.

Family Child Care Environment Rating Scale – Revised Edition [FCCERS-R] (2007)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
Infant – school-age	2.5 – 5 hours	Individual administering must read and practice the scale & have knowledge of child development.	$18

Available Languages: English

Source: Teachers College Press, Teachers College, Columbia University, New York, NY 10027 (800) 575-6566; www.teacherscollegepress.com

Administration: Observer watches family child care program and asks questions of the provider. Environment is rated on 38 7-point scales (*inadequate* to *excellent*).

Scores Available: Total score, plus subscale scores, each ranging from 1 to 7

Subscales: Space and Furnishings, Personal Care Routines, Listening and Talking, Activities, Interaction, Program Structure, Parents and Providers

Norming Sample: Not normed

Reliability: May be available in tool's introductory materials.

Concurrent Validity: Not reported; may not be other comparable measures.

Infant Toddler Environment Rating Scale - Revised [ITERS-R] (2003)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
0 – 2:6 years	2.5 – 5 hours	Individual administrator must read and practice the scale & have knowledge of child development.	$18

Available Languages: English, Spanish, German, Japanese

Source: Teachers College Press, Teachers College, Columbia University, New York, NY 10027 (800) 575-6566; www.teacherscollegepress.com

Administration: Observer watches Infant Toddler classroom and asks questions of the teacher. Environment is rated on 39 7-point scales (*inadequate* to *excellent*).

Scores Available: Total score, plus subscale scores, each ranging from 1 to 7.

Subscales: Space and Furnishings, Personal Care Routines, Listening and Talking, Activities, Interaction, Program Structure, Parents and Staff

Norming Sample: Not normed

Reliability: High (.80 or higher)

Concurrent Validity: Not reported

School-Age Care Environment Rating Scale [SACERS] (1995)

Age Range	Time to Administer	Administrator Required/Training Needed	Cost
5 – 12 years	2.5 – 5 hours	Individual administering must read and practice the scale and have knowledge of child development.	$17

Available Languages: English, French, German

Source: Teachers College Press, Teachers College, Columbia University, New York, NY 10027 (800) 575-6566; www.teacherscollegepress.com

Administration: Observer watches group-care program for school-aged children (e.g., after school care) and asks questions of the teacher/provider. Environment is rated on 49 7-point scales (*inadequate* to *excellent*).

Scores Available: Total score, plus subscale scores, each ranging from 1 to 7.

Subscales: Space and Furnishings, Health and Safety, Activities, Interaction, Program Structure, Staff Development, Special Needs Supplementary Items

Norming Sample: Not normed

Reliability: May be available in tool's introductory materials.

Concurrent Validity: Not reported (may not be other comparable measures)

Program evaluation assessments answer formative questions about the overall quality of programs, accomplished through careful descriptions of service components, participants, and resources.

Assessment Tools for Children
Ages Birth to 2 Years 11 Months

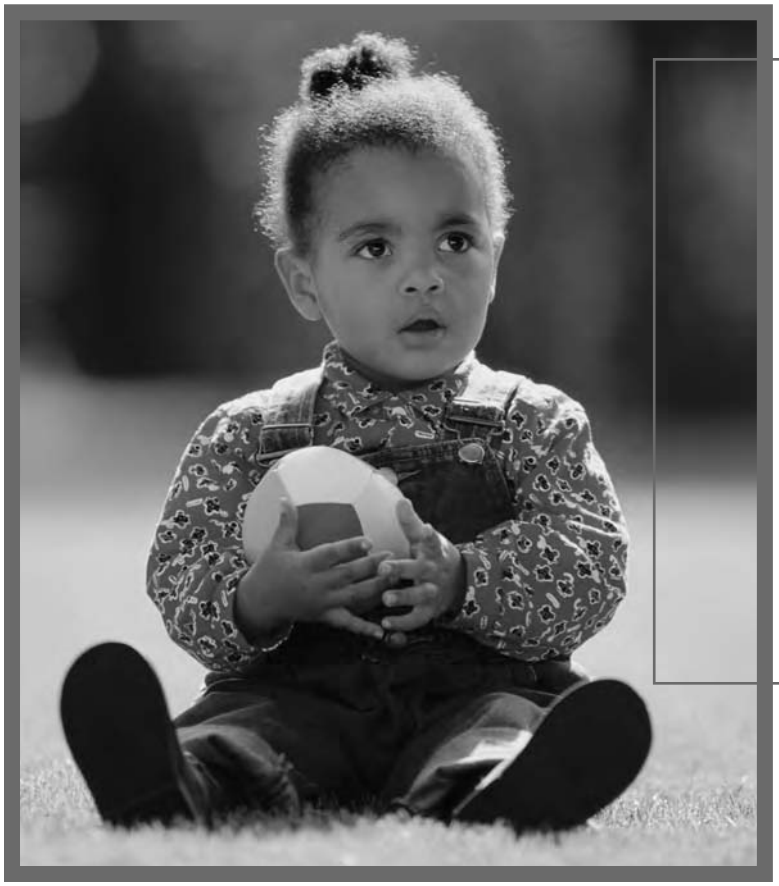
Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
121	1:6–5 yrs	Achenbach System of Empirically Based Assessment (ASEBA) - Preschool Forms			X	
122	0–89 yrs	Adaptive Behavior Assessment System, 2nd Ed. (ABAS-II)		X	X	
93	0:4–5:0 yrs	Ages & Stages Questionnaire, 2nd Ed. (ASQ)	X			
93	0:6–5:0 yrs	Ages & Stages Questionnaire: Social Emotional, 2nd Ed. (ASQ:SE)	X			
123	1:6–18 yrs	Arizona Articulation Proficiency Scale, 3rd Ed.			X	
105	0–3 yrs	Assessment, Evaluation, and Programming System (AEPS) for Birth to Three Years, 2nd Ed.		X		
123	0:0–7:11 yrs	Battelle Developmental Inventory, 2nd Ed. (BDI-II)			X	
94	0:0–7:11 yrs	Battelle Developmental Inventory, 2nd Ed., Screening Test	X			
124	0:1–3:6 yrs	Bayley Scales of Infant and Toddler Development, 3rd Ed. (Bayley-III)			X	
94	0:1–3:6 yrs	Bayley Scales of Infant and Toddler Development, 3rd Ed. (Bayley-III) Screening Test	X			
124	2 yrs–college	Behavior Assessment System for Children, 2nd Ed. (BASC-2)		X	X	
127	2:6–8 yrs	Bracken Basic Concept Scale-Revised (BBCS-R)		X	X	
95	1:0–3:0 yrs	Brief Infant Toddler Social Emotional Assessment (BITSEA)	X			
106	0–7 yrs	Brigance Diagnostic Inventory of Early Development II (IED-II)		X	X	
95	2:0–2:11 yrs	Brigance Early Preschool Screen II	X			
96	0:0–1:11 yrs	Brigance Infant & Toddler Screen	X			
107	0:1–12 yrs	Carey Temperament Scales (CTS)		X		
107	0:0–3:0 yrs	Carolina Curriculum for Infants and Toddlers with Special Needs (CCITSN), 3rd Ed.		X		
108	2–5 yrs	Carolina Curriculum for Preschoolers with Special Needs (CCPSN), 2nd Ed.		X		
129	0:6–6:0 yrs	Communication and Symbolic Behavior Scale Developmental Profile, First Normed Edition (CSBS DP)	X		X	
109	0–2:11 yrs	Creative Curriculum Developmental Continuum for Infants, Toddlers & Twos		X		
131	0–5:11 yrs	Developmental Assessment of Young Children (DAYC)			X	
98	0–6 yrs	Developmental Observation Checklist System (DOCS)	X			
110	2–5 yrs	Devereux Early Childhood Assessment (DECA)		X		
132	2:6–17:11 yrs	Differential Ability Scales-II (DAS-II)		X	X	
162	2:6–5 yrs	Early Childhood Environment Rating Scale – Revised (ECERS-R)				X

Assessment Tools for Children
Ages Birth to 2 Years 11 Months

Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
133	2–18:11 yrs	Expressive One-Word Picture Vocabulary Test (EOWPVT)			X	
133	2:6–90 yrs	Expressive Vocabulary Test, 2nd Ed. (EVT-2)			X	
98	2–16 yrs	Eyberg Child Behavior Inventory (ECBI) and Sutter-Eyberg Student Behavior Inventory-Revised (SESBI-R)	X			
163	Infant – school age	Family Child Care Environment Rating Scale – Revised (FCCERS-R)				X
99	2:9–6:2 yrs	FirstSTEp: Screening Test for Evaluating Preschoolers	X			
134	2–21:11 yrs	Goldman-Fristoe Test of Articulation, 2nd Ed. (GFTA-2)			X	
112	0–3 yrs	Hawaii Early Learning Profile (HELP) (0-3 years)		X		
113	6 wks–3 yrs	High/Scope Child Observation Record for Infants and Toddlers		X		
113	2:6–6 yrs	High/Scope Preschool Child Observation Record		X		
114	0–3 yrs	Individual Growth and Development Indicators for Infants and Toddlers	X	X		
135	0–3:0 yrs	Infant Toddler Developmental Assessment (IDA)			X	
163	0–2:6 yrs	Infant Toddler Environment Rating Scale – Revised (ITERS-R)				X
135	0–3:0 yrs	Infant Toddler Sensory Profile			X	
136	1:0–3:0 yrs	Infant Social Emotional Assessment (ITSEA)			X	
100	0:7–2:6 yrs	Infant Toddler Symptom Checklist (ITSC)	X			
139	2–21 yrs	Khan-Lewis Phonological Analysis, 2nd Ed. (KLPA-2)			X	
139	2:6–6:0 yrs	Learning Accomplishment Profile-Diagnostic (LAP-D), 3rd Ed.			X	
140	2:0–20:11 yrs	Leiter International Performance Scale-Revised (LEITER-R)			X	
141	0:8–3:1 yrs	MacArthur-Bates Communicative Development Inventories (CDI), 2nd Ed.			X	
142	0–6:6 yrs	Merrill-Palmer-Revised Scales of Development (M-P-R)			X	
143	0–5:8 yrs	Mullen Scales of Early Learning (MSEL)			X	
114	0–3:6 yrs	Ounce Scale		X		
101	0–8 yrs	Parents' Evaluation of Development Status	X			
144	0–6:0 yrs	Peabody Developmental Motor Scales		X	X	
145	2:6–90 yrs	Peabody Picture Vocabulary Test			X	
101	1:0–4:0 yrs	Pervasive Developmental Disorders Screening Test-II	X			
148	0–6:11 yrs	Preschool Language Scale			X	
150	0–3 yrs	Receptive Expressive Emergent Language Scale			X	

Assessment Tools for Children
Ages Birth to 2 Years 11 Months

Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
150	2–18:11 yrs	Receptive One-Word Picture Vocabulary Test			X	
151	0:3–80 yrs	Scales of Independent Behavior-Revised	X		X	
152	2:6–6:4 yrs	Social Competence and Behavior Evaluation			X	
153	2:0–7:3 yrs	Stanford-Binet Intelligence Scales for Early Childhood			X	
153	0:11–5:11 yrs	Temperament and Atypical Behavior Scale Screener and Assessment Tool	X		X	
154	2:0–7:11 yrs	Test of Early Language Development			X	
158	0–3:6 yrs	Toddler and Infant Motor Evaluation			X	
117	0–6 yrs	Transdisciplinary Play Based Assessment		X		
158	0–90 yrs	Vineland Adaptive Behavior Scales			X	
159	2:6–7:3 yrs	Wechsler Preschool and Primary Scale of Intelligence			X	
160	2:0–90+ yrs	Woodcock-Johnson III Normative Update Complete			X	



High-quality early education produces long-lasting benefits. With this evidence, federal, state, and local decision makers are asking critical questions about young children's education.

Assessment Tools for Children
Ages 3 Years to 4 Years 11 Months

Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
121	1:6–5 yrs	Achenbach System of Empirically Based Assessment – Preschool Forms			X	
122	0–89 yrs	Adaptive Behavior Assessment System		X	X	
93	0:4–5:0 yrs	Ages & Stages Questionnaire	X			
93	0:6–5:0 yrs	Ages & Stages Questionnaire: Social Emotional	X			
123	1:6–18 yrs	Arizona Articulation Proficiency Scale			X	
105	3–6 yrs	Assessment, Evaluation, and Programming System for Three to Six Years		X		
123	0:0–7:11 yrs	Battelle Developmental Inventory			X	
94	0:0–7:11 yrs	Battelle Developmental Inventory, Screening Test	X			
124	0:1–3:6 yrs	Bayley Scales of Infant and Toddler Development			X	
94	0:1–3:6 yrs	Bayley Scales of Infant and Toddler Development, Screening Test	X			
124	2 yrs–college	Behavior Assessment System for Children		X	X	
125	3:0–5:11 yrs	Boehm Test of Basic Concepts, Preschool		X	X	
127	2:6–8 yrs	Bracken Basic Concept Scale-Revised		X	X	
106	PreK–6th grade	Brigance Comprehensive Inventory of Basic Skills-Revised		X	X	
106	0–7 yrs	Brigance Diagnostic Inventory of Early Development II		X	X	
97	3–4 yrs	Brigance Preschool Screen II	X			
127	4–21 yrs	Bruininks-Oseretsky Test of Motor Proficiency		X	X	
107	0:1–12 yrs	Carey Temperament Scales		X		
108	2–5 yrs	Carolina Curriculum for Preschoolers with Special Needs		X		
161	PreK–3rd grade	Classroom Assessment Scoring System				X
128	3–6:11 yrs	Clinical Evaluation of Language Fundamentals – Preschool		X	X	
129	0:6–6:0 yrs	Communication and Symbolic Behavior Scale Developmental Profile	X		X	
129	3–21 yrs	Comprehensive Assessment of Spoken Language			X	
108	3–5 yrs	Creative Curriculum Developmental Continuum for Ages 3-5		X		
131	0–5:11 yrs	Developmental Assessment of Young Children			X	
97	3–6:11 yrs	Developmental Indicators for the Assessment of Learning	X			
98	0–6 yrs	Developmental Observation Checklist System	X			
110	2–5 yrs	Devereux Early Childhood Assessment		X		
132	2:6–17:11 yrs	Differential Ability Scales-II		X	X	
162	2:6–5 yrs	Early Childhood Environment Rating Scale				X

Assessment Tools for Children
Ages 3 Years to 4 Years 11 Months

Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
162	PreK–3rd grade	Early Language and Literacy Classroom Observation Scale				X
133	2–18:11 yrs	Expressive One-Word Picture Vocabulary Test			X	
133	2:6–90 yrs	Expressive Vocabulary Test			X	
98	2–16 yrs	Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory	X			
163	Infant -school age	Family Child Care Environment Rating Scale				X
99	2:9–6:2 yrs	FirstSTEp: Screening Test for Evaluating Preschoolers	X			
99	3–6:11 yrs	Fluharty Preschool Speech and Language Screening Test	X			
111	3–5 yrs	Get It, Got it, Go	X	X		
134	3–22 yrs	Gilliam Autism Rating Scale			X	
134	2–21:11 yrs	Goldman-Fristoe Test of Articulation			X	
112	3–6 yrs	Hawaii Early Learning Profile for Preschoolers		X		
113	2:6–6 yrs	High/Scope Preschool Child Observation Record		X		
137	3–18 yrs	Kaufman Assessment Battery for Children			X	
137	4:0–90 yrs	Kaufman Brief Intelligence Test			X	
138	3:0–6:11 yrs	Kaufman Survey of Early Academic and Language Skills			X	
138	4:6–90+ yrs	Kaufman Test of Educational Achievement		X	X	
139	2–21 yrs	Khan-Lewis Phonological Analysis			X	
139	2:6–6:0 yrs	Learning Accomplishment Profile-Diagnostic			X	
100	3–5 yrs	Learning Accomplishment Profile – Normed Screens	X			
140	2:0–20:11 yrs	Leiter International Performance Scale-Revised			X	
141	0:8–3:1 yrs	MacArthur-Bates Communicative Development Inventories			X	
142	0–6:6 yrs	Merrill-Palmer-Revised Scales of Development			X	
142	4–7 yrs	Metropolitan Readiness Test			X	
143	0–5:8 yrs	Mullen Scales of Early Learning			X	
114	0–3:6 yrs	Ounce Scale		X		
143	3–21:11 yrs	OWLS: Listening Comprehension Scale and Oral Expression Scale			X	
101	0–8 yrs	Parents' Evaluation of Development Status	X			
144	0–6:0 yrs	Peabody Developmental Motor Scales		X	X	
145	5:0–22:11 yrs	Peabody Individual Achievement Test-Revised, Normative Update			X	

Assessment Tools for Children
Ages 3 Years to 4 Years 11 Months

Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
145	2:6–90 yrs	Peabody Picture Vocabulary Test			X	
101	1:0–4:0 yrs	Pervasive Developmental Disorders Screening Test-II	X			
115	4 yrs (PreK)	Phonological Awareness and Literacy Screenings-PreKindergarten	X	X		
146	3–8 yrs	Pictorial Test of Intelligence			X	
147	3–6 yrs	Preschool and Kindergarten Behavior Scales			X	
147	3:0–5:11 yrs	Preschool Language Assessment Instrument			X	
148	0–6:11 yrs	Preschool Language Scale			X	
116	PreK–1st grade	Qualls Early Learning Inventory		X		
149	3:6–6:5 yrs	Ready to Learn: A Dyslexia Screener	X		X	
150	0–3 yrs	Receptive Expressive Emergent Language Scale			X	
150	2–18:11 yrs	Receptive One-Word Picture Vocabulary Test			X	
151	0:3–80 yrs	Scales of Independent Behavior-Revised	X		X	
151	3:0–10:0 yrs	Sensory Profile			X	
152	2:6–6:4 yrs	Social Competence and Behavior Evaluation			X	
102	3–6:11 yrs	Speed DIAL (Developmental Indicators for the Assessment of Learning)	X			
153	2:0–7:3 yrs	Stanford-Binet Intelligence Scales for Early Childhood			X	
117	3–5 yrs	Teacher Rating of Oral Language & Literacy		X		
153	0:11–5:11 yrs	Temperament and Atypical Behavior Scale Screener and Assessment Tool	X		X	
154	3:0–9:11 yrs	Test for Auditory Comprehension of Language			X	
154	2:0–7:11 yrs	Test of Early Language Development			X	
155	3:0–8:11 yrs	Test of Early Mathematics Ability			X	
155	3:6–8:6 yrs	Test of Early Reading Ability			X	
156	4:0–8:11 yrs	Test of Language Development – Primary			X	
156	3:0–5:11 yrs	Test of Preschool Early Literacy			X	
158	0–3:6 yrs	Toddler and Infant Motor Evaluation			X	
117	0–6 yrs	Transdisciplinary Play Based Assessment		X		

Assessment Tools for Children
Ages 3 Years to 4 Years 11 Months

Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
158	0–90 yrs	Vineland Adaptive Behavior Scales			X	
159	4:0–85 yrs	Wechsler Individual Achievement Test			X	
159	2:6–7:3 yrs	Wechsler Preschool and Primary Scale of Intelligence			X	
160	2:0–90+ yrs	Woodcock-Johnson III Normative Update Complete			X	
118	3 yrs–6th grade	Work Sampling System		X		
118	4:0–7:11 yrs	Young Children's Achievement Test		X		

Assessment Tools for Children
Ages 5 and Older

Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
122	6–18 yrs	Achenbach System of Empirically Based Assessment – School Aged Forms			X	
122	0–89 yrs	Adaptive Behavior Assessment System		X	X	
123	1:6–18 yrs	Arizona Articulation Proficiency Scale			X	
105	3–6 yrs	Assessment, Evaluation, and Programming System for Three to Six Years		X		
123	0:0–7:11 yrs	Battelle Developmental Inventory			X	
94	0:0–7:11 yrs	Battelle Developmental Inventory, Screening Test	X			
124	2 yrs–college	Behavior Assessment System for Children		X	X	
124	5:0–18:11 yrs	Behavioral and Emotional Rating Scale			X	
125	5 yrs–adult	Bilingual Verbal Ability Test - Normative Update			X	
125	3:0–5:11 yrs	Boehm Test of Basic Concepts, Preschool		X	X	
126	K–2nd grade	Boehm Test of Basic Concepts		X	X	
127	2:6–8 yrs	Bracken Basic Concept Scale-Revised		X	X	
96	PreK–6th grade	Brigance Comprehensive Inventory of Basic Skills-Revised		X	X	
127	0–7 yrs	Brigance Diagnostic Inventory of Early Development II		X	X	
96	K–1st	Brigance K & 1 Screen II	X			
127	4–21 yrs	Bruininks-Oseretsky Test of Motor Proficiency		X	X	
107	0:1–12 yrs	Carey Temperament Scales		X		
161	PreK–3rd grade	Classroom Assessment Scoring System				X
128	5–21 yrs	Clinical Evaluation of Language Fundamentals		X	X	
128	3–6:11 yrs	Clinical Evaluation of Language Fundamentals-Preschool		X	X	
129	0:6–6:0 yrs	Communication and Symbolic Behavior Scale Developmental Profile, First Normed Edition	X		X	
129	3–21 yrs	Comprehensive Assessment of Spoken Language			X	
130	5–24:11 yrs	Comprehensive Test of Phonological Processing			X	
130	6–18 yrs	Conners			X	
131	6–18 yrs	Conners Comprehensive Behavior Rating Scales			X	
131	0–5:11 yrs	Developmental Assessment of Young Children			X	
97	3–6:11 yrs	Developmental Indicators for the Assessment of Learning	X			
98	0–6 yrs	Developmental Observation Checklist System	X			
109	K–3rd grade	Developmental Reading Assessment		X		
110	2–5 yrs	Devereux Early Childhood Assessment		X		

Assessment Tools for Children
Ages 5 and Older

Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
132	2:6–17:11 yrs	Differential Ability Scales-II		X	X	
110	K–6th grade	Dynamic Indicators of Basic Early Literacy Skills	X	X		
162	PreK–3rd grade	Early Language and Literacy Classroom Observation Scale				X
132	K–3rd grade	Early Reading Diagnostic Assessment		X	X	
111	K–1st grade	ECLS-K Approaches to Learning Sub-Scale		X		
133	2–18:11 yrs	Expressive One-Word Picture Vocabulary Test			X	
133	2:6–90 yrs	Expressive Vocabulary Test			X	
98	2–16 yrs	Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory-Revised	X			
163	Infant- school-age	Family Child Care Environment Rating Scale – Revised				X
99	2:9–6:2 yrs	FirstSTEp: Screening Test for Evaluating Preschoolers	X			
99	3–6:11 yrs	Fluharty Preschool Speech and Language Screening Test	X			
134	3–22 yrs	Gilliam Autism Rating Scale			X	
134	2–21:11 yrs	Goldman-Fristoe Test of Articulation			X	
112	3–6 yrs	Hawaii Early Learning Profile for Preschoolers		X		
113	2:6–6 yrs	High/Scope Preschool Child Observation Record		X		
136	5–14 yrs	Iowa Test of Basic Skills, Forms A & B			X	
137	3–18 yrs	Kaufman Assessment Battery for Children			X	
137	4:0–90 yrs	Kaufman Brief Intelligence Test			X	
138	3:0–6:11 yrs	Kaufman Survey of Early Academic and Language Skills			X	
138	4:6–90+ yrs	Kaufman Test of Educational Achievement		X	X	
139	2–21 yrs	Khan-Lewis Phonological Analysis			X	
139	2:6–6:0 yrs	Learning Accomplishment Profile-Diagnostic			X	
140	2:0–20:11 yrs	Leiter International Performance Scale-Revised			X	
141	5:0–18:11 yrs	Lindamood Auditory Conceptualization Test			X	
142	0–6:6 yrs	Merrill-Palmer-Revised Scales of Development			X	
142	4–7 yrs	Metropolitan Readiness Test			X	
143	0–5:8 yrs	Mullen Scales of Early Learning			X	
143	3–21:11 yrs	OWLS: Listening Comprehension Scale and Oral Expression Scale			X	
144	5:0–21:11 yrs	OWLS: Written Expression Scale			X	
101	0–8 yrs	Parents' Evaluation of Development Status	X			

Assessment Tools for Children
Ages 5 and Older

Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
144	0–6:0 yrs	Peabody Developmental Motor Scales		X	X	
145	5:0–22:11 yrs	Peabody Individual Achievement Test-Revised, Normative Update			X	
145	2:6–90 yrs	Peabody Picture Vocabulary Test			X	
115	5 yrs (K)	Phonological Awareness and Literacy Screenings-Kindergarten	X	X		
116	1st–3rd grade	Phonological Awareness and Literacy Screenings-1-3	X	X		
146	5:0–9:11 yrs	Phonological Awareness Test 2		X	X	
146	3–8 yrs	Pictorial Test of Intelligence			X	
147	3–6 yrs	Preschool and Kindergarten Behavior Scales			X	
147	3:0–5:11 yrs	Preschool Language Assessment Instrument			X	
148	0–6:11 yrs	Preschool Language Scale			X	
148	K–6th grade	Process Assessment of the Learner: Diagnostic Assessment for Math		X	X	
149	K–6th grade	Process Assessment of the Learner: Diagnostic Assessment for Reading and Writing		X	X	
116	PreK –1st grade	Qualls Early Learning Inventory		X		
149	3:6–6:5 yrs	Ready to Learn: A Dyslexia Screener	X		X	
150	2–18:11 yrs	Receptive One-Word Picture Vocabulary Test			X	
151	0:3–80 yrs	Scales of Independent Behavior-Revised	X		X	
164	5–12 yrs	School-Age Care Environment Rating Scale				X
151	3:0–10:0 yrs	Sensory Profile			X	
152	2:6–6:4 yrs	Social Competence and Behavior Evaluation			X	
102	3–6:11 yrs	Speed DIAL (Developmental Indicators for the Assessment of Learning)	X			
152	K–12	Stanford Achievement Test			X	
153	2:0–7:3 yrs	Stanford-Binet Intelligence Scales for Early Childhood			X	
153	0:11–5:11 yrs	Temperament and Atypical Behavior Scale Screener and Assessment Tool	X		X	
154	3:0–9:11 yrs	Test for Auditory Comprehension of Language			X	
154	2:0–7:11 yrs	Test of Early Language Development			X	
155	3:0–8:11 yrs	Test of Early Mathematics Ability			X	
155	3:6–8:6 yrs	Test of Early Reading Ability			X	
156	4:0–8:11 yrs	Test of Language Development - Primary			X	
156	5–8 yrs	Test of Phonological Awareness			X	
156	3:0–5:11 yrs	Test of Preschool Early Literacy			X	

Assessment Tools for Children
Ages 5 and Older

Page	Age Range	Name	Section/Purpose			
			1. Screening	2. Inform / monitor instruction	3. Diagnostic	4. Evaluation
157	6:0–24:11 yrs	Test of Word Reading Efficiency		X	X	
117	0–6 yrs	Transdisciplinary Play Based Assessment		X		
158	0–90 yrs	Vineland Adaptive Behavior Scales			X	
159	4:0–85 yrs	Wechsler Individual Achievement Test			X	
159	2:6–7:3 yrs	Wechsler Preschool and Primary Scale of Intelligence			X	
160	2:0–90+ yrs	Woodcock-Johnson III Normative Update Complete			X	
118	3 yrs–6th grade	Work Sampling System		X		
118	4:0–7:11 yrs	Young Children's Achievement Test		X		

Children from early childhood through primary grades—and beyond—need to be cognitively, physically, socially, and artistically active.

Information Resources for Part II

Information about specific tools presented in this guide comes primarily from four sources:

- Publishers' websites listed in the **Source** section of each tool
- Buros Institute (9th to 17th Yearbooks, July 2007), *Ovid mental measurements yearbook*, accessed online in December 2007 and January 2008
- Child Trends (Berry, D. J., Bridges, L. J., & Zaslow, M. J.) (2004, September). *Early childhood measures profiles.* Retrieved January 21, 2008, from http://www.childtrends.org/Files//Child_Trends-2004_09_01_FR_ECMeasures.pdf
- Mathematica Policy Research, Inc. (Kisker, E. E., Boller, K., Nagatoshi, C., Sciarrino, C., Jethwani, V., Zavitsky, T., Ford, M., & Love, J. M.) (2003, April). *Resources for measuring services and outcomes in Head Start programs serving infants and toddlers.* Retrieved January 21, 2008, from http://www.acf.hhs.gov/programs/opre/ehs/perf_measures/reports/resources_measuring/resources_for_measuring.pdf

Other sources used less frequently in this guide include:

- Technical manuals for individual tools, when available
- SERVE (Niemeyer, J. & Scott-Little, C.) (2001). *Assessing kindergarten children: A compendium of assessment instruments.* Retrieved January 21, 2008, from http://www.serve.org/_downloads/publications/rdakcc.pdf
- The Early Childhood Outcomes Center. (n.d.). *Instrument crosswalks.* Retrieved January 21, 2008, from http://www.fpg.unc.edu/~eco/crosswalks.cfm

Additional information about specific tools presented in this guide comes from the following sources:

- *ECLS-K Approaches to Learning Sub-Scale:*
  U.S. Department of Education, National Center for Education Statistics. Assessment of social competence, adaptive behaviors, and approaches to learning with young children. Working Paper No. 96-18, by Samuel J. Meisels, Sally Atkins-Burnett, and Julie Nicholson. Jerry West, project officer. Washington, D.C.: 1996, Retrieved January 21, 2008, from, http://nces.ed.gov/pubs96/9618.pdf

  U.S. Department of Education, National Center for Education Statistics. ECLS-K Base Year Public-Use Data Files and Electronic Code Book. Washington, D.C. Retrieved January 21, 2008, from: http://nces.ed.gov/pubs2001/2001029rev_1_4.pdf
- *Get it, Got It, Go:*
  Early Childhood Research Institute on Measuring Growth and Development, funded through the Office of Special Education and Rehabilitative Services, U.S. Department of Education (Grant No.: H024S60010). Retrieved January 21, 2008, from ggg.umn.edu

- *Test of Preschool Early Literacy (TOPEL):*
  Lonigan, C. J. (2006). Development, assessment, and promotion of preliteracy skills. *Early Education and Development, 17*(1), 91–114.

Other sources of information used in framing this document:

- National Child Care Information Center (2005, June). *Assessment and evaluation: Becoming an educated consumer. Part I: Child assessment.* Retrieved January 26, 2008, from http://nccic.acf.hhs.gov/pubs/goodstart/assess-eval1.pdf
- National Child Care Information Center (2005, June). *Assessment and evaluation: Becoming an educated consumer. Part II: Program evaluation.* Retrieved January 26, 2008, from http://nccic.acf.hhs.gov/pubs/goodstart/assess-eval2.pdf
- National Child Care Information Center (2005, June). *Assessment and evaluation: Becoming an educated consumer. Part III: Accountability systems.* Retrieved January 26, 2008, from http://nccic.acf.hhs.gov/pubs/goodstart/assess-eval3.pdf
- National Institute for Early Education (n.d.) Research assessment database. Retrieved January 21, 2008, from http://nieer.org/assessment/list.php
- Psychological and Educational Publications, Inc. (2006). *Emotional and conduct assessments.* Retrieved January 21, 2008, from http://www.psych-edpublications.com/emotional.htm#asiep
- School Psychiatry Program & MADI Resource Center, Massachusetts General Hospital (2006). *Table of all screening tools & rating scales: Pervasive developmental disorder and autism spectrum detail.* Retrieved January 21, 2008, from http://www.mgh.harvard.edu/madiresourcecenter/schoolpsychiatry/screening_pdd.asp
- Southwest Developmental Education Laboratory. (2000). Reading assessment database for grades K-2. Retrieved January 21, 2008, from http://www.sedl.org/reading/rad/database.html
- University of New Mexico's Center for Development & Disability. New Assessment: Early Childhood Resources. (n.d.). *Early childhood Assessment Measures.* Retrieved January 21, 2008, from http://www.newassessment.org/Public/Assessments/selecttool.cfm?CategoryID=20
- U.S. Department of Health & Human Services, Administration for Children & Families, National Child Care Information Center (2004, November). *Early language & literacy classroom observation and assessment tools.* Retrieved January 21, 2008, from http://www.nccic.org/pubs/goodstart/assessment-literacy.html

Appendix A:

Details Regarding High, Adequate, and Low Ratings of Reliability, Concurrent Validity, and Sensitivity/Specificity

Reliability

Each tool's reliability is compared to the following scale:

High = .80 or higher

Adequate = .65 - .79

Low = below .65

This rating system was originally based on the Mathematica Policy Research, Inc., report referenced in Sources of Information, Part 2. However, that report did not have a *high* category, only distinguishing between measures above or below .65. The authors believed that adding a high category would be most useful for readers and .80 is a common cutpoint for establishing high reliability.

The three most common ways of measuring reliability are:

1) *internal* - the extent to which the items within the scale measure the same construct;

2) *test-retest* - the extent to which the same child gets the same score when given the same test twice, a few days apart; and

3) *inter-rater* - the extent to which different assessors/observers give the same scores on the assessment of the same child/classroom.

Test publishers often report multiple types of reliability, but there is little consistency across tools in what types of reliability are reported. For this reason, the authors decided to report whichever type of reliability (see above) was highest, so as not to penalize tests/publishers that honestly report multiple types of reliability.

Reliability for the most general scale or score (e.g., overall or total score) is compared to the rating scale. If only sub-scale reliabilities are reported, the mean or midpoint of reported range of reliabilities was used.

Concurrent Validity

For tools whose purpose is *inform instruction/monitor progress*, *diagnostic*, or *program evaluation*, concurrent validity is reported. Concurrent validity is also reported for *screening* tools when Sensitivity/Specificity information is unavailable.

Each tool's concurrent validity is compared to the following scale:

High = .70 or higher

Adequate = .50 -.69

Low = below .50

As with the reliability rating scale, these cutpoints were based on Mathematica Policy Research, Inc., report referenced in Sources of Information, Part 2 However, that report only differentiated between those that were above and below .50 The "high" category was included to provide additional information to users of this guide.

The concurrent validity for the most general scale or score (e.g., overall or total score) is compared to the rating scale If only subscale scores are available, the mean or midpoint of the range was used If concurrent validity with several other measures was reported, the highest reported values are compared to the rating scale Concurrent validity with earlier versions of the same tool were not used for this rating Note, non-significant correlations are reported as *low,* regardless of their size When publishers report validity information for multiple ages, only those that fall in the 0–8 age range are included.

Sensitivity/Specificity

For tools whose purpose is *screening*, sensitivity and selectivity are reported. Sensitivity reflects the percent of children with a disability who are referred for further testing based on screening results. Specificity reflects the percent of children without a disability who are not referred for further testing, based on screening results.

Each tool's sensitivity/specificity is compared to the following scale:

High = 80% or higher

Adequate = 65% to 79%

Low = below 65%

Appendix B:

Position Paper on Building an Effective, Accountable System in Programs for Children Birth through Age 8

The following document is a position statement of the National Association for the Education of Young Children and the National Association of Early Childhood Specialists in State Departments of Education.

Early Childhood Curriculum, Assessment, and Program Evaluation

A Joint Position Statement of the National Association for the Education of Young Children (NAEYC) and the National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE)

Introduction

High-quality early education produces long-lasting benefits. With this evidence, federal, state, and local decision makers are asking critical questions about young children's education. What should children be taught in the years from birth through age eight? How would we know if they are developing well and learning what we want them to learn? And how could we decide whether programs for children from infancy through the primary grades are doing a good job?

Answers to these questions—questions about early childhood curriculum, child assessment, and program evaluation—are the foundation of this joint position statement from the National Association for the Education of Young Children (NAEYC) and the National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE).

The Position

The National Association for the Education of Young Children and the National Association of Early Childhood Specialists in State Departments of Education take the position that policy makers, the early childhood profession, and other stakeholders in young children's lives have a shared responsibility to:

- Construct comprehensive systems of curriculum, assessment, and program evaluation guided by sound early childhood practices, effective early learning standards and program standards, and a set of core principles and values: belief in civic and democratic values; commitment to ethical behavior on behalf of children; use of important goals as guides to action; coordinated systems; support for children as individuals and members of families, cultures, and communities; partnerships with families; respect for evidence; and shared accountability.

- Implement curriculum that is thoughtfully planned, challenging, engaging, developmentally appropriate, culturally and linguistically responsive, comprehensive, and likely to promote positive outcomes for all young children.
- Make ethical, appropriate, valid, and reliable assessment a central part of all early childhood programs. To assess young children's strengths, progress, and needs, use assessment methods that are developmentally appropriate, culturally and linguistically responsive, tied to children's daily activities, supported by Professional development, inclusive of families, and connected to specific, beneficial purposes: (1) making sound decisions about teaching and learning, (2) identifying significant concerns that may require focused intervention for individual children, and (3) helping programs improve their educational and developmental interventions.
- Regularly engage in program evaluation guided by program goals and using varied, appropriate, conceptually and technically sound evidence to determine the extent to which programs meet the expected standards of quality and to examine intended as well as unintended results.
- Provide the support, Professional development, and other resources to allow staff in early childhood programs to implement high-quality curriculum, assessment, and program evaluation practices and to connect those practices with well-defined early learning standards and program standards.

Recommendations

Curriculum

Implement curriculum that is thoughtfully planned, challenging, engaging, developmentally appropriate, culturally and linguistically responsive, comprehensive, and likely to promote positive outcomes for all young children.

Indicators of Effectiveness

- *Children are active and engaged.*
  Children from early childhood through primary grades—and beyond—need to be cognitively, physically, socially, and artistically active. In their own ways, children of all ages and abilities can become interested and engaged, develop positive attitudes toward learning, and have their feelings of security, emotional competence, and linkages to family and community support
- *Goals are clear and shared by all.*
  Curriculum goals are clearly defined, shared, and understood by all "stakeholders" (for example, program administrators, teachers, and families). The curriculum and related activities and teaching strategies are designed to help achieve these goals in a unified, coherent way.
- *Curriculum is evidence-based.*
  The curriculum is based on evidence that is developmentally, culturally, and linguistically relevant for the children who will experience the curriculum. It is organized around principles of child development and learning.

- *Valued content is learned through investigation, play, and focused, intentional teaching.*
  Children learn by exploring, thinking about, and inquiring about all sorts of phenomenal material. These experiences help children investigate "big ideas," those that are important at any age and are connected to later learning. Pedagogy or teaching strategies are tailored to children's ages, developmental capacities, language and culture, and abilities or disabilities.

- *Curriculum builds on prior learning and experiences.*
  The content and implementation of the curriculum builds on children's prior individual, age-related, and cultural learning, is inclusive of children with disabilities, and is supportive of background knowledge gained at home and in the community. The curriculum supports children whose home language is not English in building a solid base for later learning.

- *Curriculum is comprehensive.*
  The curriculum encompasses critical areas of development including children's physical well-being and motor development; social and emotional development; approaches to learning; language development; and cognition and general knowledge; and subject matter areas such as science, mathematics, language, literacy, social studies, and the arts (more fully and explicitly for older children).

- *Professional standards validate the curriculum's subject-matter content.*
  When subject-specific curricula are adopted, they meet the standards of relevant Professional organizations (for example, the American Alliance for Health, Physical Education, Recreation and Dance [AAHPERD], the National Association for Music Education [MENC]; the National Council of Teachers of English [NCTE]; the Council of Teachers of Mathematics [NCTM]; the National Dance Education Organization [NDEO]; the National Science Teachers Association [NSTA]) and are reviewed and implemented so that they fit together coherently.

- *The curriculum is likely to benefit children.*
  Research and other evidence indicates that the curriculum, if implemented as intended, will likely have beneficial effects. These benefits include a wide range of outcomes. When evidence is not yet available, plans are developed to obtain this evidence.

Assessment of Young Children

Make ethical, appropriate, valid, and reliable assessment a central part of all early childhood programs. To assess young children's strengths, progress, and needs, use assessment methods that are developmentally appropriate, culturally and linguistically responsive, tied to children's daily activities, supported by professional development, inclusive of families, and connected to specific, beneficial purposes:

1) making sound decisions about teaching and learning;

2) identifying significant concerns that may require focused intervention for individual children, and

3) helping programs improve their educational and developmental interventions.

Indicators of Effectiveness

- *Ethical principles guide assessment practices.*
  Ethical principles underlie all assessment practices. Young children are not denied opportunities or services, and decisions are not made about children on the basis of a single assessment.

- *Assessment instruments are used for their intended purposes.*
  Assessments are used in ways consistent with the purposes for which they were designed If the assessments will be used for additional purposes, they are validated for those purposes.

- *Assessments are appropriate for ages and other characteristics of children being assessed.*
  Assessments are designed for and validated for use with children whose ages, cultures, home languages, socioeconomic status, abilities and disabilities, and other characteristics are similar to those of the children with whom the assessments will be used.

- *Assessment instruments are in compliance with Professional criteria for quality.*
  Assessments are valid and reliable. Accepted Professional standards of quality are the basis for selection, use, and interpretation of assessment instruments, including screening tools. NAEYC and NAECS/SDE support and adhere to the measurement standards set forth in 1999 by the American Educational Research Association, the American Psychological Association, and the National Center for Measurement in Education. When individual norm-referenced tests are used, they meet these guidelines.

- *What is assessed is developmentally and educationally significant.*
  The objects of assessment include a comprehensive, developmentally, and educationally important set of goals, rather than a narrow set of skills. Assessments are aligned with early learning standards, with program goals, and with specific emphases in the curriculum.

- *Assessment evidence is used to understand and improve learning.*
  Assessments lead to improved knowledge about children. This knowledge is translated into improved curriculum implementation and teaching practices. Assessment helps early childhood Professionals understand the learning of a specific child or group of children; enhance overall knowledge of child development; improve educational programs for young children while supporting continuity across grades and settings; and access resources and supports for children with specific needs.

- *Assessment evidence is gathered from realistic settings and situations that reflect children's actual performance.*
  To influence teaching strategies or to identify children in need of further evaluation, the evidence used to assess young children's characteristics and progress is derived from real-world classroom or family contexts that are consistent with children's culture, language, and experiences.

- *Assessments use multiple sources of evidence gathered over time.*
  The assessment system emphasizes repeated, systematic observation, documentation, and other forms of criterion- or performance-oriented assessment using broad, varied, and complementary methods with accommodations for children with disabilities.

- *Screening is always linked to follow-up.*
  When a screening or other assessment identifies concerns, appropriate follow-up, referral, or other intervention is used Diagnosis or labeling is never the result of a brief screening or one-time assessment.

- *Use of individually administered, norm-referenced tests is limited.*
  The use of formal standardized testing and norm-referenced assessments of young children is limited to situations in which such measures are appropriate and potentially beneficial, such as identifying potential disabilities. (See also the indicator concerning the use of individual norm-referenced tests as part of program evaluation and accountability.)

- *Staff and families are knowledgeable about assessment.*
  Staff are given resources that support their knowledge and skills about early childhood assessment and their ability to assess children in culturally and linguistically appropriate ways. Preservice and in-service training builds teachers' and administrators' "assessment literacy," creating a community that sees assessment as a tool to improve outcomes for children. Families are part of this community, with regular communication, partnership, and involvement.

Program Evaluation and Accountability

Regularly evaluate early childhood programs in light of program goals, using varied, appropriate, conceptually and technically sound evidence to determine the extent to which programs meet the expected standards of quality and to examine intended as well as unintended results.

Indicators of Effectiveness

- *Evaluation is used for continuous improvement.*
  Programs undertake regular evaluation, including self-evaluation, to document the extent to which they are achieving desired results, with the goal of engaging in continuous improvement. Evaluations focus on processes and implementation as well as outcomes. Over time, evidence is gathered that program evaluations do influence specific improvements.

- *Goals become guides for evaluation.*
  Evaluation designs and measures are guided by goals identified by the program, by families and other stakeholders, and by the developers of a program or curriculum, while also allowing the evaluation to reveal unintended consequences.

- *Comprehensive goals are used.*
  The program goals used to guide the evaluation are comprehensive, including goals related to families, teachers and other staff, and community as well as child-oriented goals that address a broad set of developmental and learning outcomes.

- *Evaluations use valid designs.*
  Programs are evaluated using scientifically valid designs, guided by a "logic model" that describes ways in which the program sees its interventions having both medium- and longer-term effects on children and, in some cases, families and communities.

- *Multiple sources of data are available.*
  An effective evaluation system should include multiple measures, including program data, child demographic data, information about staff qualifications, administrative practices, classroom quality assessments, implementation data, and other information that provides a context for interpreting the results of child assessments.

- *Sampling is used when assessing individual children as part of large-scale program evaluation.*
  When individually administered, norm-referenced tests of children's progress are used as part of program evaluation and accountability, matrix sampling is used (that is, administered only to a systematic sample of children) so as to diminish the burden of testing on children and to reduce the likelihood that data will be inappropriately used to make judgments about individual children.

- *Safeguards are in place if standardized tests are used as part of evaluations.*
  When individually administered, norm-referenced tests are used as part of program evaluation, they must be developmentally and culturally appropriate for the particular children in the program, conducted in the language children are most comfortable with, with other accommodations as appropriate, valid in terms of the curriculum, and technically sound (including reliability and validity). Quality checks on data are conducted regularly, and the system includes multiple data sources collected over time.

- *Children's gains over time are emphasized.*
  When child assessments are used as part of program evaluation, the primary focus is on children's gains or progress as documented in observations, samples of classroom work, and other assessments over the duration of the program. The focus is not just on children's scores upon exit from the program.

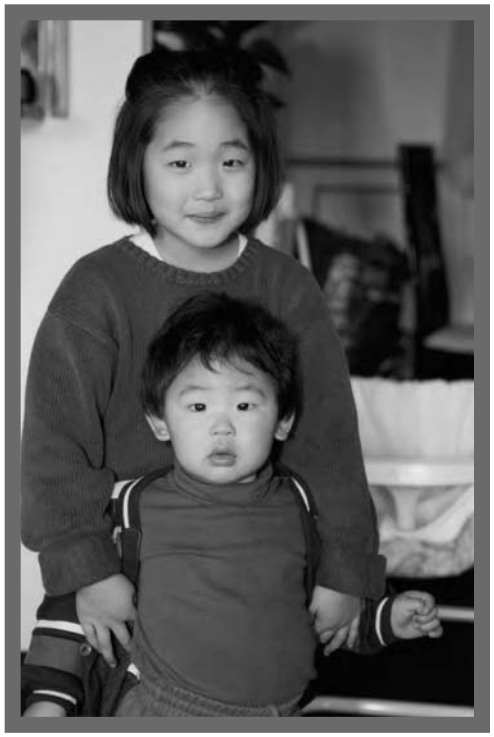
- *Well-trained individuals conduct evaluations.*
  Program evaluations, at whatever level or scope, are conducted by well-trained individuals who are able to evaluate programs in fair and unbiased ways. Self-assessment processes used as part of comprehensive program evaluation follow a valid model. Assessor training goes beyond single workshops and includes ongoing quality checks. Data are analyzed systematically and can be quantified or aggregated to provide evidence of the extent to which the program is meeting its goals.

- *Evaluation results are publicly shared.*
  Families, policy makers, and other stakeholders have the right to know the results of program evaluations. Data from program monitoring and evaluation, aggregated appropriately and based on reliable measures, should be made available and accessible to the public.

Creating Change Through Support for Programs

Implementing the preceding recommendations for curriculum, child assessment, and program evaluation requires a solid foundation. Calls for better results and greater accountability from programs for children in preschool, kindergarten, and the primary grades have not been backed up by essential supports for teacher recruitment and compensation, Professional preparation and ongoing Professional development, and other ingredients of quality early education.

The overarching need is to create an integrated, well-financed system of early care and education that has the capacity to support learning and development in all children, including children living in poverty, children whose home language is not English, and children with disabilities. Unlike many other countries, the United States continues to have a fragmented system for educating children from birth through age eight, under multiple auspices, with greatly varying levels of support, and with inadequate communication and collaboration.



Many challenges face efforts to provide all young children with high-quality curriculum, assessment, and evaluation of their programs. Public commitment, along with investments in a well-financed system of early childhood education and in other components of services for young children and their families, will make it possible to implement these recommendations fully and effectively.

This document is an official position statement of the National Association for the Education of Young Children and the National Association of Early Childhood Specialists in State Departments of Education.

Copyright © 2BirthBirth3 by the National Association for the Education of Young Children

Reprinted with permission.
Contact us at: pubaff@Nalyc.org
Approved November 2BirthBirth3

NOTES

©2008 Washington State Office of Superintendent of Public Instruction. This document may be freely reproduced without permission for non-profit, educational purposes. Use of or reference to this document should cite: *A Guide to Assessment in Early Childhood; Infancy to Age Eight. Washington State Office of Superintendent of Public Instruction, 2008.*