Understanding and Choosing Assessments and Developmental Screeners for Young Children Ages 3-5: Profiles of Selected Measures

OPRE Report # 2011-23

Understanding and Choosing Assessments and Developmental Screeners for Young Children Ages 3-5: Profiles of Selected Measures

Final Report

OPRE Report # 2011-23

June 2011

Contract Number: GS10F0030R

This report is in the public domain. Permission to reproduce is not necessary.
Suggested citation:
Halle, T., Zaslow, M., Wessel, J., Moodie, S., and Darling-Churchill, K. (2011). Understanding and Choosing Assessments and Developmental Screeners for Young Children: Profiles of Selected Measures. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

DISCLAIMER

The views expressed in this publication do not necessarily reflect the views or policies of the Office of Planning, Research and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services.

This report and other reports sponsored by the Office of Planning, Research and Evaluation are available at http://www.acf.hhs.gov/programs/opre/index.html.

Table of Contents

Introduction **5**

Summary Tables **14**

General Information about Assessments 15

Assessments: Evidence of Reliability and Validity 16

Assessments: Evidence of Reliability and Validity for Different Languages and Different Populations 17

General Information about Developmental Screeners 18

Developmental Screeners: Evidence of Reliability and Validity 19

Developmental Screeners: Evidence of Reliability and Validity for Different Languages and Different Populations 20

Profiles of Individual Measures **21**

Assessments 21

Creative Curriculum Developmental Assessment 22

Galileo Preschool Assessment Scales 30

High Scope Child Observation Record 38

Learning Accomplishment Profile—3rd Edition 46

Learning Accomplishment Profile—Diagnostic 55

*Learning Accomplishment Profile—Diagnostic, Spanish Edition** 65

*Mullen Scales of Early Learning** 75

Work Sampling System for Head Start 85

Developmental Screeners 94

Ages and Stages Questionnaire—3rd Edition 95

Ages and Stages Questionnaire—Social-Emotional 104

Battelle Developmental Inventory Screening Test 113

Brigance Screens 121

Denver II 130

*Developmental Indicators for the Assessment of Early Learning—3rd Edition** 139

Early Screening Inventory 150

Learning Accomplishment Profile—Diagnostic Screens 160

Parents' Evaluation of Developmental Status 169

Parents' Evaluation of Developmental Status: Developmental Milestones 178

Appendix A: Glossary of Terms **187**

Appendix B: Psychometric Documentation and Rationale **193**

Note: * Indicates that the measure has not been reviewed by the developer. The remaining profiles have been reviewed by the developer and updated according to his/her feedback.

Introduction

The 2007 reauthorization of Head Start requires Head Start programs to use child assessments and developmental screeners that are developmentally, linguistically, and culturally appropriate, as well as valid and reliable in the language in which they are used. This can be a challenge, since very few child assessment tools are developed or tested with linguistically and culturally diverse samples or with samples of children with disabilities (Pena & Halle, 2011; Spiker, Hebbeler, & Barton, 2011). Further, staff members in Head Start and other early childhood education programs rarely have the time, and may lack the technical training, to review and compare complex psychometric information on the quality of assessment and developmental screening tools. This compendium has been created to address this need and to promote the use of reliable and valid assessment data, wherever possible, in Head Start and other early childhood programs.

Purpose of this Compendium

This document has three purposes. First, the compendium aims to help Head Start managers and other early childhood education administrators review information regarding the reliability and validity of commonly used assessment and developmental screening tools in order to help them better select appropriate tools for the populations they serve. Second, the compendium aims more generally to increase awareness about reliability and validity and how to evaluate whether an instrument is reliable and valid for the population and purpose for which it will be used. Finally, the compendium aims to highlight areas in which the early childhood field is lacking information on reliability and validity of early childhood assessments and developmental screeners. While originally developed in response to Head Start's reauthorization, the compendium is designed to be useful to managers and staff who work in different types of early childhood education programs and who are responsible for selecting and evaluating assessment or screening instruments.

With regard to the first purpose, many manuals provide complex information regarding the reliability, validity, and appropriateness of a particular assessment instrument for different populations. This information can be detailed and is not always presented in a way that is easily understood to those who are not trained researchers. Head Start and other program managers may need help sifting through and understanding the information that is provided in these formats. This challenge is amplified by the fact that an instrument may be reliable and valid with one population but not another, or for one purpose but not another.

This compendium is meant to aid in these efforts by doing the following:

1) Summarize information from the assessment and developmental screening instruments most commonly used by Head Start programs for 3-, 4-, and 5-year-olds.

2) Share reliability and validity data as it relates to key populations of children, including dual language learners[1] and children with disabilities, as well as the children served in American Indian/Alaska Native and Migrant and Seasonal Head Start programs.

[1] It is important to note that when developers assess the reliability and validity of assessments and developmental screeners, they do not always indicate whether the children in the sample were dual

3) Translate information on reliability and validity in a way that is quicker and easier for Head Start managers and others to understand and use.

Yet, the compendium also has a broader goal: to increase understanding of reliability and validity more generally among Head Start and early childhood managers. To this end, when profiling assessments and developmental screeners, the compendium introduces different types of reliability and validity and conveys psychometric information in an accessible and easy-to-use format. This compendium reviews a select set of assessment and developmental screening tools and likely does not address every factor that managers must consider in selecting a tool. However, all programs can benefit from a greater understanding of reliability and validity, and what is used to understand whether a tool is reliable or valid for a particular population and purpose.

Finally, with regard to the third purpose, this compendium aims to highlight where greater information on the reliability and validity of assessments is needed. As shown in the summary tables in this compendium, very few of the selected assessments or developmental screeners have documented reliability or validity for key populations, including dual language learners, American Indian/Alaskan Native populations, and children with special needs. Often these populations were not included in studies used to evaluate the effectiveness of the instruments or to develop norms. When they were included in the studies, they were rarely examined separately, which is necessary for determining reliability and validity for that population. Further, while many of the tools examined have versions in languages other than English, only one instrument has detailed information on the reliability and validity of a Spanish version based on a review of the manuals. Likewise, most instruments have limited information on key types of reliability and validity (for example, predictive validity), with those types of reliability and validity that are the easiest to examine–and often easiest to achieve–being the most commonly reported.

What Are Reliability and Validity and Why Are They Important?

Before describing what is covered in this compendium, it seems critical to define reliability and validity, and to highlight why they are important to Head Start and other early childhood programs. Information on the reliability and validity of an assessment or screening tool is critical to determining whether that tool is appropriate for use in a program. If an instrument does not produce reliable or valid information, one cannot trust that information to provide a good sense of how children are doing.

But what are reliability and validity?

By **reliability**, we mean that the scores on the tool will be stable regardless of *when* the tool is administered, *where* it is administered, and *who* is administering it. Reliability answers the question: Is the tool producing consistent information across different circumstances? Reliability provides assurance that comparable information will be obtained from the tool across

---

language learners or monolingual speakers. In such cases, the summary tables in this compendium provide a rating of "No evidence –No information about this population is provided by the developer" with regard to reliability or validity for dual language learners. This simply means that the developer did not record or convey this information.

different situations. By **validity**, we mean that the scores on the tool accurately capture what the tool is meant to capture in terms of content. Validity answers the question: Is the tool assessing what it is supposed to assess?

There are many types of reliability and validity, and each has a role to play in the development of child assessment and screening tools. For example, *content validity* assures program staff that a measurement tool is assessing the behaviors or skills of interest by measuring all key indicators of those skills. *Construct validity* indicates that the items of an assessment tool or developmental screener are capturing the aspects of development that are the focus of the measure and of importance to program staff.[2] *Internal consistency reliability* refers to how closely items within a measure are related to one another; this type of reliability ensures that all of the items within a particular domain[3] actually are related to each other but still are distinct enough as to not be redundant within the measurement tool. *Convergent and divergent validity* refers to how closely different domains within the measurement tool are related to one another. Similarly, *convergent and divergent criterion validity* refers to the degree to which constructs within one measurement tool are related in an expected pattern to other established measurement tools.[4] *Predictive validity* tests how a measure relates to an important, established measurement tool that assesses the same outcome at a later point in time.

Not only should a measure capture what it is supposed to be capturing, it also should do so consistently over time and across assessors. *Inter-rater reliability* refers to whether different people administering the measurement tool can do so in a consistent way. *Test-retest reliability* tells us whether a measurement tool provides a consistent assessment of a skill, regardless of other factors such as a child's mood or health, the time of day, or the time of year that the child was assessed.

For screening tools, additional information is needed regarding how well these tools identify children who do indeed have a developmental delay (i.e., **sensitivity**), and how well they guard against misclassifying children as needing additional screening for a developmental delay who are, in fact, developing normally (i.e., **specificity**).

Researchers and professionals generally understand that not all children with or at risk for delays will be identified by a screener. While this understandably may raise questions, various circumstances, including the severity of the suspected delay, or the child's performance or mood on the day the screener is given, all affect the results. This is why ongoing observational assessment and opportunities for repeat screenings are essential.

---

[2] Based on the American Psychological Association's Standards for Educational and Psychological Testing, a *construct* is "the concept or characteristic that a test is designed to measure" (National Research Council of the National Academies, 2008, p. 186). A common method to determine construct validity is factor analysis, which sorts individual items into sets that fit together the best. Items that fit together should be measuring a single construct. Another approach to examining construct validity is to analyze the relationship between sets of items (i.e., scales) and characteristics of the child or family, such as child age or parent education, to determine whether the sets of items are related in expected ways to these child or family characteristics.

[3] A domain is a set of related skills, behaviors, or information that is classified as a single area of study or development. Domains typically cover multiple, related constructs within a broad area of study or development, such as fine motor development or approaches toward learning.

[4] Sometimes manuals refer to convergent criterion validity as concurrent validity, which could be interpreted to mean that the two measurement tools concur or "agree" in the measurement of a particular construct. However, another meaning of concurrent validity is that the two separate assessments were administered *at the same time* to measure criterion validity.

Of critical importance in understanding reliability and validity:

- **The reliability and validity of a screening or assessment instrument is dependent upon the purpose for which it is used.** Ongoing assessment aims to provide information on children's competencies or abilities over time. By comparing assessment results over time, children's developmental progress within or across domains can be monitored. This information can be used appropriately for such purposes as to guide instruction for individuals or groups of children, or to make decisions about program improvement efforts. Screening aims to identify children who need further evaluation to identify developmental delays. An instrument may provide reliable and valid information for the purpose of ongoing assessment, but be inaccurate at identifying children who may need further assessment or special services. Likewise, screening instruments are meant to identify children who need further evaluation to determine if there are delays in their development. Screening tools are rarely appropriate for assessing the developmental progress of children over time, since they cover only a limited range of development.

- **The reliability and validity of a screening or assessment instrument is dependent upon the population to whom it is given and the language in which it is administered.** It is important to know for whom a tool is reliable and valid. A tool may have been found to be reliable and valid for one group of children, but not others. For instance, its reliability may be established for children whose sole language is English, but not for dual language learners.

- **The reliability and validity of the information you get from assessment and screening instruments depend upon the instrument's implementation.** No matter how well-documented the reliability and validity of an assessment or screening tool, if an individual does not closely follow the training procedures outlined by the developer or if he or she alters the approach to implementing the assessment or screening tool, one cannot be confident that the information provided by the tool will be reliable or valid.

This document does not address every way that reliability and validity can be measured. We have chosen to report the methods for determining evidence of different forms of reliability and validity that were found in the majority of the assessment and developmental screener tools that were reviewed. Throughout the document, we introduce the different types of reliability and validity by identifying the question each type addresses. For example, the technical term "inter-rater reliability" addresses the question, "Do different raters agree when assessing the same children?" Similarly, the technical term "predictive validity" addresses the question, "How well does this assessment predict children's later academic achievement and adjustment to school?" By providing both the technical terms and the descriptive questions that are addressed, the profiles of the tools in this compendium strive to convey psychometric information in an accessible and easy-to-use format.

How to Use this Compendium

Now that we've provided an overview of reliability and validity, and its importance in selecting assessment and screening instruments, we can turn to the question of how to best use this document.

The compendium is composed of five parts:

1) **Introduction:** a review of the purpose of this compendium, the importance of reliability and validity of assessments and developmental screeners, and the organization and use of the compendium
2) **Summary Tables:** a set of tables summarizing common information from each of the assessment and screening tools examined
3) **Individual Instrument Profiles:** a set of profiles providing more detailed information for each of the assessment and screening tools reviewed
4) **Definition of Standards:** an overview of the standards used to evaluate the reliability and validity of the instruments (Appendix A)
5) **Glossary:** a glossary defining key terms used throughout this compendium (Appendix B)

Each piece of this compendium provides different information, and a manager might use the compendium differently depending upon his or her goals. Those who want to look across the most commonly used assessments for certain information–such as what domains are covered or how reliable the assessment is for children with special needs–would want to start with the summary tables. They might then choose a smaller set of instruments to examine in more detail by looking at the individual profiles for these instruments. In contrast, those who currently use one of the instruments included in the compendium and are interested in seeing detailed information on the reliability and validity of that instrument may want to turn directly to the individual profile for that instrument.

This compendium has been designed primarily to support programs that are in the process of selecting or changing their assessment or screening tools. However, programs should not interpret this compendium as recommending or requiring the use of any particular tools. For instance, if a program is satisfied with its existing tool, and knows that tool is a good complement for their curriculum and their population, the information in this compendium about particular tools may not be of interest. Likewise, a program that has recently changed its instrument may want to give staff time to get used to and implement that instrument before switching to another. It is our hope that this compendium can be helpful even to those programs not currently considering changing their assessment or screening instruments by highlighting the types of information on reliability and validity that the program managers should examine to determine the effectiveness of its assessment tool.

Guidance for Using and Interpreting the Summary Tables

The summary tables are intended to provide an overview of many different tools, either to narrow the range of tools to consider further or to provide a broad range of information for a set of tools. For both developmental screeners and assessment instruments, there are three summary tables. These three tables address the following:

- an overview of general information on the instrument, such as the age ranges covered, the languages in which the instrument is available, and whether the instrument is a direct assessment or observational tool;
- evidence of reliability and validity for the instrument, covering all types of reliability and validity, regardless of the population with which this information has been examined; and
- evidence of reliability and validity for particular populations of interest–dual language learners, children with special needs, and American Indian/Alaskan Native children.

Which Tools Are Included in the Compendium and Why?

Due to resource constraints, we could review only a limited number of assessment and screening tools. The original plan was to review 10 of each of the most commonly used assessment and screening instruments, with some flexibility to incorporate additional instruments that seemed to be "up-and-coming" or had exemplary reliability and validity. In order to identify these instruments, we reviewed reports from the 2008 Program Information Report (PIR) and the 2006 Family and Child Experiences Survey (FACES 2006) teacher survey regarding which assessments and developmental screening tools were being used by Head Start programs. This compendium is not intended to include all instruments used in Head Start programs; a particular instrument familiar to Head Start managers may be absent from this resource. Lack of information about a particular instrument in this compendium is not meant to reflect upon the quality of that tool.

It is important to note that this compendium only addresses tools for ongoing assessment and screening purposes that can be used with children in the 3- to 5-year-old age range. The review was limited to those assessment tools that cover a broad array of domains of the Head Start Child Outcomes Framework (HSCOF). We did not review instruments that only address one or two domains of the HSCOF (see http://www.hsnrc.org/CDI/pdfs/UGCOF.pdf).[5]

Among the tools that met these criteria, we looked for information about the tool's reliability and validity, as well as its availability for use by Head Start and other early childhood programs. If no information on reliability and validity was available, or the tool was not available for widespread use, we did not review it. As a result, two instruments that are among the most commonly used by Head Start programs (i.e., the Desired Results Developmental Profile –

[5] This compendium was completed prior to the December 2010 release of the revised framework, The Head Start Child Development and Learning Framework: Promoting Positive Outcomes in Early Childhood Programs Serving Children 3–5 Years Old. The revised version can be found at http://eclkc.ohs.acf.hhs.gov/hslc/ecdh/eecd/Assessment/Child%20Outcomes/HS_Revised_Child_Outcomes_Framework.pdf.

Revised, and the Hawaii Early Learning Profile) were not included in the compendium. See the section on "Tools under Development" for more information on these instruments.

For those instruments that were reviewed, we evaluated the reliability and validity information for all language versions available. So, for instance, if a separate Spanish version of the instrument was available, we looked for reliability and validity for that Spanish version. If no information was available on the reliability and validity of versions in languages other than English, we did not review those versions separately. In all but one instance, reliability and validity data were missing, so the Spanish-or other language versions-were not included in the review. The individual profile for each instrument notes languages for which the instrument is available and whether any data on reliability and validity are available for these languages.

The final list of instruments included in this review follows. Where possible, the most current version of a tool is profiled; minor variations on a tool are not included.

Ongoing Assessment Instruments

Creative Curriculum Developmental Assessment
Galileo Preschool Assessment Scales
High/Scope Child Observation Record
Learning Accomplishment Profile—3rd Edition
Learning Accomplishment Profile—Diagnostic
Learning Accomplishment Profile—Diagnostic, Spanish Edition
Mullen Scales of Early Learning
Work Sampling System

Developmental Screeners

Ages and Stages Questionnaire—3rd Edition
Ages and Stages Questionnaire—Social-Emotional
Battelle Developmental Inventory Screening Test
Brigance Preschool Screens
Denver II
Developmental Indicators for the Assessment of Early Learning—3rd Edition
Early Screening Inventory
Learning Accomplishment Profile—Diagnostic Screens
Parents' Evaluation of Developmental Status
Parents' Evaluation of Developmental Status: Developmental Milestones

The information included in each individual instrument profile was drawn from user's manuals, information available on the tools developers' websites, and information provided directly by the developer. The developer of each tool was asked to review the profile for accuracy and completeness. Profiles were updated and revised based on their input. Profiles that have not been reviewed by developers have an asterisk by the names of the tools in the Table of Contents and this information is also noted in the title on the profile.

For each child assessment and developmental screener tool within this compendium, the profiles summarize the following information:

- **Background Information**
- **Availability and Cost of Assessment**
- **Training and Other Requirements for Assessors**
- **Availability of an Information Reporting System**
- **Approaches to Parental/Family Input**
- **Appropriateness for Children from Different Backgrounds**
- **Reliability and Validity Information**
- **Sensitivity and Specificity Information (for screeners only)**
- **Availability of Guidance for Follow-up Actions (for screeners only)**
- **Comments from Review Authors**

Tools under Development

There were two tools (the Desired Results Developmental Profile – Revised and the Hawaii Early Learning Profile) that were not included in the compendium due to a lack of information on reliability and validity at the time this compendium was created. A third instrument (the Teaching Strategies GOLD assessment) was considered for inclusion, since it will be replacing the commonly used Creative Curriculum Developmental Continuum, but could not be included due to a lack of information on reliability and validity at the time of publication. Some brief information on these assessments follows:

The Desired Results Developmental Profile – Preschool is a child assessment that preschool teachers complete twice a year to measure children's progress toward the Desired Results for Children, a set of six goals for children and families used mainly in California. Results are summarized and shared with parents. Preschool teachers also use the results plan for individual and group instruction, analyze results, and continually monitor progress. More information can be found at http://www.cde.ca.gov/sp/cd/ci/drdpforms.asp or http://www.wested.org/desiredresults/training/form_drdp.htm.

The Hawaii Early Learning Profile (HELP) is a child assessment tool for children from birth through age 6 that covers development in the following domains: cognitive, language, gross motor, fine motor, social, and self-help. The revised version, HELP 3-6 Second Edition, was under development while this compendium was being developed. The HELP is an ongoing observational assessment that can be completed in a variety of settings and by teachers, parents, specialists, and others familiar with early childhood development. It is not a standardized measure and little technical information is available. One of the desired results informed by this measure is English Language Development. More information can be found at http://www.vort.com/.

The Teaching Strategies GOLD assessment system combines birth to 3 and 3- to 5- year-old instruments into one assessment tool. This system helps teachers observe, document, analyze, and evaluate children's progress as they progress through developmental steps, share this information with families, and plan instruction for individual children as well as the group. This measure may be used with any developmentally appropriate curriculum. The GOLD was developed by the same developer as the *Creative Curriculum Developmental Continuum*

Assessment System for Ages 3-5, and it will replace the *Creative Curriculum* assessment after the 2010-11 school year. Teaching Strategies, Inc. will cease to support The *Creative Curriculum Developmental Continuum Assessment System for Ages 3-5* in July 2011. More information can be found at http://www.teachingstrategies.com/page/GOLD.cfm.

References

National Research Council (2008). *Early childhood assessment: Why, what, and how*. Committee on Developmental Outcomes and Assessments for Young Children, C.E. Snow and S.B. Van Hemel, *Editors*. Board on Children, Youth, and Families, Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Peña, E. D. & Halle, T. (2011). Assessing preschool English learners: Traveling a multi-forked road. *Child Development Perspectives, 5 (1), 28-32.*

Spiker,D., Hebbeler, K. & Barton, L. (2011). Measuring quality of programs for children with disabilities. In M. Zaslow, I. Martinez-Beck, K. Tout, & T. Halle (Eds.), *Measuring quality in early childhood settings* (pp. 229-255). Baltimore, MD: Brookes Publishing.

Summary Tables

SUMMARY TABLE 1– General Information About Assessments

ASSESSMENT TITLE	Developmental Domains Covered (As listed by publisher)	Age Range	Observational or Direct Assessment	Norm-Referenced or Criterion-Referenced	Languages of Assessment Materials	Training Available Through Publisher or Developer	Technical training required to administer or score (over and above basic training on the assessment)	Scoring Options (Manual, Electronic)	Assessment Includes Parent and Family Input
Creative Curriculum Developmental Continuum Assessment	Social and Emotional Development Physical Development Cognitive Development Language Development	3-5 years	Observational	Norm-Referenced	English Spanish	Yes	No	Manual Electronic	Yes
High Scope Child Observation Record	Initiative Social Relations Creative Representation Movement & Music Language and Literacy Mathematics & Science	2-1/2 to 6 years old	Observational	Criterion-Referenced	English	Yes	No	Manual Electronic	Yes
Learning Accomplishment Profile- 3rd Edition (LAP-3)	Gross Motor Cognitive Fine Motor Language Pre-writing Self-help Cognitive Personal/Social	36 - 72 months	Observational and Direct	Criterion-Referenced	English	Yes	No	Manual Electronic	No
Learning Accomplishment Profile- Diagnostic (LAP-D)	Fine Motor Language Cognitive Gross Motor	30-72 months	Direct	Norm-Referenced	English	Yes	No	Manual Electronic	No
Learning Accomplishment Profile- Diagnostic (LAP-D) Spanish	Fine Motor Language Cognitive Gross Motor	30-72 months	Direct	Norm-Referenced	Spanish	Yes	No	Manual Electronic	No
Mullen Scales of Early Learning	Visual Reception Receptive Language Fine Motor Expressive Language Gross Motor	Birth - 68 months	Direct	Norm-Referenced	English	Yes	Yes. The measure is designed for use by clinical professionals.	Manual Electronic	No
Galileo Preschool Assessment Scales	Creative Arts Nature and Science Approaches to Learning Early Math Language and Literacy Physical Health Practices Fine and Gross Motor Development Social and Emotional Development	3-5 years	Observational and Direct	Norm-Referenced	English Spanish	Yes	No	Electronic	Yes
Work Sampling	Social and Emotional Development Approaches to Learning Creative Arts Language Development Literacy Mathematics Science Physical Health and Development	3-5 years	Observational	Criterion-Referenced	English	Yes	No	Manual Electronic	No

Note: For definitions and standards used to determine levels of evidence, see Appendices A and B.

SUMMARY TABLE 2– Assessments: Evidence of Reliability and Validity

	Reliability			Validity			
ASSESSMENT TITLE	**Inter-Rater Reliability** (Acceptable, Low/Weak, Not Examined)	**Test-Retest Reliability** (Acceptable, Low/Weak, Not Examined)	**Internal Consistency Reliability** (Acceptable, Low/Weak, Not Examined)	**Content Validity** (Content was reviewed by experts)	**Construct Validity** (Strong/High, Moderate, Weak/Low, Not Examined)	**Convergent Validity** (Strong/High, Moderate, Weak/Low, Not Examined)	**Predictive Validity** (Evidence of prediction to later achievement)
Creative Curriculum Developmental Continuum Assessment	**Not examined by the developer**	**Not examined by the developer**	**Acceptable**	**Yes, content was reviewed by experts**	**Strong**	**Not examined by the developer**	**Not examined by the developer**
High-Scope Child Observation Record (COR)	**Acceptable**	**Not examined by the developer**	**Acceptable**	**Not examined by the developer**	**Strong**	**Moderate**	**Not examined by the developer**
Learning Accomplishment Profile- 3rd Edition (LAP-3)	**Acceptable**	**Acceptable**	**Acceptable**	**Yes, content was reviewed by experts**	**Moderate**	**Moderate**	**Not examined by the developer**
Learning Accomplishment Profile- Diagnostic (LAP-D)	**Acceptable**	**Acceptable**	**Acceptable**	**Not examined by the developer**	**Strong**	**Strong**	**Not examined by the developer**
Learning Accomplishment Profile- Diagnostic (LAP-D) Spanish	**Acceptable**	**Acceptable**	**Acceptable**	**Not examined by the developer**	**Moderate**	**Strong**	**Not examined by the developer**
Mullen Scales of Early Learning	**Acceptable**	**Acceptable**	**Acceptable**	**Not examined by the developer**	**Strong**	**Strong**	**Yes, there is evidence of prediction to later achievement**
Galileo Preschool Assessment Scales	**Acceptable (only Language, Literacy and Math)**	**Acceptable**	**Acceptable**	**Yes, content was reviewed by experts**	**Strong**	**Not examined by the developer**	**Not examined by the developer**
Work Sampling	**Not examined by the developer**	**Acceptable (only Language, Literacy, and Math)**	**Acceptable (only Language, Literacy, and Math)**	**Yes, content was reviewed by experts**	**Strong**	**Moderate**	**Not examined by the developer**

Note: Ratings reported in this table reflect the majority finding when developers examined separate domains for the different types of reliability or validity. For example, if construct validity was examined for the cognitive, language, physical, and social domains and 3 of the 4 domains were found to have "strong" evidence of construct validity while the fourth domain was "moderate", the aspect was rated as having "strong" construct validity. See individual profiles for detailed findings. For definitions and standards used to determine levels of evidence, see Appendices A and B.

SUMMARY TABLE 3 – Assessments: Evidence of Reliability and Validity for Different Languages and Different Populations

	Reliability and Validity for Different Languages		Reliability and Validity for Different Populations		
ASSESSMENT TITLE	Evidence of Reliability and Validity in English?	Evidence of Reliability and Validity in Other Languages?	Evidence of Reliability and Validity for Dual Language Learners?	Evidence of Reliability and Validity for Children with Special Needs?	Evidence of Reliability and Validity for American Indian/Alaskan Native Children?
Creative Curriculum Developmental Continuum Assessment	Yes	No	No evidence[1]	No evidence[1]	No evidence[2]
High-Scope Child Observation Record (COR)	Yes	No	No evidence[1]	No evidence[1]	No evidence[1]
Learning Accomplishment Profile- 3rd Edition (LAP-3)	Yes	No	No evidence[1]	No evidence[2]	No evidence[2]
Learning Accomplishment Profile-Diagnostic (LAP-D)	Yes	N/A	No evidence[1]	No evidence[2]	No evidence[3]
Learning Accomplishment Profile-Diagnostic (LAP-D) Spanish	N/A	Yes (Spanish)	No evidence[1]	No evidence[2]	No evidence[1]
Mullen Scales of Early Learning	Yes	No	No evidence[1]	Yes	No evidence[1]
Galileo Preschool Assessment Scales	Yes	No	No evidence[1]	No evidence[1]	No evidence[2]
Work Sampling	Yes	No	No evidence[1]	No evidence[2]	No evidence[1]

KEY

YES: At least one measure of acceptable reliability or validity is presented by the developer.

NO: The developer did not examine whether the measure was reliable or valid for this population.

[1] No information about this population is provided by the developer.

[2] While this population was included in the total sample of children, separate analyses for this sub-group were not conducted by the developer.

[3] This population was NOT included in the sample of children examined by the developer

N/A = Not Applicable

For definitions and standards used to determine levels of evidence, see Appendices A and B.

SUMMARY TABLE 4 – General Information About Developmental Screeners

DEVELOPMENTAL SCREENER TITLE	Developmental Domains Covered (As listed by publisher)	Age Range	Languages of Developmental Screener Materials	Training Available Through Publisher or Developer	Must Be Administered by Someone with Technical Background	Scoring Options (Manual, Electronic)	Screener Includes Parent and Family Input	Screener Includes Guidance on Follow-Up Steps
Denver II	Personal-Social Language Fine Motor-Adaptive Gross Motor	0 months to 6 years	English Spanish	Yes	No	Manual Electronic	Yes	Yes
Ages and Stages Questionnaire (ASQ-3)	Communication Problem Solving Gross Motor Fine Motor Personal-Social	1 - 66 months	English Spanish	Yes	No	Manual Electronic	Yes	Yes
Ages and Stages Questionnaire: Social-Emotional	Self-regulation Adaptive functioning Compliance Autonomy Communication Interaction with people Affect	6 - 60 months	English Spanish	Yes	No	Manual Electronic	Yes	Yes
Battelle Developmental Inventory Screening Test	Adaptive Cognitive Personal-Social Motor Communication	Birth through age 7	English Spanish	Yes	No	Manual Electronic	No	Yes
Brigance Preschool Screen	Academics/pre-academics Expressive language Gross motor Receptive language Fine motor Social-emotional skills Self-help	Birth through 1st grade	English	Yes	No	Manual Electronic	Yes	Yes
DIAL-3 (Developmental Indicators for the Assessment of Learning)	Motor Self-Help Development Concepts Social Development Language	3 years to 6 years, 11 months	English Spanish	Yes	No	Manual Electronic	Yes	Yes
Early Screening Inventory [ESI-R]	Visual-Motor/Adaptive Language and Cognition Gross Motor	ESI-P: 3 years to 4 years, 5 months ESI-K: 4 years, 6 months to 5 years, 11 months	English Spanish	Yes	No	Manual Electronic	Yes	Yes
Learning Accomplishment Profile-Diagnostic Screens	Gross Motor Cognitive Fine Motor Language	3-5 years	English Spanish	Yes	No	Manual Electronic	No	No
Parents' Evaluation of Developmental Status	Global/Cognitive Fine Motor Expressive Language and Articulation Receptive Language Gross Motor Behavior Social-Emotional Self-Help School	Birth through age 8	English (Forms also translated into 14 other languages.)	Yes	No	Manual Electronic	Yes	Yes
Parents' Evaluation of Developmental Status-Developmental Milestones	Expressive Language Social-Emotional Receptive Language Self-Help Fine-Motor Gross Motor Academic: Pre-Reading; Pre-Math, and Written Language	Birth through age 7 years, 11 months	English Spanish	Yes	No	Manual Electronic	Yes	Yes

Note: For definitions and standards used to determine levels of evidence, see Appendices A and B.

SUMMARY TABLE 5 – Screeners: Evidence of Reliability and Validity

	Reliability			Validity				
SCREENER TITLE	**Inter-Rater Reliability** (Acceptable, Low/Weak, Not Examined)	**Test-Retest Reliability** (Acceptable, Low/Weak, Not Examined)	**Internal Consistency Reliability** (Acceptable, Low/Weak, Not Examined)	**Content Validity** (Content was reviewed by experts)	**Construct Validity** (Strong/High, Moderate, Low/Weak, Not Examined)	**Convergent Validity** (Acceptable, Low/Weak, Not Examined)	**Sensitivity*** (High, Moderate, Low)	**Specificity*** (High, Moderate, Low)
Denver II	**Acceptable**	**Acceptable**	**Not examined by the developer**	**Not examined by the developer**	**Not examined by the developer**	**Not examined by the developer**	**Moderate**	**Low**
Ages and Stages Questionnaire (ASQ-3)	**Acceptable**	**Acceptable**	**Acceptable**	**Yes, content was reviewed by experts**	**Not examined by the developer**	**Moderate**	**Moderate**	**Moderate**
Ages and Stages Questionnaire: Social-Emotional	**Not examined by the developer**	**Acceptable**	**Acceptable**	**Yes, content was reviewed by experts**	**Not examined by the developer**	**Strong**	**Moderate**	**High**
Battelle Developmental Inventory Screening Test	**Not examined by the developer**	**Not examined by the developer**	**Acceptable**	**Yes, content was reviewed by experts**	**Not examined by the developer**	**Not examined by the developer**	**Moderate**	**Moderate**
Brigance Preschool Screen	**Acceptable**	**Acceptable**	**Acceptable**	**Yes, content was reviewed by experts**	**Strong**	**Strong**	**Moderate for 3-4 year olds High for 5 year olds**	**Moderate**
DIAL-3 (Developmental Indicators for the Assessment of Learning)	**Not examined by the developer**	**Acceptable**	**Not examined by the developer**	**Not examined by the developer**	**Not examined by the developer**	**Strong**	**Low**	**High**
Early Screening Inventory [ESI-R]	**Acceptable**	**Acceptable**	**Not examined by the developer**	**Not examined by the developer**	**Not examined by the developer**	**Not examined by the developer**	**High**	**Moderate**
Learning Accomplishment Profile-Diagnostic Screens	**Acceptable**	**Acceptable**	**Acceptable**	**Yes, content was reviewed by experts**	**Not examined by the developer**	**Strong**	**Moderate**	**Not examined by the developer**
Parents' Evaluation of Developmental Status	**Acceptable**	**Acceptable**	**Acceptable**	**Yes, content was reviewed by experts**	**Not examined by the developer**	**Strong**	**Moderate**	**Moderate**
Parents' Evaluation of Developmental Status-Developmental Milestones	**Acceptable**	**Acceptable**	**Acceptable**	**Yes, content was reviewed by experts**	**Not examined by the developer**	**Strong**	**Moderate**	**Moderate**

* Sensitivity and specificity refer to the accuracy with the measures identifying children at-risk for developmental problems.

Note: Ratings reported in this table reflect the majority finding when developers examined separate domains for the different types of reliability or validity. For example, if construct validity was examined for the cognitive, language, physical, and social domains, and 3 of the 4 domains were found to have "strong" evidence of construct validity while the fourth domain was "moderate", the aspect was rated as "strong" overall. See individual profiles for detailed findings. For definitions and standards used to determine levels of evidence, see Appendices A and B.

SUMMARY TABLE 6 – Developmental Screeners: Evidence of Reliability and Validity for Different Languages and Different Populations

	Reliability and Validity for Different Languages		Reliability and Validity for Different Populations		
DEVELOPMENTAL SCREENER TITLE	Evidence of Reliability and Validity in English?	Evidence of Reliability and Validity in Other Languages?	Evidence of Reliability and Validity for Dual Language Learners?	Evidence of Reliability and Validity for Children with Special Needs?	Evidence of Reliability and Validity for American Indian/Alaskan Native Children?
Denver II	Yes	No	No evidence[1]	Yes	No evidence[2]
Ages and Stages Questionnaire (ASQ-3)	Yes	No	No evidence[1]	Yes	No evidence[2]
Ages and Stages Questionnaire: Social-Emotional	Yes	No	No evidence[1]	No evidence[1]	No evidence[2]
Battelle Developmental Inventory Screening Test	Yes	No	No evidence[1]	Yes	No evidence[2]
Brigance Preschool Screen	Yes	No	No evidence[1]	No evidence[1]	No evidence[2]
DIAL-3 (Developmental Indicators for the Assessment of Learning)	Yes	No	No evidence[1]	Yes	No evidence[1]
Early Screening Inventory-Revised [ESI-R]	Yes	No	No evidence[1]	No evidence[1]	No evidence[1]
Learning Accomplishment Profile-Diagnostic Screens	Yes	No	No evidence[1]	No evidence[1]	No evidence[2]
Parents' Evaluation of Developmental Status	Yes	No	No evidence[1]	No evidence[2]	No evidence[1]
Parents' Evaluation of Developmental Status-Developmental Milestones	Yes	No	No evidence[1]	Yes	No evidence[2]

KEY

YES: At least one measure of acceptable reliability or validity is presented by the developer.
NO: The developer did not examine whether the measure was reliable or valid for this population.
[1] No information about this population is provided by the developer.
[2] While this population was included in the total sample of children, separate analyses for this sub-group were not conducted by the developer.
[3] This population was NOT included in the sample of children examined by the developer

For definitions and standards used to determine levels of evidence, see Appendices A and B.

Assessment Profiles

The Creative Curriculum for Preschool Developmental Continuum Assessment System for Ages 3-5

Developers: Diane Trister Dodge, Laura Colker, and Cate Heroman
Publisher: Teaching Strategies, Inc. http://www.creativecurriculum.net/

Purpose:
The Creative Curriculum for Preschool Developmental Continuum Assessment System for Ages 3-5 (Developmental Continuum) is an assessment system that reflects the curricular goals and objectives of The Creative Curriculum for Preschool.

Developmental domains addressed in the assessment, as stated by the publisher:

- Social and emotional development
  - o Sense of self
  - o Responsibility for self and others
  - o Prosocial behavior
- Physical development
  - o Gross motor
  - o Fine motor
- Cognitive development
  - o Learning and problem solving
  - o Logical thinking
  - o Representation and symbolic thinking
- Language development
  - o Listening and speaking
  - o Reading and writing

Intended age range:
3-5 years

Number of items:
The full assessment contains 50 items.

Background

The Creative Curriculum Development Continuum Assessment system helps teachers observe, document, analyze, and evaluate children's progress as they move through developmental steps; share this information with families; and plan instruction for individual children as well as groups. This measure may be used independently of the curriculum.

Is the assessment a direct assessment or an ongoing observational tool?

The Developmental Continuum is an ongoing observational tool that includes portfolio and anecdotal record keeping throughout the year.

Language(s) developed for:

The assessment was developed for English-speaking children, but some test administration materials are also available in Spanish.

What is the appropriate time period between administering, recording, or reviewing the data?

The manual recommends ongoing collection of information and reviewing information about where each child is on each objective at three points during the year.

Availability and Cost of the Assessment

Is the assessment available to programs without restrictions?

Yes, it is available without restrictions.

What is the cost of the assessment?

As of 2010, The Creative Curriculum Developmental Continuum Assessment Toolkit for Ages 3-5 costs $114.95. It includes materials and forms needed to assess up to 25 children at three summary checkpoints each year and share information with families. Included are forms for individual children, class summary forms, and report forms. The Toolkit may be used on its own or with a software reporting tool, CC-PORT (The Creative Curriculum Progress and Outcomes Reporting Tool).

Training and Other Requirements for Assessors

Is training available on how to administer and score the assessment? Who offers the training?

Yes, customized, onsite training on how to administer and score the *Developmental Continuum* is available through Teaching Strategies, Inc. However, Teaching Strategies, Inc. will cease to support *The Creative Curriculum Developmental Continuum Assessment System for Ages 3-5* in July 2011. A new assessment called *Teaching Strategies GOLD™* will be replacing *The Creative Curriculum Developmental Continuum Assessment System for Ages 3-5* after the 2010 school year.

Is it necessary to have a professional background or technical training over and above training on the assessment to administer or complete the assessment?

No, it is not necessary to have a technical background or training to complete the assessment. Teachers can complete the assessment. The manual that accompanies the assessment includes detailed instructions on how to administer this assessment system.

Is it necessary to have a professional background or technical training over and above training on the assessment to score the assessment?

No, a teacher can score the assessment without technical background or training. The manual includes detailed instructions on how to compile and evaluate the information collected using this assessment system.

Are regular checks on administration required or recommended to ensure appropriate administration? If so, when and by whom?

No, they are not required.

Information Reporting System for the Assessment

Electronic Data Entry. *Does the assessment come with a process for entering information from the assessment electronically?*

Yes, two tools are available to enter information electronically. CreativeCurriculum.net is the online version of the *Developmental Continuum* and is sold as an annual license on a per-child basis. Pricing is based on the number of children included in the subscription and currently ranges from $13.95-$19.95 per child.

Teaching Strategies also offers *The Creative Curriculum Progress and Outcomes Reporting Tool* (*CC-PORT™ 2.0*) software for $99.95 (as of publication date).

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

CreativeCurriculum.net allows the user to generate planning and progress reports, based on child observations entered into the system, at the individual, classroom, or program level.

CC-PORT™ 2.0 is a software program that creates progress reports about groups of preschool children based on information collected through *The Creative Curriculum Developmental Continuum Assessment System for Ages 3-5*. After completing assessments of children for each class, the data are entered into *CC-PORT*; compiled; and aggregated into two reports (complete with charts and graphs) on groups of children at the class, program, and agency level.

The reports show overall progress in the four developmental areas of *The Creative Curriculum for Preschool* and each of the Head Start general domain areas and the required domain elements and indicators. Data can be consolidated and reports produced at the classroom, center, program, delegate agency, and grantee level.

Child Outcomes Framework. *Does the publisher map the domains in the assessment onto the domains in the Head Start Child Outcomes Framework (HSCOF)?*

Yes, the publisher maps the domains of this assessment onto the domains in the HSCOF.

Instructional Support. *Does the assessment provide guidance about follow-up steps teachers can use to help individual children progress?*

Yes, the manual explains how determining a child's position on the *Developmental Continuum* can inform planning to support individual child's development. . The manual also discusses how teachers can use the assessment results in parent conferences or meetings. It does not give specific activities for teachers to use to help children progress.

Planning Support. *Does the assessment come with guidance to help teachers determine follow-up steps involving planning at the classroom level?*

Yes, the assessment provides a class summary sheet and the manual provides guidance on interpreting the summary to help plan classroom activities. For example, for skills that many children in the class are still developing, the manual suggests incorporating different learning materials and activities into the program.

Administrator Support. *Does the assessment come with guidance to help program administrators plan follow-up steps involving program improvement?*

The manual does not provide guidance for administrators. However, specific guidance for program administrators is available for programs using either CreativeCurriculum.net or *CC-PORT™* and can be found online.

Approaches to Family/Parent Input

Tools for Family Input. *Does the assessment include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

Yes, as part of conferences, families are invited to share their observations of their child. This information is recorded on the Child Progress and Planning Report, along with a summary of the teacher's observations.

Sharing Results. *Does the assessment include recommendations on how to share assessment results with a child's family?*

Yes, sharing this information with families is considered one of the main purposes of this assessment. The manual provides guidance on how to compile and present assessment information to parents using the Child Progress and Planning Report. This report summarizes development in each of the four domains (social/emotional, physical, cognitive, and language) by providing examples of a child's skills in each area. The form also has blank fields for teachers and families to complete together (Family Comments & Observations, and Next Steps at School and at Home).

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this an assessment with developmental norms?*

Yes, the current version of the *Developmental Continuum* is an assessment with developmental norms from 2003.

Which populations were included in the norming sample?

These norms were developed with 3- to 5-year-old children in selected Head Start programs that satisfied the following criteria:

- Adopted the 4th edition of *The Creative Curriculum for Preschool*
- Received *The Creative Curriculum for Preschool* training
- Demonstrated a reasonably complete implementation of *The Creative Curriculum for Preschool*
- Used *The Creative Curriculum Developmental Continuum Assessment System for Ages 3-5* as suggested
- Recorded *The Creative Curriculum Developmental Continuum Assessment System for Ages 3-5* data in an electronic format

Twenty-five program sites were selected and assessment data were collected for 1,590 children during the winter checkpoint. The table on the next page provides more information about these children.

Availability of Versions in Languages Other than English. *Is the assessment available in languages other than English? Which languages?*

Yes, the measure has been translated into Spanish, but there is no separate research on the reliability and validity of the Spanish version.

How were versions in languages other than English developed?

The developers did not provide information about how the Spanish version was developed.

Is there any evidence that versions in languages other than English were developed with a representative group of children speaking that language, either as monolinguals or bilinguals?

No, the developers did not provide information about the translation of the Spanish version.

Options for Use with Special and Diverse Populations

What are the findings on the reliability and validity of versions of the assessment in languages other than English?

There are no findings on the reliability and validity of the Spanish version.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

The *Developmental Continuum* provides descriptions of different skill levels for each objective. The manual states that there are children who may go beyond the scope of this developmental continuum and others whose development is not at a typical level. For children who are developmentally delayed, the manual identifies forerunner skills. These are examples only, and children may present wide variation in how their development relates to certain objectives.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted with diverse populations to determine the appropriateness of this assessment for these populations?*

The developers did not examine appropriateness for diverse populations in this way.

Characteristics of 2003 Norming Sample
Number of children in the sample: 1,590

	Percentage of Children
Age of Children	
3 years old	25.09
4 years old	55.79
5 years old	19.12
Gender	
Male	51.4
Female	48.6
Racial/Ethnic Group	
White	33.3
African American	30.2
Hispanic Origin	29.9
Other (includes Native American)	6.6
Head Start Region	
I	0.0
II	11.7
III	13.9
IV	18.6
V	27.0
VI	0.0
VII	2.8
VIII	0.8
IX	25.2
X	0.0
Primary Language in the Home	
English	76.6
Spanish	20.3
Other	3.1

Reliability and Validity Information

What is known about the reliability and validity of the assessment…

In English?

There is information about reliability and validity of the assessment in English. This information is outlined in later sections of this profile.

In other languages?

The developers did not examine the reliability and validity for children assessed using the Spanish version of the measure.

For dual language learners?

Children with a primary language other than English in the home are included in the sample (20.3% Spanish and 3.1% Other). However, the developers have not examined the reliability and validity of the assessment for this population separately.

For children with special needs?

The developers do not provide information about children with special needs and have not examined the reliability and validity for this population.

For American Indian/Alaskan Native children?

While American Indian/Alaskan Native children were included in the sample of children (1%) on which the assessment was tested, the reliability and validity for this group have not been examined separately.

For children of migrant and seasonal farm workers?

The developers do not provide information about children of migrant and seasonal farm workers and have not examined the reliability and validity for this population.

Reliability: Does the assessment obtain the same results, consistently, under the same conditions with the same children?

Inter-Rater Reliability. *Do different raters agree when they are assessing the same children?*

The developers did not examine the agreement between different raters assessing the same children.

Test-Retest Reliability. *How consistent are scores if the assessment is administered once and then administered again soon?*

The developers did not examine the consistency of scores between administrations.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

Items intended to reflect the same skills or behaviors have acceptably strong interrelationships. For example, all of the items within the cognitive domain were strongly related to one another. This was examined with the study sample of 1,590 children.

Validity: Does the assessment measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the assessment do a good job of reflecting what the assessment is supposed to be assessing?*

Yes, 40 experts were consulted during the development of this assessment.

Construct Validity. *How closely related to each other are sets of items within the assessment that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

Sets of items that are intended to address similar skills and behaviors are moderately to strongly related to each other. Of the 50 items in this tool, three items were related across two to three different domains as they measured several areas of development at the same time. These three items were:

- (#21) "Uses tools for writing and drawing," which related to items that measure physical, cognitive, and language-related development
- (#24) "Shows persistence in approaching tasks," which related to items that measure social and cognitive development
- (#40) "Understands and follows oral directions," which related to items that measure social, cognitive, and language-related development

The remaining 47 items were organized into four distinct developmental domains: social/emotional, physical, cognitive, and expressive language. Information about how scores on sets of items relate to children's age as expected is not provided.

Convergent Validity. *Is this assessment closely related to other well-established assessments aimed at measuring the same skills and behaviors?*

The developers did not examine the relationship of this assessment to other well-established assessments aimed at measuring the same skills and behaviors.

Predictive Validity. *How well does this assessment predict children's later academic achievement and adjustment to school? For what groups of children has this been examined?*

The developers did not examine the relationship between scores and later development.

Comments

With regard to this measure's construct validity, the developer elected to group the expressive language items within the *Developmental Continuum's* Language domain with eight emergent literacy items that were found to relate more closely to the cognitive domain.

A new assessment called *Teaching Strategies GOLD™* will be replacing *The Creative Curriculum Developmental Continuum Assessment System for Ages 3-5* after the 2010-11 school year. Teaching Strategies, Inc. will cease to support *The Creative Curriculum Developmental Continuum Assessment System for Ages 3-5* in July 2011. For more information, see http://www.teachingstrategies.com/page/GOLD.cfm.

References

Dodge, D.T., Colker, L., and Heroman, C. (2001). *A Teacher's Guide to Using the Creative Curriculum Developmental Continuum Assessment System.* Teaching Strategies, Inc.

Lambert, R.G. (2003). *The Creative Curriculum Developmental Continuum for Ages 3 - 5: The Assessment Component of the Creative Curriculum for Preschool Technical Report.* University of North Carolina.

The Galileo Pre-K Online System for the Electronic Management of Learning

Developer: Assessment Technology Incorporated
Publisher: Assessment Technology Incorporated

http://www.ati-online.com/

Purpose:
The purpose of the Galileo Pre-K Online web-based software application is to support a set of management processes aimed at promoting learning.

Developmental domains addressed in the assessment, as stated by the publisher:

- Creative arts
- Approaches to learning
- Early math
- Language and literacy
- Nature and science
- Physical health practices
- Fine and gross motor development
- Social and emotional development

Intended age range:
Galileo System:
Infancy-5 years

Preschool Assessment (in this profile):
3-5 years

Number of items:
The Preschool Assessment contains 481 developmental capabilities or items.

Background

The Preschool Assessment Scales are one part of a larger Galileo Instructional Improvement System comprising several components, including curriculum, lesson planning, assessments, and reporting. This profile focuses on the Preschool Assessment Scales, including goal setting, planning, implementation, and evaluation.

Is the assessment a direct assessment or an ongoing observational tool?

The Galileo system has both an ongoing observational and a direct assessment component.

If the assessment is a direct assessment, how long does it take to administer the measure?

The developer does not provide information regarding the time required to administer the measure.

Language(s) developed for:

English and Spanish

What is the appropriate time period between administering, recording, or reviewing the data?

Since the preschool assessment scales are ongoing, the scores are updated as the teacher sees that a child has mastered a skill rather than administering the measure at given dates. Typically, early childhood programs using Galileo will use the program's reports to record data on children from three to six times a year.

Availability and Cost of the Assessment

Is the assessment available to programs without restrictions?

Yes, the measure is available to programs without restrictions.

What is the cost of the assessment?

As of 2010, Galileo Pre-K Online is provided to programs through an annual renewable license fee. The cost information is not available online; for more information, contact the publisher.

Training and Other Requirements for Assessors

Is training available on how to administer and score the assessment? Who offers the training?

The Galileo website details several different trainings for different parts of the system. Costs for these trainings vary based upon method of delivery, such as whether training is done on-site or online.

Is it necessary to have a professional background or technical training over and above training on the assessment to administer or complete the assessment?

No, it is not necessary to have a professional background or technical training over and above training on Galileo to administer or complete the assessment.

Is it necessary to have a professional background or technical training over and above training on the assessment to score the assessment?

No, it is not necessary to have a professional background or technical training to score the assessment. Galileo is designed for ease of use by early childhood teachers, specialists, and administrators. In addition, the publisher provides comprehensive professional development and ongoing technical support for programs implementing Galileo Pre-K Online. Many of the scoring and reporting functions in Galileo occur automatically and do not require the teacher to do additional work.

Are regular checks on administration required or recommended to ensure appropriate administration? If so, when and by whom?

Yes, staff and specialists should build in time during the year to review how to administer the assessment, including how to collect information about a child from multiple sources.

Information Reporting System for the Assessment

Electronic Data Entry. *Does the assessment come with a process for entering information from the assessment electronically?*

Yes, information on children's learning can only be entered directly into the Galileo Pre-K Online web-based application.

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

The Galileo Online Reporter is an online tool used to manage data from the child assessments. Specifically, there is a developmental profile report, which shows progress on all goals set for a specific child, and a developmental milestone report, which shows progress on a specific subset of goals. Reports can be generated at the level of the individual child, classroom, center, and for higher levels if applicable, such as a group of centers or schools. The reports also can be organized according to demographic information such as gender, special needs, who the child lives with, and primary language spoken. Additionally, achievement-level reports can be generated to show the number of children who have reached a certain goal set by a teacher, director, or other administrator. Online reports also can help teachers set goals for children and plan lessons based on children's progress and development.

Child Outcomes Framework. *Does the publisher map the domains in the assessment onto the domains in the Head Start Child Outcomes Framework (HSCOF)?*

Yes, the publisher maps the domains of the Galileo Preschool Assessment Scales onto the HSCOF.

Instructional Support. *Does the assessment provide guidance about follow-up steps teachers can use to help individual children progress?*

Yes, the measure does have guidance about follow-up steps teachers can take to help individuals progress as well as information on planning at the classroom level.

Based on the assessment information, the Galileo system can recommend activities that are specifically targeted toward the areas of development where a child needs to improve. Specifically, the activities are broken down into different levels of "plan now," "plan soon," and "plan later," depending on when a child should be progressing on specific developmental skills. These levels provide guidance so the teacher knows in what order the activities should occur to promote development in a certain area or on a certain task. The Galileo system has libraries of activities that are linked to the assessment areas so that appropriate activities can be selected. Further, there is also a tool for teachers to use to develop their own activities and add them into their library, all of which is done electronically.

Planning Support. *Does the assessment come with guidance that helps teachers determine follow-up steps involving planning at the classroom level?*

All of the methods mentioned for helping individual children can be applied to the classroom level. Since the electronic system is capable of reporting data by classroom, a teacher can look at trends at the classroom level and determine which areas of development should be a focus for the near future.

Administrator Support. *Does the assessment come with guidance to help program administrators plan follow-up steps involving program improvement?*

Yes, the electronic system provides reports of how all children are doing in a program, and also has the capability to produce reports based a variety of other characteristics such as gender and age. The system also provides guidance to help administrators plan activities (including goal setting and planning) tailored to children's needs.

Approaches to Family/Parent Input

Tools for Family Input. *Does the assessment include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

Yes, the Galileo system allows for parent/family input on a child's development. Teachers can record information from the child's family in the electronic system. For example, if a parent reports that a child identified 10 letters correctly at home, the teacher can indicate that the parent saw this, but that the teacher has not yet witnessed this behavior.

Sharing Results. *Does the assessment include recommendations on how to share assessment results with a child's family?*

Yes, the measure includes recommendations on how to share assessment results with family. One of the electronically generated reports, called the Developmental Milestone Report, shows how a child is progressing within a certain developmental area. A teacher can share this with parents and suggest activities to be done at home to strengthen areas where a child may be progressing slowly. There is also an Online Parent Center within the electronic system where parents can log in and access up-to-date information about their child's learning and classroom experiences. They can see what children will be doing at preschool by viewing a teacher's lesson plan. Parents can also browse activity libraries to see the activities the child will be doing. In addition, the Times for Learning Library within the Activity Library houses numerous activities that parents can use with their children at home.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this an assessment with developmental norms?*

Yes, Galileo has developmental norms.

Which populations were included in the norming sample?

There is no information about the characteristics of the sample on which the norms are based.

Availability of Versions in Languages Other than English. *Is the assessment available in languages other than English? Which languages?*

The measure is available in Spanish.

How were versions in languages other than English developed?

The Florida Head Start Association Research Committee translated the Galileo 3-5 Assessment Scales into Spanish.

Is there any evidence that versions in languages other than English were developed with a representative group of children speaking that language, either as monolinguals or bilinguals?

There is no separate reliability or validity information regarding the use of the assessment with Spanish-speaking children.

What are the findings on the reliability and validity of versions of the assessment in languages other than English?

There is no separate reliability or validity information regarding the use of the assessment with Spanish-speaking children.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

While the developer states that the assessment can be used with children who are developing typically or atypically, it provides no information about accommodations.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted with diverse populations to determine the appropriateness of this assessment for these populations?*

No, the developer has not examined the appropriateness of this assessment for diverse populations in this way.

Reliability and Validity Information

What is known about the reliability and validity of the assessment…

In English?

To examine the reliability and validity of the assessment in English, the developer tested two groups of children, one in 1998 and one in 2001. In 1998, the developer examined some aspects of reliability and validity with 2,528 children in early childhood programs in Ohio. Children ranged from 3 to 6 years. Thirty-six percent of the children were African American, 50% were Caucasian, 8% were Hispanic, 1% were Asian, 5% were biracial, fewer than 1% were American Indian, and fewer than 1% were classified as other. In 2001, the sample included 3,092 children from early childhood programs in Florida, Indiana, Kentucky, Ohio, Oregon, Tennessee, and Texas. Children ranged in age from 3 to almost 6 years. Fifty-two percent of the children were male. The racial/ethnic makeup of the group was 43% African American, 40% Caucasian, and 17% Hispanic. The reliability and validity findings for this sample are outlined in later sections of this profile.

In other languages?

The developer has not examined the reliability and validity for children assessed using the Spanish version of the measure.

For dual language learners?

The developer does not provide information about dual language learners and has not examined the reliability and validity of the assessment for this population.

For children with special needs?

The developer does not provide information about children with special needs and has not examined the reliability and validity of this assessment for this population.

For American Indian/Alaskan Native children?

While American Indian/Alaskan Native children were included in the 1998 sample (less than 1%), the developer has not examined reliability and validity of this assessment when used with American Indian/Alaskan Native children.

For children of migrant and seasonal farm workers?

The developer does not provide information about children of migrant and seasonal farm workers and has not examined the reliability and validity for this population.

Reliability: Does the assessment obtain the same results, consistently, under the same conditions with the same children?

Inter-Rater Reliability. *Do different raters agree when they are assessing the same children?*

Yes, for three of the subscales, different raters agree when they are assessing the same children. In order to test this, two observers assessed children on the early math and language and literacy Scales. The two observers agreed more than 80% of the time, so the relationship between the two scores is acceptable. This was examined with some of the children in the 1998 sample of children described earlier. The children were from three Head Start programs in Ohio. Three children were chosen from each of the 318 classrooms in the Head Start programs. After they were trained, the lead teacher and the assistant teacher both assessed the same children and then the scores were compared.

Test-Retest Reliability. *How consistent are scores if the assessment is administered once and then administered again soon?*

Scores on all four of the LAP-D domains met the criteria for acceptable consistency when the assessment was administered twice. The agreement on scores was acceptable on all the scales, but agreement was a little lower on the language and literacy scales. The developer examined this with the 2001 sample of children described above. The developers do not provide further information about the teachers.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

For all eight of the domains or categories in the preschool assessment scales, items that are meant to reflect the same set of skills or behaviors as other items meet the criteria for acceptable relationships. This was examined with the 2001 sample of children described above.

Validity: Does the assessment measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the assessment do a good job of reflecting what the assessment is supposed to be assessing?*

Yes, the developer consulted experts when it was working to develop the preschool assessment scales. The developer spoke with educators and parents, and looked at research in the child development field to get input about what to include in the measure.

Construct Validity. *How closely related to each other are sets of items within the assessment that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

Sets of items in the Galileo scales intended to address similar skills and behaviors are strongly related to each other. Information about how scores on sets of items relate to children's age as expected is not provided.

Convergent Validity. *Is this assessment closely related to other well-established assessments aimed at measuring the same skills and behaviors?*

The developer has not examined this information for the Galileo Preschool Assessment Scales.

Predictive Validity. *How well does this assessment predict children's later academic achievement and adjustment to school?*

The developer has not examined this information for the Galileo preschool assessment scales.

Comments

While Galileo has not been tested for reliability and validity for children with special needs, it does feature capabilities that permit users to sort data by different variables (such as children with special needs, program types, and English language learners). Assessment Technology Incorporated (the program's developer) designed its ongoing research program to document the performance of diverse groups of children, including those served by programs addressing the Office of Special Education Programs requirements to satisfy the 2004 Individuals with Disabilities Education Improvement Act, Part B.

References

Bergan, J. Richard, Burnham, C.G., Feld, J.K., & Bergan, J. Robert. (2009). *The Galileo Pre-K Online System for the Electronic Management of Learning: Technical Manual*. Assessment Technology Incorporated. Retrieved September 30, 2009, from http://www.ati-online.com/galileoPreschool/PrePublications.html

Feld, J.K. (2006). *Administrative decision making in early childhood education: How do I use the reports available in online reporter?* Assessment Technology Incorporated. Retrieved June 15, 2009, from www.ati-online.com

High/Scope Preschool Child Observation Record (COR)

Developer: High/Scope Foundation
Publisher: High/Scope Foundation

http://www.highscope.org/Content.asp?ContentId=113

Purpose:

The High/Scope Preschool Child Observation Record (COR) is an observation-based instrument providing systematic assessment of young children's knowledge and abilities in all areas of development.

Developmental domains addressed in the assessment, as stated by the publisher:

- Initiative
- Social relations
- Creative representation
- Movement and music
- Language and literacy
- Mathematics and science

Intended age range:
2 1/2 - 6 years

Number of items:
32

Background

The High/Scope Preschool Child Observation Record (COR) may be used in any developmentally oriented program and is not limited to use with other High/Scope materials.

Is the assessment a direct assessment or an ongoing observational tool?

The COR is an ongoing observational tool.

Language(s) developed for:

The manual and other materials were developed in English; however, the child information and developmental summary and the OnlineCOR include options for scoring anecdotes when the child's primary language is other than English. The parent guide and parent report form are available in Spanish. Translations of the materials may be available from other countries such as Canada (French), Taiwan (Chinese), and Korea (Korean).

Availability and Cost of the Assessment

Is the assessment available to programs without restrictions?

Yes, the measure is available without restrictions.

What is the cost of the assessment?

As of 2010, the print-version kit cost $174.95 and the CD-ROM-version kit cost $199.95. The CD-ROM can be purchased separately for $99.95. Both kits include observation items manuals, user guide, *What's Next? Planning Children's Activities Around Preschool Child Observations* guide, parent guide booklets (including a Spanish version), preschool COR poster, and the High/Scope preschool key experiences poster. In addition, the print-version kit includes 25 child anecdotes booklets, 1 class summary form, 25 child information and developmental summary forms, and 50 parent report forms (including a Spanish version). These materials and forms may be purchased separately (child information and developmental summary forms, $11.95 for 25; family report forms, $16.95 for 50). Pricing plans for OnlineCOR.net are also available. Costs vary based on the number of children being assessed, and whether a basic or a premium plan is selected.

Training and Other Requirements for Assessors

Is training available on how to administer and score the assessment? Who offers the training?

Yes, the publisher offers a variety of training options on how to administer and score the High/Scope COR. Information can be found on the High/Scope website.

Is it necessary to have a professional background or technical training over and above training on the assessment to administer or complete the assessment?

A teacher can administer the measure and a technical background is not necessary. However, users should familiarize themselves with the manual, procedures, and materials for the tool to be effective.

Is it necessary to have a professional background or technical training over and above training on the assessment to score the assessment?

A teacher or administrator can score the measure without a technical background once he or she is familiar with the manual, procedures, and materials.

Are regular checks on faithful administration required or recommended? If so, when and by whom?

The developer does not provide any information regarding the performance of regular checks on faithful administration.

Information Reporting System for the Assessment

Electronic Data Entry. *Does the assessment come with a process for entering information from the assessment electronically?*

Yes, the COR CD-ROM software package is a computer version of the revised preschool COR. This program enables teachers to use their computers to record and store their observations of children's development, and to generate COR scores and reports based on this information. Another option is to use OnlineCOR.net, a web-based COR. This option has the same functions as the CD-ROM, as well as providing families with a vehicle for participating in the observation of their children. Many more reporting and support features are available through the online system. Costs vary based on the number of children being assessed and basic and premium plans are offered.

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

The COR CD-ROM software package allows programs to generate COR scores and reports based on the observations that have been entered. Reports may be generated at the individual or group level. OnlineCOR.net offers reports such as a tally sheet, individual growth profiles, teacher-family journals and work sampling archives, category reports, and family reports. Programs can consolidate results across classrooms and across the program at up to seven levels. Additionally, COR data can be summarized in terms of the Head Start Outcomes with the COR-Head Start Outcomes Reporter. This software translates COR results into statistics, charts, and graphs that meet Head Start's reporting requirements.

Child Outcomes Framework. *Does the publisher map the domains in the assessment onto the domains in the Head Start Child Outcomes Framework (HSCOF)?*

Yes, the publisher maps the domains in the COR onto the domains in the HSCOF.

Instructional Support. *Does the assessment provide guidance about follow-up steps teachers can use to help individual children progress?*

Yes, the measure provides guidance to help individual children progress. After teachers have discussed, recorded, and scored anecdotes about the children in their class, program staff can use ideas in the booklet *What's Next? Planning Children's Activities Around COR Observations* on a daily basis to plan ways to support each child's development.

Planning Support. *Does the assessment come with guidance to help teachers determine follow-up steps involving planning at the classroom level?*

The manual states that COR results may be useful to teachers in assessing growth for children in groups, as well as evaluating the curriculum. A classroom-level summary form is available which allows users to summarize preschool COR results for their whole group for up to three administrations of the COR. The manual also recommends using *What's Next? Planning Children's Activities Around COR Observations* to guide planning at the classroom level. General guidance is also provided for using the COR in multicultural and multilingual classrooms.

Administrator Support. *Does the assessment come with guidance to help program administrators in planning for follow-up steps involving program improvement?*

The manual states that the COR is a tool which may reflect either a single classroom's effectiveness or an entire program's progress in meeting the developmental needs of children. Administrators may use COR results to document program changes over time. Details about planning in response to this documentation are not provided. High/Scope provides additional training on using COR results to create and implement a program improvement plan.

Approaches to Family/Parent Input

Tools for Family Input. *Does the assessment include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

Yes, the COR provides guidance for gathering and incorporating parental/family input. The manual encourages the sharing of anecdotes between the program and the family to supplement the collection of observations of the child.

Sharing Results. *Does the assessment include recommendations on how to share assessment results with the child's family?*

The COR includes a family report form to allow classroom staff to summarize and share observations and developmental materials collected with families.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this an assessment with developmental norms?*

The COR does not have developmental norms.

Which populations are included in the norming sample?

The COR does not have a norming sample.

Availability of Versions in Languages Other than English. *Is the assessment available in languages other than English? Which languages?*

The manual and other materials were developed only in English; however, translations of the materials may be available from other countries such as Canada (French), Taiwan (Chinese), and Korea (Korean).

How were versions in languages other than English developed?

This assessment is not available from the publisher in languages other than English.

What are the findings on the reliability and validity of versions of the assessment in languages other than English?

This assessment is not available from the publisher in languages other than English.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

The only suggested accommodation is that, for children functioning in the sensory-motor level of development, the infant-toddler COR may be more appropriate.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

No, the appropriateness of this assessment for diverse populations has not been examined in this way.

Reliability and Validity Information

What is known about the reliability and validity of the assessment…

In English?

The COR was tested on 393 preschool children ages 3 years, 0 months to 5 years, 5 months. The COR was used with 160 children in the spring of 2002, and with 233 children the following fall. Additional details are not available about characteristics and background of the children in this sample. The reliability and validity findings for this sample are outlined in later sections of this profile.

In other languages?

The COR is not available in other languages.

For dual language learners?

The developer does not provide any information about dual language learners and has not examined the reliability and validity of the assessment for this population.

For children with special needs?

The developer does not provide any information about children with special needs and has not examined the reliability and validity of the assessment for this population.

For American Indian/Alaskan Native children?

The developer does not provide any information about American Indian/Alaskan Native children and has not examined the reliability and validity of the assessment for this population.

For children of migrant and seasonal farm workers?

The developer does not provide any information about children of migrant and seasonal farm workers and has not examined the reliability and validity of the assessment for this population.

Reliability: Does the assessment obtain the same results, consistently, under the same conditions with the same children?

Inter-Rater reliability. *Do different raters agree when they are assessing the same children?*

To test whether children get the same score on the COR when being assessed by different raters, a subset of 41 children were observed and rated by two different people. Evaluation of the scoring completed by two different people concluded that the COR has acceptable inter-rater reliability.
The developers do not provide any information about the characteristics or training of the teachers who were involved in the study. The manual says only that they were pairs of teachers and assistant teachers from 10 classrooms, 20 teachers total, in a single program.

Test-Retest Reliability. *How consistent are scores if the assessment is administered once and then administered again soon?*

The developers have not examined the consistency of scores if the COR is administered once and then again soon.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

The relationships between items intended to reflect the same set of skills or behaviors was examined with the sample described above—a group of 160 children in the spring and 233 children in the fall of 2002. Overall, there were acceptable relationships between the items within each of the four subscales identified by researchers. In other words, items in each domain of the COR, which are intended to reflect the same set of skills or behaviors, are related. For example, all of the items within the language and literacy domain–such as understanding and using speech or having knowledge of books, sounds, or letters–were strongly related to one another.

Validity: Does the assessment measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the assessment do a good job of reflecting what the assessment is supposed to be assessing?*

The developers do not provide information about consultation with experts.

Construct Validity. *How closely related to each other are sets of items within the assessment that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

To test whether items within the measure that aim to address similar skills and behaviors are related, and determine how these items should be organized, researchers used a procedure that examines how well each individual item fits within its particular domain. Specifically, the procedure determined that the COR's 32 items would work well if combined into four categories:

- Initiative and social relations
- Creative representation and movement and music
- Language and literacy
- Mathematics and science

The COR has six broad categories, despite the fact that this analysis suggested just four categories. Moderate to strong relationships were demonstrated among the items as grouped in these four categories. For example, items known to represent initiative (such as 'making choices and plans,' 'solving problems with materials,' and 'initiating play') and items known to represent social relations (such as 'relating to adults and children,' 'resolving interpersonal conflict,' and 'understanding and expressing feelings') all were found to be closely related skills, and the analyses suggested they could be grouped within a combined initiative/social relations category. However, the developers felt it would still be useful to record and summarize information on the children in these areas separately, and these categories remain distinct in the actual COR assessment. The developers do examine whether scores on sets of items relate to children's age as expected.

Convergent Validity. *Is this assessment closely related to other well-established assessments aimed at measuring the same skills and behaviors?*

In order to examine how closely related the COR is to other well-established assessments in the field, the Cognitive Skills Assessment Battery (CSAB) was given to 28 of the children in the fall study. The CSAB produces three scores—*basic* information (correctly identifying name, address, birth date, and telephone number); *cognitive skills* (such as identifying body parts, colors, and shapes; recalling words and sentences; muscle coordination; vocabulary; and discriminating symbols, sounds, and words); and *response during assessment* (such as task persistence, attention span, and confidence). The COR language and literacy domain was strongly related to all three CSAB domains. Moderate to strong relationships also were determined between the COR initiative and social relations category and all three domains of the CSAB. For the COR creative representation and movement and music and the mathematics and science domains, moderate to strong relationships were found with the CSAB *cognitive skills* and *response during assessment* domains. The COR creative representation and movement and music and the mathematics and science domains showed weak relationships with the CSAB *basic information* domain.

Predictive Validity. *How well does this assessment predict children's later academic achievement and adjustment to school?*

The developer has not examined the extent to which this assessment predicts children's later academic achievement and adjustment to school.

References

Preschool Child Observation Record User Guide. (2003). Ypsilanti, MI: High/Scope Educational Research Foundation.

Child Observation Record (COR): Information for Decision Makers. (2005). Ypsilanti, MI: High/Scope Educational Research Foundation.

Learning Accomplishment Profile-3rd Edition (LAP-3)

Developers: Anne R. Sanford and Janet G. Zelman
Publisher: Kaplan Early Learning Company

http://chtop.org/Products/The-LAP-3.html

Purpose:
The purpose of the Learning Accomplishment Profile-3 (LAP-3) is to help teachers, clinicians, and parents assess children's skill development in several domains of development.

Developmental domains addressed in the assessment, as stated by the publisher:

- Gross motor
- Fine motor
- Prewriting
- Cognitive
- Language
- Self-help
- Personal/social

Intended age range: 36-72 months

Number of items:
The LAP-3 Assessment contains 383 items/developmental skills.

Background

Four of the tools that are included in this document are from the Learning Accomplishment System (LAP). The four tools are distinct from each other, but make up a comprehensive system of assessment and developmental screening. The LAP tools profiled include: 1) The Learning Accomplishment System-3rd Edition (LAP-3), a criterion-referenced assessment tool, meaning a child's scores on the assessment are compared to developmental benchmarks; 2) Learning Accomplishment System-Diagnostic (LAP-D), which is not a diagnostic tool, but rather a norm-referenced assessment, meaning a child's scores on the assessment are compared to the scores of a group of children the assessment was developed with and tested on; 3) LAP-D Spanish version; and 4) Learning Accomplishment System-Diagnostic Screener (LAP-D Screen), which is a shorter version of the LAP-D assessment that is used for screening for potential developmental delays.

Is the measure a direct assessment or an ongoing observational tool?

The LAP-3 can be used as both a direct assessment and an ongoing observational tool. It may be administered at specified checkpoints (e.g., beginning, middle, end of year) or used for ongoing observation, depending on the purpose and goals for which it is being used. For example, the LAP-3 may be administered at the beginning of the year to obtain a baseline of a child's development. As the year progresses, users may choose to administer the full assessment at specified points in time or document the acquisition of new skills throughout the year.

If the measure is a direct assessment, how long does it take to administer the assessment?

Generally, it takes 45 to 90 minutes to administer the LAP-3. It may take longer for children functioning significantly below age level.

Language(s) developed for:

The LAP-3 was developed in English.

Availability and Cost of the Assessment

Is the assessment available to programs without restrictions?

Yes, the assessment is available to programs without restrictions.

What is the cost of the assessment?

As of 2010, the full LAP-3 kit costs $474.95. The kit contains all materials necessary to administer the assessment to approximately 20 children. The materials also can be purchased separately, not in a kit. The Administrator's Manual costs $19.95. A set of 20 Scoring Booklets costs $22.95. The Technical Manual costs $17.95. The Illustration Manual costs $59.95.

Training and Other Requirements for Assessors

Is training available on how to administer and score the assessment? Who offers the training?

Yes, Kaplan Early Learning Company offers training on the LAP System. Information is available on the Kaplan website, http://www.kaplanco.com/services/profDev_onSiteTraining.asp

Is it necessary to have a professional background or technical training over and above training on the assessment to administer or complete the assessment?

It is not necessary to have a technical background over and above training on the assessment to administer the LAP-3. The manual recommends that teachers, paraprofessionals, clinicians, special educators, psychologists, occupational and physical therapists, speech-language pathologists, and others familiar with child development administer the assessment.

Is it necessary to have a professional background or technical training over and above training on the assessment to score the assessment?

Teachers and other assessment administrators can score the LAP-3 without a technical background. The assessment can be scored on paper using the LAP-3 Scoring Booklet, or electronically using the LAP-3 Computer Scoring Assistant (PC, Web-based, and Personal Digital Assistants software).

Are regular checks on faithful administration required or recommended? If so, when and by whom?

The developers do not provide information regarding the performance of regular checks on faithful administration.

Information Reporting System for the Assessment

Electronic Data Entry. *Does the assessment come with a process for entering information from the assessment electronically?*

Yes. LAP-3 software is available in both web and CD-ROM formats. LAP-3 software is also available for PDA to help teachers with the collection and recording of assessment data.

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

Yes, the LAP-3 software generates individual assessment results and summaries, classroom profiles, group progress charts, links to developmentally appropriate activities, and individual, classroom, and center analyses of assessment results in relation to the Head Start Child Outcomes.

Child Outcomes Framework. *Does the publisher map the domains in the assessment onto the domains in the Head Start Child Outcomes Framework (HSCOF)?*

Yes, the publishers map the domains in the LAP-3 onto the domains in the HSCOF.

Instructional Support. *Does the assessment provide guidance about follow-up steps teachers can use to help individual children progress?*

Yes, the assessment provides specific guidance for improving individual child progress. The LAP-3 Assessment Kit contains a set of 383 Learning Activity Cards that correspond to each item in the seven domains of the LAP-3. Each card presents one of more activities teachers and parents can do with children to enhance the acquisition or progression of a specific developmental skill from the LAP-3.

Planning Support. *Does the assessment come with guidance to help teachers determine follow-up steps involving planning at the classroom level?*

The manual does not provide follow-up guidance, but such guidance is discussed during training.

Administrator Support. *Does the assessment come with guidance to help program administrators plan follow-up steps involving program improvement?*

The manual does not provide follow-up guidance for program administrators, but such guidance is discussed during training.

Approaches to Family/Parent Input

Tools for Family Input. *Does the assessment include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

No, the LAP-3 does not include specific tools or guidance regarding parental/family input on a child's skills and development. However, this topic is discussed during training.

Sharing Results. *Does the assessment include recommendations on how to share assessment results with a child's family?*

No, the LAP-3 does not include recommendations on how to share results of the assessment with a child's family. However, this topic is discussed during training.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this an assessment with developmental norms?*

No, the LAP-3 does not yield norm-referenced scores regarding a child's level of functioning.

Which populations are included in the norming sample?

The LAP-3 does not have a norming sample.

Availability of Versions in Languages Other than English. *Is the assessment available in languages other than English? Which languages?*

No, the LAP-3 is not available in languages other than English.

How were versions in languages other than English developed?

There are no versions of the LAP-3 in languages other than English.

Is there any evidence that versions in languages other than English were developed with a representative group of children speaking that language, either as monolinguals or bilinguals?

There are no versions of the LAP-3 in languages other than English.

What are the findings on the reliability and validity of versions of the assessment in languages other than English?

There are no versions of the LAP-3 in languages other than English.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

The manual suggests that adaptations may be made for children with special needs, but it does not describe any specific accommodations.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

The developers did not provide information about testing for appropriate use with diverse populations in this way.

Reliability and Validity Information

What is known about the reliability and validity of the assessment…

In English?

To test the reliability and validity of the LAP-3 in English, developers tested it on a sample of 363 children of different races, ethnicities, ages, genders, and from different geographic locations, family composition, household incomes, and parental levels of education. About half (51.5%) of the children were female. These children were recruited from various educational settings including center-based care programs, Head Start programs, public school kindergartens, family day care programs, and individual homes. The sites were located across the United States, and represent the Northeast, South, Central, and Northwest regions. This sample was representative of children in the United States based on population projections for the year 2000 by the U.S. Census Bureau (1995). A sample of atypically developing children was included to examine whether the LAP-3 could be used appropriately with children with disabilities. See the table on the next page for more information about these children. The reliability and validity findings for this sample are outlined in later sections of this profile.

In other languages?

The developers did not examine the reliability and validity of the LAP-3 in other languages because the LAP-3 is not available in other languages.

For dual language learners?

The developers do not provide any information about dual language learners and have not examined the reliability and validity for this population.

For children with special needs?

While children with special needs were included in the sample of children the assessment tested, reliability and validity of the LAP-3 for this group has not been examined. During the reliability and validity study, the assessment was tested on a subsample of 28 children who had been professionally diagnosed and were receiving special education services. These children ranged in age from 33 to 73 months. They were 39.3% females and 60.7% males, and were 10.7% African American, 14.3% Asian and Pacific Islander, 10.7% Hispanic, 53.6% White, and 10.7% Other in terms of race/ethnicity. Of the 28 children in the sample, 8 children had developmental delays, 2 children had motor disabilities, 7 children had speech and language disabilities, 3 children had Autism, 1 child had ADHD, and 7 had multiple disabilities. If information regarding the individual child's developmental level of functioning is not available, administrators should begin the assessment at half the child's chronological age. The relationship between the scores for children with disabilities and their chronological age is lower than for typically developing children, indicating that the LAP-3 discriminates children's skill levels independently of their age. Thus, while the LAP-3 can be used with children with disabilities, there is no separate reliability or validity information for this population.

For American Indian/Alaskan Native children?

While American Indian/Alaskan Native children were included in the sample the assessment tested (2.2% of children), the developers did not examine the reliability and validity for this group of the LAP-3.

For children of migrant and seasonal farm workers?

The developers do not provide any information about children of migrant and seasonal farm-workers and have not examined the reliability and validity for this population.

Demographic Characteristics of 2004 Reliability and Validity Study Sample
(Continues on next page)
Number of children in the sample: 363

	Percentage of Children
Age of Children	
30-35 months	3.0
36-41 months	9.9
42-47 months	14.3
48-53 months	21.2
54-59 months	20.1
60-65 months	13.8
66-72 months	14.3
73-78 months	3.3
Racial/Ethnic Group	
White	58.2
African American	19.0
Hispanic Origin	6.3
Asian and Pacific Islander	5.5
American Indian, Eskimo, Aleut	2.2
Other[6]	8.8
Gender	
Male	48.5
Female	51.5
Maternal Education	
Less than High School	3.9
High School	24.2
Some college	11.0
Associate's degree	2.2
Bachelor's degree	27.3
Master's degree	16.2
Doctoral degree	9.6
No report	5.5
Paternal Education	
Less than High School	3.9
High School	23.4
Some college	6.6
Associate's degree	1.9
Bachelor's degree	24.2
Master's degree	12.4
Doctoral degree	12.9
No report	14.6

[6] Children classified as Other were multiracial according to the following distribution: Hispanic/White=12; African American/White=6, and other multiracial groups=14.

Demographic Characteristics of 2004 Reliability and Validity Study Sample (continued)

	Percentage of Children
Household Income[7]	
Below $10K	9.9
$10K to $20K	11.3
$20K to $30K	5.5
$30K to $40K	9.1
$40K+	59.0
Geographic Distribution	
Northeast	11.6
South	47.4
Central	30.3
Northwest	10.7
Number of Children in Each Type of Care Setting[8]	
Center-Based	60.1
Head Start Programs	23.7
Public School Kindergartens	11.0
Family Day care Programs	1.9
Individual Homes	3.3

[7] Of the 363 children in the sample, 344 families reported annual income, which is why the percentage does not equal 100%.

[8] There were 29 child care settings where data collection took place for this reliability and validity study.

Reliability: Does the assessment obtain the same results, consistently, under the same conditions with the same children?

Inter-Rater Reliability. *Do different raters agree when they are assessing the same children?*

Yes, two raters who were adequately trained on the assessment should get a similar score when administering the LAP-3 to the same child. To test whether children get the same score on the LAP-3 when assessed by different raters, the assessment was administered to a subset of 33 children by two different examiners on two separate occasions one to three weeks apart. This sample included children ages 33 to 73 months old, 51.5% females and 48.5% males, and was 18.2% African American, 9.1% Asian and Pacific Islander, 6.1% Hispanic origin, 60.6% White, and 6.1% Other in terms of race/ethnicity. The developers did not provide information about the characteristics of the assessment examiners. The results show that agreement between two different raters when they are assessing the same children is acceptable.

Test-Retest Reliability. *How consistent are scores if the assessment is administered once and then administered again soon?*

The LAP-3 was administered by the same examiner on two separate occasions one to three weeks apart for a subset of 40 children from the overall project sample. The age range of this sample was 37 to 72 months, and included both typically and atypically developing children. The sample consisted of 55% females and 45% males, and was 5% African American, 5% Asian and Pacific Islander, 5% Hispanic origin, 65% White, and 15% Other in terms of race/ethnicity. No information is provided about the characteristics of the assessment examiners. The results of the assessments showed that the scores in all the domains of the LAP-3 were very similar on the first and second assessment for all children. This suggests that the consistency of individual scores is acceptable over short intervals of time.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

Relationships between items on the LAP-3 were examined with a sample of 251 children. The developers did not provide additional information about the characteristics of this sample. The relationships between items that are intended to reflect the same set of skills or behaviors are acceptable. There is an acceptable relationship between all items within each subscale. The relationships between items within subscales for the individual age groups are also acceptable.

Validity: Does the instrument measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the assessment do a good job of reflecting what the assessment is supposed to be assessing?*

During the development of the LAP-3, a panel of experts in child development and early childhood education reviewed each item of the assessment. The majority of the items were viewed as appropriate for the domain and developmental age range. Items deemed unsatisfactory were dropped or moved as recommended by the expert reviewers.

Construct Validity. *How closely related to each other are sets of items within the assessment that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

To see how closely the sets of items in the LAP-3 are related to each other, the scores in the different sets of items were compared to each other for children in each age group. The results of these comparisons showed moderate to weak relationships between the scores for different sections. For example, when comparing the scores on the gross motor domain to the scores on the self-help domain for all ages, the scores were very different and not strongly related. This means that the gross motor and self-help domains are capturing very different skills and behaviors in children of all ages, as expected. However, the scores on the fine motor and pre-writing sections were moderately related, suggesting that there are some similar skills used to respond to the items in both these domains, as expected.

These results suggest that while the different domains of the LAP-3 are somewhat related, they are also measuring somewhat independent aspects of development. Further, sets of items that might be anticipated to be more closely related were found to be closely related, while those viewed as most different were not as closely related. The developers do examine whether scores on sets of items relate to children's age as expected.

Convergent Validity.
Is this assessment closely related to other well-established assessments aimed at measuring the same skills and behaviors?

In order to examine how closely related the LAP-3 is to other well-established assessments in the field, 230 children from the sample described in the table above were administered both the LAP-3 and the Battelle Developmental Inventory (BDI), either during the same testing session or in two sessions within a very short period of time. The results of the assessments show that there is a moderate relationship between the LAP-3 and the BDI for conceptually related domains.

Predictive Validity. *How well does this assessment predict children's later academic achievement and adjustment to school?*

The developers did not predict children's later academic achievement or adjustment to school.

References

Hardin, B.J., & Peisner-Feinberg, E. (2004). *The Learning Accomplishment Profile-Third Edition (LAP-3): Examiners manual and technical report*. Chapel Hill Training Outreach Project, Inc. and Kaplan Early Learning Company.

Learning Accomplishment Profile-Diagnostic 3rd Edition (LAP-D)

Developers: David LeMay, Patricia Griffin, and Anne R. Sanford
Publisher: Kaplan Press

http://chtop.org/Products/The-LAP-D.html

Purpose:
The Learning Accomplishment Profile-Diagnostic is a norm-referenced assessment that is used to observe a child's level of development and functioning.

Developmental domains addressed in the assessment, as stated by the publisher:

- Fine motor
  - o Manipulation
  - o Writing
- Cognitive
  - o Matching
  - o Counting
- Language
  - o Comprehension
  - o Naming
- Gross motor
  - o Body movement
  - o Object movement

Intended age range:
30-72 months

Number of items:
The full assessment contains 226 items that are hierarchically arranged by developmental level. It is not always necessary to administer the full assessment based on the child's level of development.

Background

Four of the tools that are included in this document are from the Learning Accomplishment Profile (LAP). The four tools are distinct from each other, but make up a comprehensive system of assessment and developmental screening. The LAP tools profiled include: 1) Learning Accomplishment System 3rd Edition (LAP-3), a criterion-referenced assessment, meaning child's scores on the assessment are compared to developmental benchmarks; 2) Learning Accomplishment Profile-Diagnostic (LAP-D) which is not a diagnostic tool, but rather a norm-referenced assessment, meaning child's scores on the assessment are compared to the scores of a group of children the assessment was developed with and tested on; 3) LAP-D Spanish version; and 4). Learning Accomplishment Profile-Diagnostic Screener (LAP-D Screen), which is a shorter version of the LAP-D assessment that is used for screening for potential developmental delays.

Is the assessment a direct assessment or an ongoing observational tool?

The LAP-D is a direct assessment.

If the assessment is a direct assessment, how long does it take to administer the measure?

On average, the assessment takes one to one and a half hours to administer. However, the time can vary depending on the child's level of development.

Language(s) developed for:

The LAP-D was developed in English. There is also a Spanish version of the assessment that has a separate profile in this compendium.

Availability and Cost of the Assessment

Is the assessment available to programs without restrictions?

Yes, the assessment is available to programs without restrictions.

What is the cost of the assessment?

As of 2010, the full LAP-D kit costs $799.00. The kit includes the manual and all of the items needed to complete and score the assessment for 10 children. It is also possible to purchase additional scoring sheets (in packs of 10) as well as refills for some of the items that require special paper or materials.

Training and Other Requirements for Assessors

Is training available on how to administer and score the assessment? Who offers the training?

Yes, Kaplan Early Learning Company offers training on the LAP System. Information is available on the Kaplan website http://www.kaplanco.com/services/profDev_onSiteTraining.asp, however the website does not detail which LAP assessment tools are covered in the training.

Is it necessary to have a professional background or technical training over and above training on the assessment to administer or complete the assessment?

While it is not necessary to have a professional background or additional technical training to administer the assessment, the developers recommend that people who administer the assessment be familiar with child development; examples of such individuals are clinicians, teachers, special educators, and psychologists. Additionally, anyone who administers the assessment should be familiar with the *Examiner's Manual and Technical Report* before administering it.

Is it necessary to have a professional background or technical training over and above training on the assessment to score the assessment?

No, according to the developers, it is not necessary to have a professional background or technical training (in addition to training on the assessment) to score the assessment. Additionally, anyone who scores the assessment should be familiar with the *Examiner's Manual and Technical Report* before scoring it.

Are regular checks on administration required or recommended to ensure appropriate administration? If so, when and by whom?

The manual does not offer guidance regarding the performance of regular checks on faithful administration.

Information Reporting System for the Assessment

Electronic Data Entry. *Does the assessment come with a process for entering information from the assessment electronically?*

Yes, the assessment has an electronic scoring system that can be purchased separately (for information about the range of prices, see http://www.redesetgrow.com/products.html). Information from the LAP-D assessment can be scored electronically or on paper. On paper, a scoring summary profile is created for each child that summarizes the scores on all of the subscales in the LAP-D. This summary also indicates the percentile rank and age equivalency based on the child's score. The electronic version is available in web and CD formats. It is also possible to have the software on a Personal Data Assistant (PDA), which can be used instead of the scoring pad for direct entry of data onto a computer.

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

Yes, the electronic system can assist in analyzing data for individual children or groups. The software can generate individual assessment results and summaries, classroom profiles, parent reports, group progress charts, links to developmentally appropriate activities, and analyses related to the Head Start Child Outcomes.

Child Outcomes Framework. *Does the publisher map the domains in the assessment onto the domains in the Head Start Child Outcomes Framework (HSCOF)?*

Yes, the publishers map the domains in the LAP-D onto the domains in the HSCOF.

Instructional Support. *Does the assessment provide guidance about follow-up steps teachers can use to help individual children progress?*

Yes, the manual gives some general suggestions regarding follow up steps that teachers can use for individual children, but these suggestions are not tailored to a child's score on the assessment.

The results of the LAP-D can be used as a basis for an Individualized Education Plan (IEP), Individualized Family Service Plan (IFSP), or any other form of individualized instruction if the child's score suggests that he or she might benefit from an IEP. More specifically, the results from the LAP-D can help shape long-term goals and short-term objectives for a child. Weak scores in subscales (broad developmental areas) of the LAP-D can indicate where long-term goals should be focused. Additionally, assessing the child during and/or after implementing individualized instruction can show growth in certain developmental areas. An additional publication that can be purchased entitled *Guide for Developing an Individual Education Program (IEP) with the LAP-D*, for $14.99.

Planning Support. *Does the assessment come with guidance to help teachers determine follow-up steps involving planning at the classroom level?*

Yes, the assessment provides general guidance for teachers, but the suggestions are not tailored to the classrooms' scores on the assessment. An additional set of materials, entitled *LAP-D Pupil/Teacher/Parent Planning Cards*, can be purchased for $49.95. Each card includes a skill from the LAP-D assessment followed by activities in art, blocks, dramatic play, library, music, outdoors, sand/water, science/discovery, and table games. These activities are intended to expand and reinforce skill development. Each card references the LAP-D skill area and development age.

Administrator Support. *Does the assessment come with guidance to help program administrators plan for follow-up steps involving program improvement?*

No, the assessment does not come with guidance for program improvement.

Approaches to Family/Parent Input

Tools for Family Input. *Does the assessment include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

No. The assessment does not include specific tools for gathering input from parents or other family members. The only person who can provide input on a child's skills and development is the person administering the assessment.

Sharing Results. *Does the assessment include recommendations on how to share assessment results with a child's family?*

Yes, the assessment includes recommendations on how to share the results with a child's family. Specifically, the manual lists many suggestions about sharing the results, including the importance of establishing rapport with families and providing some background information about the LAP-D. Administrators should ask parents or other family members how they think their child is progressing; this is one time when questions can be asked about a child's behavior at home, but it does not change the results of the assessment. When giving parents or family members the results of the LAP-D, the administrator should provide context for the results of the assessment. In addition to speaking with the parents or family members, it is valuable to provide a written summary of the assessment.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this an assessment with developmental norms?*

Yes, the LAP-D is an assessment with developmental norms. The assessment was normed in 2005. This sample is described below and in the table on the following page.

Which populations were included in the norming sample?

The norming sample included 2,099 children of different races, ethnicities, ages, and genders, from different geographic locations, family compositions, household incomes, and parental levels of education.

These children were recruited from various educational settings, including center-based care programs, Head Start programs, private schools, public schools, and other settings. The sites were located across the United States representing the Northeast, South, Central and Southwest. This sample was representative of children in the United States based on data from the U.S. Census from the year 2000. Of the 2,099 children in the reliability and validity study, 1,124 spoke English and 975 spoke Spanish.

This profile focuses on the English version of the Learning Accomplishment Profile. The table on the next page provides more information about these children. Please see the Learning Accomplishment Profile-Diagnostic, Spanish Edition Profile in this compendium for details about the Spanish-speaking sample.

Options for Use with Special and Diverse Populations

Availability of Versions in Languages Other than English. *Is the assessment available in languages other than English? Which languages?*

Yes, the assessment is available in Spanish.

How were versions in languages other than English developed?

The Spanish version of the LAP-D was translated and adapted from the English version using a consensus method in which many experts worked together to create the best possible translation and adaptation of the assessment.

What are the findings on the reliability and validity of versions of the assessment in languages other than English?

See the Learning Accomplishment Profile-Diagnostic, Spanish Edition Profile in this compendium for more information about reliability and validity for the Spanish version.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

According to the developer, the LAP-D is an appropriate assessment to use with children who are diagnosed with special needs. Children who may not be functioning at age level due to developmental delays still can be assessed with this tool as long as their level of functioning is above the minimum age for the assessment. However, they may not begin at the item corresponding with their chronological age; The manual provides guidance as to where to begin for these children. This may result in a longer administration time for some children. There are no additional suggestions for administering the assessment for children with special needs.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

No, the appropriateness of this assessment for diverse populations has not been examined in this way.

Characteristics of 2005 Norming Sample (continues on next page)
Number of children in the sample: 1,124

	Percentage of Children
Age of Children	
30-35 months	8.9
36-41 months	11.03
42-47 months	16.01
48-53 months	16.10
54-59 months	19.3
60-65 months	16.28
66-72 months	12.37
Racial/Ethnic Group	
White	47.9
African American	11.6
Hispanic Origin	28
Asian and Pacific Islander	1.8
American Indian, Eskimo, Aleut	1
Other	9.7
Maternal Education[9]	
Less than High School	10.1
High School	32.8
Associates Degree	17
Bachelors Degree	21.5
Masters Degree	9.2
Doctoral Degree	1.1
Paternal Education[10]	
Less than High School	8.7
High School	31.7
Associates Degree	11.8
Bachelors Degree	18.4
Masters Degree	6.8
Doctoral Degree	1.6
Household Income[11]	
Under $10K	7.78
$10K to $20K	9.59
$20K to $30K	6.68
$30K to $40K	4.76
$40K to $50K	4.64
$50K to $60K	3.83
$60K to $70K	4.70
$70K to $80K	5.69
$80K+	11.1
Number of Adults in the Home	
1	17.8
2	71.8
3 or more	10.4

[9] Only 1,030 mothers reported their education level.
[10] 888 fathers reported their education level.
[11] Only 950 of the English-speaking families reported household income levels.

Characteristics of 2005 Norming Sample (continued)

	Percentage of Children
Number of Children in the Home	
1	23.9
2	43.1
3 or more	33
Type of Care Setting Children Attended	
Center-Based	50
Head Start	18.5
Private School	14.6
Public School	9.2
Other	7.7

Reliability and Validity Information

What is known about the reliability and validity of the assessment…

In English?

There is information about reliability and validity of the assessment in English. This information is outlined in later questions of this profile.

In other languages?

For information on the Spanish version of the Learning Accomplishment Profile-Diagnostic, see the profile of the Spanish version in this compendium.

For dual language learners?

Although there were children who spoke Spanish in the sample used to test the LAP-D, they were given the Spanish version and a separate profile is available with that information. The developers do not indicate whether dual language learners were included in the sample and have not examined the reliability and validity of this instrument for dual language learners separately.

For children with special needs?

While children with special needs were included in the norming sample (4.36% of children), the publishers do not provide separate information on the reliability and validity of the LAP-D for children with special needs. This sample was referred to as the Atypical Development Sample. These children were diagnosed by a specialist and all of them were receiving some form of special education. The disabilities included speech/language disabilities, social-emotional disabilities, motor or other health disabilities, behavioral disabilities, and other disabilities defined by the state.

For American Indian/Alaskan Native children?

While American Indian/Alaskan Native children were included in the norming sample (1% of children), there is no separate information about the reliability and validity for this group.

For children of migrant and seasonal farm workers?

The reliability and validity for children of migrant and seasonal farm workers have not been examined.

Reliability: Does the assessment obtain the same results, consistently, under the same conditions with the same children?

Inter-Rater Reliability. *Do different raters agree when they are assessing the same children?*

Yes, two raters who have been adequately trained on the assessment can get a similar score when administering the LAP-D on the same child. To test whether children get the same score on the LAP-D when being assessed by different raters, two different raters each assessed a subset of the English-speaking children from the 2005 sample. On average, scores from the second assessment, with a different rater, agreed with scores from the first assessment.

The groups for which inter-rater reliability was tested include 58 children with diverse racial and ethnic backgrounds. Compared to the total sample in the 2005 study, this group had a somewhat higher number of Hispanics and Whites (31 percent and 52 percent, respectively) and a slightly lower number of African Americans (8 percent). Two percent of these children had a diagnosed special need. They ranged from 30-71 months of age. No information about the characteristics of the assessors was given.

Test-Retest Reliability. *How consistent are scores if the assessment is administered once and then administered again soon?*

Scores on all four of the LAP-D domains met the criteria for acceptable consistency when the assessment was administered twice (one to three weeks apart). This was done with 163 English-speaking children and 155 Spanish-speaking children, but the percentages are only for the English-speaking sample. This group of children was relatively comparable to the 2005 study, with more White children (51.53 percent). No information about the characteristics of the assessors was given.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

Within each of the domains addressed by the LAP-D, the strength of the relationships between items intended to reflect the same set of skills met the criteria for acceptable relationships. The weakest relationship among items was in the object movement subscale of the gross motor domain, but it still met the criteria for an acceptable relationship. This was examined with the majority of the English-speaking children (total =1,075 out of 1,124) in the 2005 study. Further information about the children in the English-speaking sample can be found in the table earlier in this profile.

Validity: Does the assessment measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the assessment do a good job of reflecting what the assessment is supposed to be assessing?*

The developers do not provide information about expert consultation.

Construct Validity. *How closely related to each other are sets of items within the assessment that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

Sets of items that are intended to address similar skills and behaviors are moderately to strongly related to each other. There are strong relationships between the cognitive domain and both the fine motor and language domains. Strong relationships also exist between the items in the language domain and those in both the fine motor and cognitive domains. Moderate relationships exist between items in the fine motor domain and those in the language domain. There were weak relationships between sets of items that aimed to address different skills. For example, there was a weak relationship between items that address language skills and those that address gross motor skills. The developers do examine whether scores on sets of items relate to children's age as expected.

Convergent Validity. *Is this assessment closely related to other well-established assessments aimed at measuring the same skills and behaviors?*

Overall, the LAP-D has demonstrated acceptable relationships when compared with three other assessments.

LAP-D and the Developmental Indicators for the Assessment of Learning, 3rd Edition (DIAL-3): A set of 85 children from the English-speaking sample in the 2005 study was given the LAP-D and the DIAL-3 screener at either the same assessment session or within a very short period of time. The DIAL-3 screener is a comprehensive, norm-referenced screener of child development (for more information, see the DIAL-3 developmental screener profile in this compendium). Only the motor concepts and language scales on the DIAL-3 were used in this study since they were the most conceptually related to the LAP-D domains. The scores on the LAP-D and the DIAL-3 showed a strong relationship across these two domains. The strongest relationship was between the fine motor domain on the LAP-D and the Motor domain on the DIAL-3. The developers do not provide additional information about the demographic characteristics of these 85 children.

LAP-D and Woodcock Johnson-Revised (WJ-R): Another set of 231 children was given the WJ-R assessment at either the same session or within a very short period of time as the LAP-D. The WJ-R is a norm-referenced set of assessments used to assess cognitive abilities. All four of the LAP-D domains were compared with three domains from the WJ-R: letter-word identification, applied problems, and dictation. While there was wide variation in the strengths of the relationships across the domains of the LAP-D and the WJ-R, all met the established criteria for strong relationships. No additional information about these 85 children in provided.

LAP-D and Peabody Picture Vocabulary Test, 3rd Edition (PPVT-III): A third group of 984 children was given the PPVT-III assessment at either the same session as the LAP-D or within a very short period of time. The PPVT-III is a norm-referenced assessment used to assess receptive vocabulary. All four of the domains on the LAP-D were compared to the PPVT-III. The comparison of scores yielded a strong relationship between the Cognitive and Language domains of the LAP-D and the PPVT-III assessment. The Fine Motor and Gross Motor domains were high, but were less strongly related to the PPVT. No additional information about the demographic characteristics of these 984 children is provided.

Overall, these results suggest that the scores generated from the LAP-D are, for the most part, comparable to the scores on other well-established assessments in the field that measure similar developmental skills. Children who score high on the specific developmental skills captured by these other assessments should score high on the LAP-D and vice versa.

Validity: Does the assessment measure what it is supposed to?

Predictive Validity. *How well does this assessment predict children's later academic achievement and adjustment to school? For what groups of children has this been examined?*

The relationship between children's LAP-D scores and their later academic achievement or adjustment to school has not been examined.

Comments

It was surprising that when comparing the LAP-D to the PPVT-III, the relationships between the Fine Motor and Gross Motor domains were relatively strong. Since the Fine Motor and Gross Motor domains are less clearly related to receptive vocabulary, it might be anticipated that the relationship would not be as strong.

References

Hardin, B.J., Peisner-Feinberg, E.S., & Weeks, S.W. (2005). Learning Accomplishment Profile Diagnostic Edition (Revised 3rd Edition): Examiner's Manual and Technical Report. Chapel Hill Training Outreach Project, Inc. Kaplan Early Learning Company.

Learning Accomplishment Profile- Diagnostic 3rd Edition, Spanish Language Edition (LAP-D Spanish)*

Developers: David LeMay, Patricia Griffin, and Anne R. Sanford
Publisher: Kaplan Press

http://chtop.org/Products/The-LAP-D.html

Purpose:
The Learning Accomplishment Profile-Diagnostic Spanish Language Edition is a norm-referenced assessment that is used to chart a Spanish-speaking child's developmental level of functioning.

Developmental domains addressed in the assessment, as stated by the publisher:

- Fine motor
  - Manipulation
  - Writing
- Cognitive
  - Matching
  - Counting
- Language
  - Comprehension
  - Naming
- Gross motor
  - Body movement
  - Object movement

Intended age range: 30-72 months

Number of items:
The full assessment contains 226 items that are hierarchically arranged by developmental level. It is not always necessary to administer all items.

Background

Four of the tools that are included in this document are from the Learning Accomplishment System (LAP). The four tools are distinct from each other, but are from a comprehensive system of assessment and developmental screening. The Learning Accomplishment System-3rd Edition (LAP-3) is a criterion-referenced assessment, meaning the child's scores on the assessment are compared to developmental benchmarks. The Learning Accomplishment System-Diagnostic (LAP-D) is not a diagnostic tool, but is a norm-referenced assessment, meaning that a child's scores on the assessment are compared to the scores of a group of children for whom the assessment was developed. There is a separate profile for the LAP-D assessment in Spanish. Finally, there is a profile for the Learning Accomplishment System-Diagnostic Screener (LAP-D Screen), a shorter version of the LAP-D assessment that is used for screening for potential developmental delays. This profile addresses the Spanish version of the LAP-D.

Is the assessment a direct assessment or an ongoing observational tool?

The LAP-D Spanish is a direct assessment.

If the assessment is a direct assessment, how long does it take to administer the measure?

On average, the assessment takes one to one and a half hours to administer. However, the time can vary depending on a child's level of development.

Language(s) developed for:

This assessment was developed for English language use, but was translated and adapted for Spanish language use. A consensus group of experts agreed on translations that are appropriate for seven different Spanish-speaking countries. The Spanish version then was given to a committee of native speakers with a background in early childhood education or similar field. The also developers completed a pilot test, which is discussed later in this profile.

What is the appropriate time period between administering, recording, or reviewing the data?

The manual indicates that the assessment should be administered at specific times throughout the year–based on the program's needs–in order to document progress in a child's development. For example, it could be administered at the beginning, middle, and end of the year.

Availability and Cost of the Assessment

Is the assessment available to programs without restrictions?

Yes, the assessment is available to programs without restrictions.

What is the cost of the assessment?

As of 2010, the full LAP-D Spanish kit costs $799.00. The kit includes the manual and all of the items needed to complete and score the assessment for 10 children. It is also possible to purchase additional scoring sheets for $9.95 (in packs of 30), as well as refills for some of the items that require special paper or materials.

Training and Other Requirements for Assessors

Is training available on how to administer and score the assessment? Who offers the training?

Yes, Kaplan Early Learning Company offers training on the LAP. Information is available on the Kaplan website, however, the website does not detail which LAP assessment tools are covered in the training. For more information, see http://www.kaplanco.com/services/profDev_onSiteTraining.asp.

Is it necessary to have a professional background or technical training over and above training on the assessment to administer or complete the assessment?

While it is not necessary to have a professional background or additional technical training to administer the assessment, the developers recommend that people who administer the assessment be familiar with child development; examples of such individuals are clinicians, teachers, special educators, and psychologists. Additionally, anyone who administers the assessment should be familiar with the Examiner's Manual and Technical Report before administering it.

Is it necessary to have a professional background or technical training over and above training on the assessment to score the assessment?

No, according to the developers, it is not necessary to have a professional background or technical training (in addition to training on the assessment) to score the assessment. Those who score the assessment should be familiar with the Examiner's Manual and Technical Report before scoring it.

Are regular checks on administration required or recommended to ensure appropriate administration? If so, when and by whom?

The manual offers no guidance regarding performance of regular checks of administration.

Information Reporting System for the Assessment

Electronic Data Entry. *Does the assessment come with a process for entering information from the assessment electronically?*

Yes, the assessment has an electronic scoring system that can be purchased separately (for prices, see http://www.redesetgrow.com/products.html). The information collected from the LAP-D Spanish assessment can be scored electronically or on paper. On paper, a scoring summary profile is created for each child that summarizes the scores on all of the subscales in the LAP-D. This summary also indicates the percentile rank and age equivalency based on the child's score. The electronic version is available in web and CD formats. It is also possible to have the software on a Personal Data Assistant (PDA), which can be used instead of the scoring pad for direct entry of data onto a computer.

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

Yes, the electronic system can assist in analyzing data for individual children or groups. The software can generate individual assessment results and summaries, classroom profiles, parent reports, group progress charts, links to developmentally appropriate activities, and analyses related to the Head Start Child Outcomes.

Child Outcomes Framework. *Does the publisher map the domains in the assessment onto the domains in the Head Start Child Outcomes Framework (HSCOF)?*

Yes, the publisher did map the domains in the LAP-D Spanish onto the domains in the HSCOF.

Instructional Support. *Does the assessment provide guidance about follow-up steps teachers can use to help individual children progress?*

Yes, the manual gives some general suggestions regarding follow-up steps teachers can use for individual children. However, it does not provide specific activities or programs and these suggestions are not tailored to a child's score on the assessment.

Planning Support. *Does the assessment come with guidance to help teachers determine follow-up steps involving planning at the classroom level?*

No, the assessment does not come with guidance for follow-up at the classroom level.

Administrator Support. *Does the assessment come with guidance to help program administrators plan for follow-up steps involving program improvement?*

No, the assessment does not come with guidance for follow-up steps for program administrators.

Approaches to Family/Parent Input

Tools for Family Input. *Does the assessment include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

No, the assessment does not include specific tools for gathering input from parents or other family members. The only person who can provide input on a child's skills and development is the person administering the assessment.

Sharing Results. *Does the assessment include recommendations on how to share assessment results with a child's family?*

Yes, the assessment includes recommendations on how to share the results with a child's family. Specifically, the manual lists many suggestions about sharing the results, including establishing rapport with families and providing some background information about the LAP-D Spanish. The assessment recommends that the parents or family members be asked about how they think their child is progressing. This is one time where the administrator can ask about a child's behavior at home, but it does not change the results of the assessment. When giving parents or family members the results of the LAP-D, the administrator should provide context for the results. In addition to speaking with parents or family members, it is valuable to provide a written summary of the assessment.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this an assessment with developmental norms?*

Yes, the LAP-D Spanish is an assessment with developmental norms. The assessment was normed in 2005. This sample is described below and in the table on the following page.

Which populations were included in the norming sample?

The norms for the Spanish version were created based on 975 Spanish-speaking children among the larger sample of 2,099, as detailed in the LAP-D English-version profile. This sample was representative of the cultural and ethnic diversity within the Latino population of the United States as of the 2000 Census. Mothers' country of origin was used to determine children's country of origin.

Availability of Versions in Languages Other than English. *Is the assessment available in languages other than English? Which languages?*

Yes, this assessment is in Spanish.

How were versions in languages other than English developed?

While the LAP-D was developed in English, it was translated and adapted for Spanish language use. A consensus group of experts agreed on translations that are appropriate for seven different Spanish-speaking countries. The Spanish version then was given to a committee of native speakers with a background in early childhood education or a similar field. A pilot study was completed and is discussed later in this profile.

Is there any evidence that versions in languages other than English were developed with a representative group of children speaking that language, either as monolinguals or bilinguals?

Yes, as seen in the characteristics of the sample in the chart in this profile, a representative group of Spanish-speaking children were included in the development of the LAP-D Spanish. Additionally, 43 Spanish-speaking children completed a pilot study, however no information is provided about bilingual status.

What are the findings on the reliability and validity of versions of the assessment in languages other than English?

The findings of the LAP-D Spanish are presented in subsequent pages in this profile.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

According to the publisher, children who may not be functioning at age level can still be assessed with this assessment as long as their level of functioning is at a level above the minimum age for the assessment. However, they may not begin at the item corresponding with their chronological age. The manual gives guidance as to where to begin for these children. This may result in a longer administration time with some children.

What are the findings on reliability and validity of the assessment for children with special needs?

While children with special needs were included in the sample of children on which the LAP-D Spanish was tested, there is no separate reliability or validity information for this group. There are no additional suggestions for administering the assessment on children with special needs.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

Yes, through consultation with experts, translators, and through a pilot study, the developers examined the appropriateness of the translation of the LAP-D for multiple Spanish-speaking subpopulations.

Characteristics of 2005 Norming Sample (continues on next page)
Number of children in the sample: 975

	Percentage of Children
Cultural Background	
Central and South American	15.1
Cuban	2.9
Mexican	43.1
Puerto Rican	1.4
Other[12]	37.5
Maternal Education[13]	
Less than High School	34.8
High School	24.7
Associates Degree	7.7
Bachelors Degree	3.8
Masters Degree	0.9
Doctoral Degree	0.2
Paternal Education[14]	
Less than High School	35.5
High School	25.7
Associates Degree	8.1
Bachelors Degree	2.9
Masters Degree	1.2
Doctoral Degree	0.3
Household Income[15]	
Under $10K	9.83
$10K to $20K	16.30
$20K to $30K	8.10
$30K to $40K	2.97
$40K to $50K	1.67
$50K to $60K	0.43
$60K to $70K	0.001
$70K to $80K	0.43
$80K+	1.48
Number of Adults in the Home	
1	10.1
2	64.3
3 or more	25.6

[12] Other was reported in these categories: not reported=330 (32.8%), Dominican Republic=10 (1%), Other=26 (2.67%).
[13] Only 811 mothers reported their education level.
[14] Only 719 fathers reported their education level.
[15] Only 667 (41.2%) of the Spanish-speaking families reported household income levels, which is why the total is less than 100%.

Characteristics of 2005 Norming Sample (continued)

	Percent of Children
Number of Children in the Home	
1	15.3
2	36.7
3 or more	48
Type of Care Setting Children Attended (130 sites total)	
Center-Based	50
Head Start	18.5
Private School	14.6
Public School	9.2
Other	7.7

Reliability and Validity Information

What is known about the reliability and validity of the assessment…

In English?

Reliability and validity information on the Learning Accomplishment Profile-Diagnostic English version is available for children who speak English. More details are provided in the Learning Accomplishment Profile-Diagnostic English version profile in this document.

In other languages?

Reliability and validity information is available for children who speak Spanish. More details are provided in the sections below.

For dual language learners?

The reliability and validity of the LAP-D Spanish was normed on a sample of Spanish-speaking children residing in the United States and representative of the cultural and ethnic diversity within the Latino population of the United States as of the 2000 Census. The manual does not indicate whether the Spanish version is reliable and valid for Spanish monolingual children, nor does it indicate the level of verbal fluency of the norming sample in Spanish or English at the time of testing.

For children with special needs?

While children with special needs were included in the norming sample (2.9% of children), the publisher does not provide separate information on the reliability and validity of the LAP-D for children with special needs. This sample was referred to as the atypical development sample. A specialist diagnosed these children, all of whom were receiving some form of special education. Their disabilities included speech/language, social-emotional, motor, other health, behavioral, and other disabilities defined by the state.

For American Indian/Alaskan Native children?

There were no American Indian/Alaskan Native children in the sample and the publisher did not report on reliability and validity for this population.

For children of migrant and seasonal farm workers?

The reliability and validity for children of migrant and seasonal farm-workers have not been examined.

Reliability: Does the assessment obtain the same results, consistently, under the same conditions with the same children?

Inter-Rater Reliability. *Do different raters agree when they are assessing the same children?*

Yes, two raters who were adequately trained on the assessment strongly agreed when administering the LAP-D Spanish on the same child. To test whether children get the same score on the LAP-D Spanish when being assessed by different raters, a subset of children completed the assessment twice with two different people administering it. This was examined with 89 children from the larger sample. Compared to the 2005 sample, this sample included many more children of Latino background (64%), although there is no specific information about their ethnic background. No information about the teachers was given.

Test-Retest Reliability. *How consistent are scores if the assessment is administered once and then administered again soon? What about much later?*

Scores on all four of the LAP-D domains met the criteria for acceptable consistency when the assessment was administered twice (one to three weeks apart).This was examined with 155 Spanish-speaking children who were representative of the 2005 sample of children.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

Overall, the items in the LAP-D Spanish that are intended to reflect the same set of skills or behaviors met the criteria for acceptable relationships. However, when broken down by age group, some of the relationships were weak, specifically, for Spanish-speaking children aged 66-72 months in the fine motor domain. This was examined with the entire typically developing Spanish-speaking children in the sample.

Validity: Does the assessment measure what it is supposed to?

Content Validity. *Do experts agree that the items in the assessment do a good job of reflecting what the assessment is supposed to be assessing?*

The developers did not provide information about expert agreement.

Construct Validity. *How closely related to each other are sets of items within the assessment that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

When taking the age of the children into consideration, there are moderate relationships between the sets of items that aim to address similar skills and behaviors and, as expected, weak relationships between sets of items that aim to address different skills. For example, there was a weak relationship between items that addressed language skills and gross motor skills. Furthermore, items that addressed the fine motor skill of writing were not strongly related to items that asked about the fine motor skill of manipulating. This was examined with all of the children in the 2005 sample. The developers do examine whether scores on sets of items relate to children's age as expected.

Convergent Validity. *Is this assessment closely related to other well-established assessments aimed at measuring the same skills and behaviors?*

Overall, the LAP-D Spanish demonstrated strong relationships with these three assessments:

LAP-D Spanish and the Developmental Indicators for the Assessment of Learning, Spanish 3rd Edition (DIAL-3): A set of 112 children from the Spanish-speaking sample was given the LAP-D Spanish and the Spanish version of the DIAL-3 screener either at the same assessment session or within a very short period of time. The DIAL-3 screener is a comprehensive, norm-referenced screener of child development (for more information, see the DIAL-3 developmental screener profile in this compendium). Only the motor concepts and Language scales on the DIAL-3 were used in this study since conceptually, they were the most closely related to the LAP-D Spanish domains. The scores on the LAP-D Spanish and the DIAL-3 showed a strong relationship across most of the domains in the LAP-D Spanish. The strongest relationship was between the fine motor domain on the LAP-D Spanish and the motor domain on the DIAL-3. The developers did not provided additional information about these 112 children.

LAP-D Spanish and Batería Woodcock-Johnson-Munoz Revised (Batería-R): A set of 178 children from the Spanish-speaking sample was given the Batería-R assessment either at the same session or within a very short period of time. The Batería-R is a norm-referenced set of assessments used to evaluate cognitive abilities. Three domains from the Batería-R were used in this study: letter-word identification, applied problems, and dictation. While there were strong relationships between the scores on the LAP-D Spanish and the Batería-R, the relationships were not as strong as between the LAP-D Spanish and the DIAL-3. There was wide variation in the strengths of the relationships across the domains of the LAP-D Spanish and the Batería-R, but most fall into the high range. The only relationship that falls in the moderate range is between the LAP-D Spanish gross motor domain and the Batería-R letter-word identification domain. The developers did not provide additional information about the characteristics of the 178 children for whom these relationships were examined.

Validity: Does the assessment measure what it is supposed to?

LAP-D and Test de Vocabulario en Imágenes Peabody (TVIP): The full group of Spanish-speaking children (n=975) were administered the TVIP assessment either at the same session as the LAP-D Spanish or within a very short period of time. The TVIP is a norm-referenced assessment used to evaluate receptive vocabulary, which is the set of words that a person hears and can associate with a picture or definition. The comparison of scores demonstrated a strong relationship between the cognitive and language domains of the LAP-D Spanish and the TVIP assessment. The fine motor and gross motor domains were still in the high range, but were less strongly related to the TVIP scores, showing that those areas are more dissimilar to the skills the TVIP measures. It is to be expected that the non-language-related domains on the LAP-D would not be as strongly related to the TVIP since it is a language assessment.

Overall, these results suggest that the scores generated from the LAP-D Spanish are, for the most part, comparable to the scores on other well-established measures in the field that assess similar developmental skills. Spanish-speaking children who score high on the specific developmental skills captured by these other assessments should score high on the LAP-D Spanish and vice versa.

Predictive Validity. *How well does this assessment predict children's later academic achievement and adjustment to school? For what groups of children has this been examined?*

The developers have not examined how well this assessment predicts children's later academic achievement and adjustment to school.

Comments

It was surprising that when comparing the LAP-D Spanish to the TVIP, the relationships between the fine motor, cognitive, and gross motor domains were relatively strong. Since the fine motor cognitive and gross motor domains are less directly related to receptive vocabulary, it seems that the relationship should be weaker.

Overall, the strength of the relationships among domains and subscales in the LAP-D Spanish were weaker than in the LAP-D English.

References

Hardin, B.J., Peisner-Feinberg, E.S., & Weeks, S.W. (2005). Learning Accomplishment Profile Diagnostic Edition (Revised 3rd Edition): Examiner's Manual and Technical Report. Chapel Hill Training Outreach Project, Inc. Kaplan Early Learning Company.

Mullen Scales of Early Learning: AGS Edition*

Developer: Eileen M. Mullen, EdD
Publisher: Pearson, Inc.

pearsonassessments.com/mullen.aspx

Purpose:
The Mullen Scales of Early Learning: AGS Edition, an assessment of cognitive development for infants and preschool children, assesses a child's visual, linguistic, and motor abilities.

Developmental domains addressed in the assessment, as stated by the publisher:

- Visual reception
- Fine motor
- Gross motor
- Receptive language
- Expressive language

Intended age range:
The Gross Motor Scale is administered from birth through 33 months; each cognitive scale is administered from birth through 68 months.

Number of items:
Information on the number of items is not available.

Background

Is the assessment a direct assessment or an ongoing observational tool?

The Mullen Scales of Early Learning is a direct assessment.

If the assessment is a direct assessment, how long does it take to administer the measure?

The assessment typically can be administered in approximately 30 minutes for 3-year-olds, and 60 minutes for 5-year-olds. Administration time may vary depending on the age and other characteristics of the child.

Language(s) developed for:

The Mullen Scales of Early Learning were developed in English.

Availability and Cost of the Assessment

Is the assessment available to programs without restrictions?

Yes, the assessment is available to programs without restrictions.

What is the cost of the assessment?

As of 2010, the Mullen Scales of Early Learning Complete Kit cost $770.00. The full kit contains the manual, the item administration book, the stimulus book, and record forms (in packages of 25). The kit also contains several manipulatives that are needed to administer the scales. The manual contains a list of additional manipulatives that the administrator will need to supply. Additional record forms (in packages of 25) can be purchased separately for $39.90.

Training and Other Requirements for Assessors

Is training available on how to administer and score the assessment? Who offers the training?

Yes, training is available through Psych Corps, a division of Pearson, the publisher of the Mullen Scales. Some information is available online (http://www.pearsonassessments.com/pai/ca/training/training.htm), but the information is not specific to the Mullen Scales.

Is it necessary to have a professional background or technical training over and above training on the assessment to administer or complete the assessment?

According to the manual, the Mullen Scales were designed for use by professionals who have training and practical experience in the assessment of infants and young children. Potential qualified users include clinical psychologists, school psychologists, special educators, speech pathologists, occupational and physical therapists, physicians, nurses, and social workers.

Is it necessary to have a professional background or technical training over and above training on the assessment to score the assessment?

As stated above, the Mullen Scales should be administered by someone with training and experience in the assessment of young children. The manual and the item administration book give detailed information about how to correctly score the assessment.

Are regular checks on faithful administration required or recommended? If so, when and by whom?

The developer does not provide information regarding regular checks on faithful administration.

Information Reporting System for the Assessment

Electronic Data Entry. *Does the assessment come with a process for entering information from the assessment electronically?*

Mullen ASSIST computer software can be purchased additionally for $228.00. The software helps administrators calculate and convert raw scores. Scores from the Mullen Scales can be reported as standard scores, percentile ranks, age equivalents, descriptive categories, and developmental stages. Mullen ASSIST also provides a personal information summary, a score summary, an early learning composite[1], a score narrative, and recommended activities.

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

Programs can generate reports at the level of the individual child only. The ASSIST software program for use with the Mullen Scales allows users to input individual item scores, and the program calculates overall individual scores and provides interpretive information for the five Mullen scales and the early learning composite. The ASSIST offers several report options, including the opportunity to write the report into a word-processing program so the results can be incorporated into another document.

Child Outcomes Framework. *Does the publisher map the domains in the assessment onto the domains in the Head Start Child Outcomes Framework (HSCOF)?*

No, the domains on the Mullen Scales are not mapped onto the domains in the HSCOF.

Instructional Support. *Does the assessment provide guidance about follow-up steps teachers can use to help individual children progress?*

Yes, the manual gives some general suggestions regarding follow-up steps teachers can use for individual children. Assessment results can be used to help develop Individual Education Plans (IEPs) and can provide information about activities and reinforcement to be used in Individual Family Service Plans (IFSPs). However, there are no activities or programs indicated for improving visual, language, or motor development (the areas of development covered by this assessment) tailored to specific scores.

Planning Support. *Does the assessment come with guidance to help teachers determine follow-up steps involving planning at the classroom level?*

No, the assessment does not come with guidance for planning at the classroom level.

Administrator Support. *Does the assessment come with guidance to help program administrators plan follow-up steps involving program improvement?*

No, the assessment does not come with guidance for program improvement.

[1]The early learning composite is a summary score for all four cognitive scales.

Approaches to Family/Parent Input

Tools for Family Input. *Does the assessment include specific tools and/or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

No, the assessment does not include specific tools or guidance for incorporating parental/family input on a child's skills and development.

Sharing Results. *Does the assessment include recommendations on how to share assessment results with a child's family?*

No, the assessment does not include recommendations on how to share assessment results with a child's family.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this an assessment with developmental norms?*

Yes, the Mullen Scales is an assessment with developmental norms. The norms were developed in two phases over a period of eight years. From June 1981 to February 1986, the developers collected a sample of children from the Northeast. Between December 1987 and April 1989, the developers obtained samples from the other regions of the country. This sample is described below and in the table on the following pages.

Which populations were included in the norming sample?

The norming sample for the Mullen Scales included a nationally representative sample of 1,849 children ranging in age from 2 days to 69 months. The sample was obtained from more than 100 sites in four geographic regions in the United States: Northeast, South, West, and North and South Central. Testing took place in kindergartens, day care centers, nursery programs, and home settings, in both urban and rural areas. All children in the sample came from homes where English was the primary language. Children with known physical or mental disabilities were not included in the standardization sample. The developer also used this sample to examine reliability and validity.

Availability of Versions in Languages Other than English. *Is the assessment available in languages other than English? Which languages?*

This assessment is not available in languages other than English.

How were versions in languages other than English developed?

This assessment is not available in languages other than English.

What are the findings on the reliability and validity of versions of the assessment in languages other than English?

This assessment is not available in languages other than English.

Options for Use with Special and Diverse Populations

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

The Mullen Scales of Early Learning include guidelines for assessing children who have attention problems that appear modality-related (verbal, visual, tactile-kinesthetic). For example, verbal directions can be condensed when a child has language-processing difficulties. The assessment examiner also can provide a felt board to define a child's working space if the child has problems processing visual information. The developer does not provide additional information regarding specific accommodations for children with other special needs.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

No, the developer has not examined the appropriateness of this assessment for diverse populations in this way.

Characteristics of Norming Sample (data collected 1981-1986 and 1987-1989)
Number of children in the sample: 1,849

	Percentage of Children
Age of Children	
1-2 months	6.0
3-4 months	5.9
5-6 months	5.5
7-8 months	5.2
9-10 months	5.3
11-12 months	4.5
13-14 months	4.8
15-20 months	7.1
21-26 months	7.3
27-32 months	8.4
33-38 months	6.4
39-44 months	6.8
45-50 months	7.4
51-56 months	7.3
57-62 months	5.6
63-68 months	5.8
Gender	
Male	51.3
Female	48.7
Racial/Ethnic Group	
White	75.2
African American	14.4
Hispanic	7.4
Other	3.0
Father's Occupation	
Professional	26.1
Technicians, Sales, Small Business	21.3
Skilled and Semi-Skilled Manual	32.3
Unskilled Manual and Laborers	20.3
Geographic Region	
South	25.0
Northeast	39.6
West	20.1
North and South Central	15.2
Community Size	
Urban	81.6
Rural	18.4

Reliability and Validity Information

What is known about the reliability and validity of the assessment…

In English?

There is information about reliability and validity of the assessment in English. This information is outlined in later sections of this profile.

In other languages?

The Mullen Scales of Early Learning are not available in other languages.

For dual language learners?

The developer does not provide any information about dual language learners and has not examined the reliability and validity of the assessment for this population.

For children with special needs?

The developer provides evidence that the assessment predicts to later achievement for children with special needs. Mullen, Freeman, and Merenda (1987) conducted a two-year study of children with a diagnosis of learning disability or developmental delay to examine how well the Mullen Scales of Early Learning predict children's readiness for school. The assessment results showed that the Mullen visual reception and fine motor scores were closely related to the visual, prereading, and quantitative scores of the Metropolitan Readiness Test (MRT), 4th Edition, Form P, Level 2, an assessment used to measure preschool and kindergarten performance. Mullen receptive and expressive language scores were closely related to MRT language, prereading, and quantitative scores. The developer has not examined other types of reliability and validity for this population.

For American Indian/Alaskan Native children?

The developer does not provide any information about American Indian/Alaskan Native children and has not examined the reliability and validity of the assessment for this population.

For children of migrant and seasonal farm workers?

The developer does not provide any information about children of migrant and seasonal farm workers and has not examined the reliability and validity of the assessment for this population.

Reliability: Does the assessment obtain the same results, consistently, under the same conditions with the same children?

Inter-Rater Reliability. *Do different raters agree when they are assessing the same children?*

Yes, two raters who have been adequately trained on the assessment should get similar scores when administering the Mullen Scales to the same child. To test whether children get the same score on the Mullen Scales when being assessed by different raters, two raters alternated roles of administrator/scorer and scorer: One rater administered the assessment to a child, but each rater scored the assessment independently, yielding two sets of scores for the one child. The test sample consisted of a subset of 181 children from the larger norming sample. The children, aged 1 to 44 months, were divided into four age groups: 0 to 6 months, 7 to 12 months, 13 to 24 months, and 25 to 44 months. The developers do not provide additional information about the characteristics of the children in this sample. The developers do not provide information about the assessment raters. The results indicate acceptable consistency between scores (for children up to 44 months) when different raters are assessing the same children.

Test-Retest Reliability. *How consistent are scores if the assessment is administered once and then administered again soon?*

To test how consistent scores are over time, the Mullen Scales of Early Learning were administered twice to 47 children aged 25 to 56 months. These children were a subset of the larger norming sample. The developers do not provide additional information about the characteristics of the children in this sample. The developers do not provide information about the assessment administrators. The time in between tests was about 1 to 2 weeks, with an average of 11 days. The results show acceptable consistency of scores in all of the cognitive domains over this period of time.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

The relationship among items within each of the five Mullen Scales was examined with a subset of 1,737 children from the larger norming sample. For children ages 33 to 68 months, the relationships among items intended to reflect the same set of skills or behaviors are acceptable, except in the visual reception domain for children ages 51 to 56 months, which has a weak relationship among items. This indicates that the items in this assessment are largely measuring related skills and behaviors.

Validity: Does the assessment measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the assessment do a good job of reflecting what the assessment is supposed to be assessing?*

The developers did not provide information regarding expert consultation.

Construct Validity. *How closely related to each other are sets of items within the assessment that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

For children ages 33 to 68 months, relationships between sets of items on the Mullen Scales of Early Learning intended to reflect similar skills or behaviors are closely related, and items intended to reflect different skills or behaviors are less closely related. Relationships between the receptive and expressive language domains are high, while relationships between the fine motor domain and the other cognitive domains are lower. This indicates that the sets of items have strong relationships where expected. The developers do examine whether scores on sets of items relate to children's age as expected.

Convergent Validity. *Is this assessment closely related to other well-established assessments aimed at measuring the same skills and behaviors?*

A sample of 65 children between 15 and 59 months were given the Mullen Receptive language and expressive language scales, together with the measures of auditory comprehension and verbal ability from the Preschool Language Assessment. Receptive language was more closely related to auditory comprehension, while expressive language was more closely related to verbal ability (a measure of expressive language). The results show that the Mullen Scales have strong relationships to established measures aimed at measuring similar skills and behaviors (for children up to 59 months).

Predictive Validity. *How well does this assessment predict children's later academic achievement and adjustment to school?*

Mullen, Freeman, and Merenda (1987) conducted a two-year study of children with a diagnosis of learning disability or developmental delay to examine how well this assessment predicts children's readiness for school. A sample of 131 children from the Warwick (Rhode Island) Child Find Preschool special education program were administered the Mullen Scales at 4 or 5 years of age. After an average of 12 months, the children also were administered the Metropolitan Readiness Test (MRT), 4th Edition, Form P, Level 2, an assessment used to measure preschool and kindergarten performance. The results of these assessments showed that the Mullen visual reception and fine motor scores were closely related to the MRT visual, prereading, and quantitative scores. Mullen receptive and expressive language scores were closely related to MRT language, prereading, and quantitative scores. These results provide evidence of the Mullen Scales as predictors of preschool and kindergarten achievement only for children with special needs. The developer has not examined the extent to which the Mullen Scales predict later academic achievement and adjustment to school for other populations.

References

Mullen, E.M. (1995). Mullen Scales of Early Learning Manual, AGS Edition. Minneapolis, MN: NCS Pearson, Inc.

Mullen, E.M., Freeman, A., & Merenda, P.F. (1987). [A study of the predictive validity of the Mullen Scales of Early Learning]. Unpublished data.

Work Sampling for Head Start

Developers: Margo L. Dichtelmiller, Judy R. Jablon, Samuel J. Meisels, Dorothea B. Marsden
Publisher: Pearson Clinical Assessment

http://www.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=PAworksampl&Mode=summary

Purpose:
The purpose of the Work Sampling for Head Start is to help teachers document and evaluate children's skills, knowledge, and behaviors using actual classroom experiences and interactions.

Developmental domains addressed in the assessment, as stated by the publisher:
See right.

Intended age range:
3-5 years

Number of items:
Work Sampling for Head Start contains 51 items for 3-year-olds and 55 items for 4-year-olds.

Background

Is the assessment a direct assessment or an ongoing observational tool?

The Work Sampling System is an ongoing observational tool.

Language(s) developed for:

The assessment was developed for English-speaking children, but some materials are also available in Spanish.

Domains:

- Social and emotional development
  - Self concept
  - Self control
  - Cooperation
  - Social relationships
  - Knowledge of families and communities
- Approaches to learning
  - Initiative and curiosity
  - Engagement and persistence
  - Reasoning and problem-solving
- Language development
  - Listening and understanding
  - Speaking and communicating
- Literacy
  - Book knowledge and appreciation
  - Print and alphabet awareness
  - Early writing
  - Patterns
  - Measurement
- Mathematics
  - Problem-solving
  - Number and operations
  - Geometry and spatial sense
- Science
  - Scientific skills and methods
  - Scientific knowledge
- Creative arts
  - Music
  - Art
  - Movement
  - Dramatic play
  - Appreciation
- Physical health and development
  - Gross motor skills

Availability and Cost of the Assessment

Is the assessment available to programs without restrictions?

Yes, the assessment is available to programs without restrictions.

What is the cost of the assessment?

As of 2010, the Work Sampling System for Head Start kit (P3 and P4) cost $156.95. The kit includes the teacher's manual, guidelines, 30 checklists, rubrics, wall charts, and reproducible family report and documentation masters for both 3- and 4 year-olds. Additional checklists can be purchased (in packs of 10) for $27.75. The Work Sampling System for Head Start Kindergarten kit cost $156.95. This kit includes the teacher's manual, guidelines, 30 checklists, rubrics, wall charts, and reproducible family report and documentation masters for 5-year-olds. Additional checklists can be purchased (in packs of 10) for $24.00. Internet-based Work Sampling Online is also available.

Training and Other Requirements for Assessors

Is training available on how to administer and score the assessment? Who offers the training?

Yes, training is available on how to complete and score the Work Sampling System. There is also training on the Work Sampling Online System. Training is offered by Psych Corps, which is within Pearson Education, Inc., the publisher of the assessment. More information can be found on the publisher's website (http://www.worksamplingonline.com/School/Home/Info/Services.cfm).

Is it necessary to have a professional background or technical training over and above training on the assessment to administer or complete the assessment?

A teacher can complete the Work Sampling System and it is not necessary to have a technical background. The Teacher's Manual contains information about how to use the Developmental Guidelines and Checklists, how to prepare various observation methods and tools, how to set up an organized storage system, as well as guidance about making final ratings.

Is it necessary to have a professional background or technical training over and above training on the assessment to score the assessment?

A teacher can score the Work Sampling System without a technical background.

Are regular checks on faithful administration required or recommended? If so, when and by whom?

The developer states that supervision by those experienced in the use of Work Sampling is important.

Information Reporting System for the Assessment

Electronic Data Entry. *Does the assessment come with a process for entering information from the assessment electronically?*

The Work Sampling System for Head Start contains checklists that are scored with paper and pencil. There is also a Work Sampling System that is available online. The annual license fee for the electronic version is $19.95 for programs with fewer than 100 children. The cost of the online Work Sampling System for larger programs can be obtained from the publisher.

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

Yes, Work Sampling Online allows users to create summary reports for individual children. These include the Head Start Family Report, which is shared with parents/family at conferences. Work Sampling Online suggests ratings for performance (based on the final checklists) and provides narratives from the teacher's comments on the checklists. Once the information for each child has been entered, Work Sampling Online generates a class profile, a report that shows the groups of children at each level in a classroom. Teachers then can plan activities based on children's individual needs or by those indicators with a not applicable (N/A) or did not observe (DNO) rating. Teachers also can generate reports at the classroom level or by group type (e.g., all 4-year-olds or all girls). They also can produce outcome reports that summarize patterns of progress on the Work Sampling System to identify areas for program improvement.

Child Outcomes Framework. *Does the publisher map the domains in the assessment onto the domains in the Head Start Child Outcomes Framework (HSCOF)?*

Yes, the Work Sampling for Head Start assessment is closely aligned to the HSCOF. The Work Sampling for Head Start Teacher's Manual contains an appendix that shows the alignment of each Work Sampling for Head Start item with the original items provided in the Child Outcomes Framework.

Instructional Support. *Does the assessment provide guidance about follow-up steps teachers can use to help individual children progress?*

Yes, the manual gives some general suggestions regarding follow-up steps teachers can use for individual children. However, there are no specific activities or programs indicated for improving certain areas of development. Results from the Work Sampling can also be used in the development of Individual Education Plans (IEPs) required for children with special needs.

Planning Support. *Does the assessment come with guidance to help teachers determine follow-up steps involving planning at the classroom level?*

No, the Work Sampling System for Head Start does not come with guidance for planning at the classroom level.

Administrator Support. *Does the assessment come with guidance that helps program administrators plan for follow-up steps involving program improvement?*

No, the Work Sampling System for Head Start does not come with guidance for program improvement.

Approaches to Family/Parent Input

Tools for Family Input. *Does the assessment include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

No, Work Sampling for Head Start does not include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development.

Sharing Results. *Does the assessment include recommendations on how to share assessment results with the child's family?*

Yes, Work Sampling for Head Start includes a family report. This report is designed to summarize the information on the Checklist and supplement it with teachers' comments. There are two versions of the Family Report. The standard form includes ratings and teacher comments. The narrative family report is entirely narrative and does not include any ratings. These forms are produced on carbonless paper in three copies: one for the family, one for the teacher, and one for the child's cumulative file. They are also available electronically and in both English and Spanish.

The manual for Work Sampling for Head Start includes guidance on completing the family report, including how to fill in identifying information, how to write comments, how to proofread the reports, and how to share the information with families at a conference.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this an assessment with developmental norms?*

Work Sampling for Head Start is not an assessment with developmental norms. Instead, the Work Sampling for Head Start Guidelines are criterion-referenced and describe reasonable expectations for children at a given age. These criteria are based on information from the Head Start Outcomes Framework, as well as national curriculum groups, and child development research.

Which populations are included in the norming sample?

The Work Sampling System does not have a norming sample.

Availability of Versions in Languages Other than English. *Is the assessment available in languages other than English? Which languages?*

Work Sampling for Head Start is not available in languages other than English, however question and answer report forms for families are available in Spanish.

How were versions in languages other than English developed?

Work Sampling for Head Start is not available in languages other than English.

What are the findings on the reliability and validity of versions of the assessment in languages other than English?

Work Sampling for Head Start is not available in languages other than English.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

Modifications to the work sampling checklist may include interpreting items or deleting certain items. The manual also suggests supplementing this assessment with other more specialized assessments for children with special needs.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

The developers did not provide any information regarding cognitive testing or focus groups conducted with diverse populations.

Reliability and Validity Information

What is known about the reliability and validity of the assessment…

In English?

The reliability and validity of this assessment were tested in a study (published in 2008) of children enrolled in the St. Paul Public Schools CHOICE program, an Early Reading First federally funded project. Participants in the study were enrolled in three School Readiness classrooms; 12 Head Start classrooms in a single center; and one community-based classroom operated by the YMCA. The study sample included 59 participants in 2004-2005 (53% enrolled in Head Start) and 53 in 2005-2006 (60% enrolled in Head Start). The age range for the children was 3.77 to 4.90 years. The sample consisted of slightly more boys (54.5%) than girls. The breakdown of the sample by race/ethnicity was 62.5% Black, 8.9% Hispanic, and 8.9% Other. Most of the children received free or reduced-price meals (94.6%). Children with special needs constituted 11.6% of the sample. The reliability and validity findings for this sample are outlined in later sections of this profile.

In other languages?

The Work Sampling System for Head Start is not available in other languages.

For dual language learners?

The developers do not provide any information about dual language learners and have not examined the reliability and validity of this assessment for this population.

For children with special needs?

While children with special needs were included in the sample of children in the reliability and validity study (11.6% of children), the developers have not examined the reliability and validity of the assessment for this population separately.

For American Indian/Alaskan Native children?

The developers do not provide any information about American Indian/Alaskan Native children and have not examined the reliability and validity of this assessment for this population.

For children of migrant and seasonal farm workers?

The developers do not provide any information about children of migrant and seasonal farm workers and have not examined the reliability and validity of this assessment for this population.

Reliability: Does the assessment obtain the same results, consistently, under the same conditions with the same children?

Inter-Rater Reliability. *Do different raters agree when they are assessing the same children?*

The developers have not examined the agreement between raters when they are assessing the same child.

Test-Retest Reliability. *How consistent are scores if the assessment is administered once and then administered again soon?*

The developers evaluated the Work Sampling for Head Start data for the language, literacy, and mathematics subscales both in the fall and in the spring with an interval of approximately 6.5 months. The length of time between administrations does not meet the criteria for test-retest reliability (3 months or less).

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

Meisels et al. (2008) examined relationships between items within the language, literacy, and mathematics subscales with the sample of children from the St. Paul Public Schools CHOICE Program (described above). The relationships between items within these subscales were acceptable. The relationships between items in other domains were not reported on in this study.

Validity: Does the assessment measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the assessment do a good job of reflecting what the assessment is supposed to be assessing?*

Yes, teachers and child development experts conducted content analysis throughout the development of Work Sampling.

Construct Validity. *How closely related to each other are sets of items within the assessment that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

The developers do not provide any information about the relationship between sets of items that address similar and different skills and behaviors or on how scores on sets of items relate to child age.

Convergent Validity. *Is this assessment closely related to other well-established assessments aimed at measuring the same skills and behaviors?*

The Test of Early Reading Ability-Third Edition (TERA-3) and the Test of Early Mathematics Ability (TEMA-3) were used to examine how closely related Work Sampling for Head Start is to other well-established assessments aimed at measuring the same skills and behaviors. The TERA-3 is an individually administered assessment of young children's reading achievement. It includes subtests in alphabet, conventions, and meaning, and was designed for use with children aged 3 years, 6 months to 8 years, 6 months. The TEMA-3 is an individually administered assessment of early mathematical achievement that is appropriate for children aged 3 years through 8 years, 11 months. The assessment focuses on knowledge of counting, calculation, conventions, and number facts.

The TERA-3 and TEMA-3 were administered to the Meisels et al. (2008) study sample to establish the reliability and validity of Work Sampling for Head Start in relation to normative assessments of achievement. The results show that all of the relationships were within a moderate range. The relationships between the measures in mathematics were lower than in language and literacy.

Predictive Validity. *How well does this assessment predict children's later academic achievement and adjustment to school?*

The developers do not provide any information about how well this assessment predicts children's later academic achievement or adjustment to school.

References

Dichtellmiller, M.L., Jablon, J.R., Meisels, S.J., & Marsden, D.B. (2001). Work sampling for Head Start: A teacher's manual. New York: Pearson Education Inc.

Meisels, S.J., Atkins-Burnett, S., Xue, Y., Nicholson, J., Bickel, D.D., & Son, S. (2003) Creating a system of accountability: The impact of instructional assessment on elementary children's achievement test scores. *Education Policy Analysis Archives, 11*(9).

Meisels, S.J., Xue, Y., Shamblott, M. (2008). Assessing language, literacy, and mathematics skills with Work Sampling for Head Start. *Early Education and Development*, *19* (6), 963-981.

Developmental Screener Profiles

Ages and Stages Questionnaires-3rd Edition (ASQ-3)

Developers: Jane Squires and Diane Bricker
Publisher: Paul H. Brookes Publishing Co., Inc

www.agesandstages.com

Developmental domains addressed in the developmental screener, as stated by the publisher:

- Gross motor
- Fine motor
- Problem solving
- Personal-social
- An overall section addresses general parental concerns.

Intended age range: 1-66 months

Number of items: Each of the 21 questionnaires contains 30 items. There is also an overall section addressing general parental concerns.

In what settings can this developmental screener be used (e.g., centers, homes, medical facilities, other)? Settings in which the ASQ-3 can be used include screening clinics, education and child care facilities, home settings, and doctors' offices or clinics.

Background

Purpose:

The Ages and Stages Questionnaires-3rd Edition (ASQ-3) is a developmental screening system made up of 21 age-specific questionnaires completed by parents or primary caregivers of young children (16 questionnaires are for use with children under the age of 3). The questionnaires can identify children who are in need of further assessment to determine whether they are eligible for early intervention or early childhood special education services. Where possible, this profile will focus on information related to the 36-, 42-, 48-, 54-, and 60-month questionnaires (3-5 years).

What is the appropriate time period between administering, recording, or reviewing the data?

The ASQ-3 manual suggests that it is ideal to screen children at regular intervals, from 2 months to 5 years, 6 months, if possible. Ideally, children should be screened initially at 2 and 4 months, then at 4-month intervals until they are 24 months old, and at 6-month intervals until they are 5 years, 6 months old. The developers do not recommend screening children more frequently than every 4-6 months (except at the 2- and 4- month intervals) unless there is some reason to suggest that more frequent screening would be useful (e.g., the child has suffered a serious illness, parents feel their child has changed, etc.).

How long does it take to administer the developmental screener?

The ASQ-3 questionnaires are completed by parents. Each questionnaire can be completed in 10-15 minutes.

Language(s) developed for:

The ASQ-3 was developed in English and translated into Spanish. Earlier editions of the ASQ are available in French and Korean. Translations of the ASQ-3 are in development in a number of languages; however, the developers did not provide information about which languages will be available.

Availability and Cost of the Developmental Screener

Is the developmental screener available to programs without restrictions?

Yes, the ASQ-3 is available to programs without restrictions.

What is the cost of the developmental screener?

As of 2010, the ASQ-3 Starter Kit, which includes 21 paper masters of the questionnaires (in English or in Spanish), scoring sheets, a CD-ROM with printable PDF questionnaires, the ASQ-3 User's Guide, and a laminated ASQ-3 Quick Start Guide, cost $249.95. The starter kit contains all 21 questionnaires, however only five of these questionnaires are for use with 3-5 year olds. Additional copies of the 21 questionnaires (in English or in Spanish) can be purchased separately for $199.95.

Training and Other Requirements for Assessors

Is training available on how to administer and score the developmental screener? Who offers the training?

Yes, ASQ-3 training is available through the publisher. Training DVDs are available that show staff how to screen, score, and interpret the results of the ASQ-3. Programs may also arrange for onsite seminars or attend the training seminars held every year by the developers of ASQ-3. Detailed information is available on the company's website (http://www.agesandstages.com/training/).

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to administer or complete the developmental screener?

No, it is not necessary to have a professional background or technical training to complete the ASQ-3. The ASQ-3 was developed as a parent-completed screening tool, and having parents and caregivers complete the screener is the preferred method. Completing a questionnaire independently requires reading skills at a 4th- to 6th-grade reading level. If parents or caregivers are unable to complete questionnaires independently (due to cognitive disability, limited reading skills, etc.), teachers and program staff can provide support. The manual does suggest that all ASQ-3 users become familiar with the information in the manual, in particular, the information regarding administering the ASQ-3 which appears in chapter 6.

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to score the developmental screener?

A parent, caregiver, or teacher can score the ASQ-3 without a professional background or technical training. The manual does suggest that ASQ-3 users become familiar with the information in the manual, in particular the information regarding scoring the ASQ-3.

Are regular checks on administration required or recommended to ensure appropriate administration? If so, when and by whom?

There is no information in the manual regarding the performance of regular checks on administration.

Information Reporting System for the Developmental Screener

Electronic Data Entry. *Does the developmental screener come with a process for entering information from the screener electronically?*

Yes, the ASQ-3 can be used with online systems called the ASQ Pro (for single sites) and the ASQ Enterprise (for multisite programs). These online management systems help with screening administration, automated scoring, and information storage. An annual subscription to the ASQ Pro costs $149.95. An annual subscription to the ASQ Enterprise costs $499.95

Electronic Reports. *Can programs generate electronic reports of individual children's data?*

Yes, the ASQ Pro and the ASQ Enterprise online systems can store questionnaire results and follow-up decisions in individual child records. The ASQ Enterprise can also generate multisite reports to show trends across programs.

Approaches to Family/Parent Input

Tools for Family Input. *Does the developmental screener include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

The ASQ-3 questionnaires were designed to be completed by parents. They indicate "yes," "sometimes," or "not yet" regarding whether the child exhibits certain skills or behaviors within five areas: communication, gross motor, fine motor, problem-solving, and personal-social. The final overall section provides space for parents and caregivers to note any general concerns.

Sharing Results. *Does the developmental screener include recommendations on how to share developmental screener results with the child's family?*

Yes, the ASQ-3 manual gives suggestions about how to communicate results of the screening with families. There are suggestions for families of children whose scores indicate typical development and for children whose results indicate the need for further assessment. An example of a feedback letter for parents and caregivers of children whose scores indicate typical development is found in Appendix D (in English and in Spanish) of the manual. The manual suggests that providing feedback to families with children whose scores indicate the need for further assessment should always be done in person due to the sensitive nature of the conversation.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this a developmental screener with developmental norms?*

Yes, the ASQ-3 is a screener with developmental norms. The sample on which the norms are based included 15,138 children and their families, and is representative of the U.S. population in geography and ethnicity, and includes representation across socioeconomic groups.

Which populations are included in the norming sample?

Norms for the ASQ-3 were developed using questionnaire data collected between January 2004 and June 2008. This norming sample was 53 percent male and 47 percent female. 54 percent of mothers in the sample had at least four years of college, whereas 12 percent had an associate's degree, 23 percent had a high school education, and 3.5 percent had not completed high school. The majority of the reporting caregivers for this sample indicated incomes greater than $40,000. See the table on the next page for more information about these children.

Availability of Versions in Languages Other than English. *Is the developmental screener available in languages other than English? Which languages?*

The ASQ-3 is available in Spanish. Previous editions of the ASQ are available in French and Korean.

How were versions in languages other than English developed?

In order to develop the Spanish translation of the ASQ-3, pediatric experts, developmental pediatricians, and practitioners working with young children and families who speak a variety of Spanish dialects reviewed the Spanish-language version of the second edition of the ASQ. Translation errors that were found in the second edition were corrected and minor wording changes were made.

The ASQ-3 Spanish questionnaires have been tested with Spanish-speaking parents in various geographic regions of the United States; however, separate cutoff scores have not been developed for children of Spanish-speaking parents.

What are the findings on the reliability and validity of versions of the developmental screener in languages other than English?

The reliability and validity of the translations of the ASQ-3 have not been examined.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

The ASQ-3 is based on parent, family, or teacher report and therefore the developers do not provide information regarding accommodations for screening children with identified or suspected special needs.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted with diverse populations to determine the appropriateness of this developmental screener for these populations?*

The developers do not provide information about whether the appropriateness of the ASQ-3 for diverse populations was addressed in this way.

Risk Levels. *What terminology is used to describe risk levels (e.g., delay, no delay, at risk, caution, rescreen, okay, etc.)?*

The risk levels on the ASQ-3 are described as "typical development," "need for monitoring," or "need for further assessment."

Characteristics of 2008 Norming Sample
Number of children in the sample: 15,138

	Percentage of Children
Race/Ethnicity	
White	66.4
African American	11.6
Latino/Hispanic	10.5
Asian/Pacific Islander	3.9
Native American/Alaskan	1.1
Other	1.1
Mixed	4.5
Unknown	0.9
Gender	
Male	52.6
Female	47.4
Maternal Education	
Less than High School Graduation	3.5
High School Graduation	22.7
Associate's Degree	12.0
4 Years of College or Above	54.0
Unknown	7.7
Family Income	
$0-$12,000	12.8
$12,001-$24,000	9.3
$24,001-$40,000	13.7
More than $40,000	57.1
Unknown	7.0

Reliability and Validity Information

What is known about the reliability and validity of the developmental screener…

In English?

There is information about reliability, validity, sensitivity, and specificity of the ASQ-3 in English. This information is outlined in responses to later questions in this profile.

In other languages?

While the ASQ-3 has been translated into Spanish, the developers have not examined the reliability, validity, sensitivity, and specificity for the Spanish version of the ASQ-3.

For dual language learners?

The developer does not provide any information about this population, and the reliability, validity, sensitivity, and specificity of the ASQ-3 for dual language learners have not been examined.

For children with special needs?

There is information about the sensitivity and specificity of the ASQ-3 for children with special needs. The extent to which the ASQ-3 correctly identifies children at risk for developmental delays was examined with a sample of 257 children participating in early intervention or early childhood special education programs in California, New York, and Oregon. The results of the screenings suggest that the ASQ-3 is moderately accurate at correctly identifying children who are at risk for developmental delays. Additionally, the extent to which the ASQ-3 correctly identifies children not at risk for developmental delays was examined with the same sample. Results show that the ASQ-3 is moderately to highly accurate in correctly identifying children who are not at risk for developmental delays. The developers did not provide additional information about the characteristics of this sample. The developers have not examined other types of reliability and validity for this population.

For American Indian/Alaskan Native children?

While American Indian/Alaskan Native children were included in the norming sample (1.1 percent of children), there is no separate information about the reliability, validity, sensitivity, and specificity for this specific group.

For children of migrant and seasonal farm workers?

The developer does not provide any information about this population, and the reliability, validity, sensitivity, and specificity of the ASQ-3 for children of migrant and seasonal farm workers have not been examined.

Reliability: Does the instrument obtain the same results, consistently, under the same conditions with the same children?

Interrater Reliability. *Do different raters agree when they are assessing the same children?*

To test whether different raters agree when they are screening the same children, questionnaires completed by parents were compared with questionnaires completed by trained test examiners for the same children. The results showed acceptable agreement between parents and trained examiners when completing the ASQ-3 for the same children. The strongest agreement was in the personal-social area and the weakest agreement was in the communication area. This may be due to parents and test examiners observing different types of behavior in different settings while completing the communication area.

The agreement between raters was examined with 107 children based on the parents' and examiners' completion of the ASQ-3. This sample was taken from the norming sample. The developers did not provide specific information about the characteristics of the children in this analysis. The developers did not provide demographic information on the trained examiners.

Test-Retest Reliability. *How consistent are scores if the developmental screener is administered once and then administered again soon?*

The consistency of scores on the ASQ-3 is acceptable if it is administered once and then again soon. This was tested by comparing two questionnaires completed by the same parent at a two-week interval. Questionnaires completed by 145 parents taken from the norming sample were included in this analysis (no specific information is provided about the characteristics of this sample). Parents did not have access to the first questionnaire when they completed the second one, and did not know whether the scores indicated a need for further follow-up. The results of the comparisons of the two questionnaires show that the scores were consistent.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

The developers did not examine relationships between the items within a developmental area. However, the developers did examine the relationships between developmental area scores and overall scores on the ASQ-3. This information is summarized under "Construct Validity" in the next section of this profile.

Validity: Does the instrument measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the developmental screener do a good job of reflecting what the developmental screener is supposed to be assessing?*

Yes, experts, parents, and practitioners were consulted during the development of items for the ASQ-3.

Construct Validity. *How closely related to each other are sets of items within the developmental screener that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

The developers have not examined relationships between sets of items that aim to address similar skills and behaviors.

The developers did examine the relationships between developmental area scores and overall score on the ASQ-3 for 20 questionnaire age intervals. The results showed strong relationships between developmental area scores and overall ASQ-3 scores.

Information about whether scores on sets of items relate to children's age as expected is not provided.

Convergent Validity. *How strongly do the scores of this developmental screener show a relationship to the scores of other developmental screeners of similar domains?*

Please see response below to how accurately the developmental screener correctly identifies children at risk for developmental delays.

Scores for Further Evaluation. *Are specific scores used to identify whether further evaluation is needed? How are these scores determined by the developer?*

Yes, cutoff scores for the five areas of development covered in each questionnaire age interval have been determined using data from 18,572 questionnaires.

The manual indicates several different levels of cutoff scores that a program can choose to use when interpreting the scores.

Sensitivity. *How accurately does the developmental screener correctly identify children at risk for developmental delays?*

To test how accurately the ASQ-3 correctly identifies children at risk for developmental delays, both the ASQ-3 and the Battelle Developmental Inventory (BDI) were administered to two groups of children: those not receiving special education services and presumed to be developing without problems (322 children), and those participating in early intervention or early childhood special education programs in California, New York, and Oregon (257 children). The BDI was administered to both groups of children by trained examiners. The ASQ-3 was completed by parents or caregivers. The results of the screenings suggest that the ASQ-3 is moderately accurate at correctly identifying children at risk for developmental delays. The accuracy of identifying children at risk for developmental problems depends on the age of the children. For children ages 27-36 months, the ASQ-3 is 85.9 percent accurate at correctly identifying children at risk for developmental delays. For children ages 42-60 months, it is 82.5 percent accurate. Additionally, accuracy varies depending on which cutoff scores have been used.

Specificity. *How accurately does the developmental screener correctly identify children who are not at risk for developmental delays?*

The ASQ-3 is moderately to highly accurate at correctly identifying children who are not at risk for developmental delays. The accuracy of identifying children not at risk for developmental problems depends on the age of the children. For children ages 27-36 months, the ASQ-3 is 85.7 percent accurate at correctly indentifying children not at risk for developmental delays. For children ages 42-60 months, the ASQ-3 is 92.1 percent accurate. Additionally, accuracy varies depending on which cutoff scores have been used.

Follow-Up Guidance

Program Follow-Up Steps. *Does the developmental screener come with guidance about follow-up steps based on the results?*

Yes, the ASQ-3 Information Summary sheet provides a list of potential actions that may follow the screening, based on the child's scores and the parent's responses to the overall questions. For example, if the child's scores indicate typical development, children can be rescreened at 4- to 6-month intervals, and parents can be given suggestions for activities to do with their children to support their continued development. If a child's scores indicate the need for further assessment, a referral to a community agency or specialist may be made.

Family Follow-Up Steps. *Does the developmental screener include recommendations on how families might follow-up on the results of the screening?*

The recommended follow-up steps included on the ASQ-3 Information Summary Sheet provide recommendations for how families might follow up on the results of the screening. In addition, parents can use the activities that are included in the manual for children with typical results or for children who need monitoring and/or referrals. Children may benefit from practicing the skills targeted in these activities.

References

Squires, J., Twombly, E., Bricker, D., & Potter, L. (2009). Ages and Stages Questionnaires: Third Edition. Baltimore, MD: Paul H. Brookes Publishing.

Ages and Stages Questionnaires-Social-Emotional

Developers: Jane Squires, Diane Bricker, and Elizabeth Twombly
Publisher: Paul H. Brookes Publishing Co., Inc.
http://www.brookespublishing.com/store/books/squires-asqse/index.htm

Developmental domains addressed in the developmental screener, as stated by the publisher:

- Self-regulation
- Compliance
- Communication
- Adaptive functioning
- Autonomy
- Affect
- Interaction with people

Intended age range: 6-60 months

Number of items:
The ASQ-SE is a series of eight separate questionnaires based on age intervals:
6 months (19 items),
12 months (22 items),
18 months (26 items),
24 months (26 items),
30 months (29 items),
36 months (31 items),
48 months (33 items),
and 60 months (33 items).

In what settings can this developmental screener be used (e.g., centers, homes, medical facilities, other)?
The ASQ-SE can be used in home settings, clinical settings (e.g., primary health care clinics, immunization clinics, mental health clinics), center-based settings (e.g., child care, preschool), and other settings (e.g., health fairs, school screenings, community Child-Find activities).

Background

Purpose:

The Ages and Stages Questionnaires-Social Emotional (ASQ-SE) is a developmental screener designed to complement the Ages and Stages Questionnaires by providing information specifically addressing the social and emotional behavior of children.[1] The ASQ-SE identifies infants and young children whose social or emotional development requires further evaluation to determine if a referral for intervention services is necessary.

What is the appropriate time period between administering, recording, or reviewing the data?

The ASQ-SE is intended for use at six month intervals between 6 months and 3 years of age, and then at one year intervals through age 5.

How long does it take to administer the developmental screener?

The ASQ-SE questionnaires are completed by parents. The questionnaires take approximately 10-15 minutes to complete.

Language(s) developed for:

The ASQ-SE was developed in English and translated into Spanish.

[1] For more information see the Ages and Stages Questionnaire profile in this document.

Availability and Cost of the Developmental Screener

Is the developmental screener available to programs without restrictions?

Yes, the ASQ-SE is available to programs without restrictions.

What is the cost of the developmental screener?

As of 2010, a complete ASQ-SE Starter Kit costs $194.95. This kit contains everything needed to start screening children with the ASQ-SE: eight photocopiable print masters of the questionnaires and scoring sheets, a CD-ROM with printable PDF questionnaires, and the ASQ-SE User's Guide. The Starter Kit is also available with Spanish questionnaires. Additional master copies of the eight questionnaires (in English and Spanish) can be purchased separately for $149.95.

Training and Other Requirements for Assessors

Is training available on how to administer and score the developmental screener? Who offers the training?

Yes, training is available through the publisher on how to administer and score ASQ-SE. There are many different types of training available including onsite seminars and training by DVD. Detailed information is available on the company's website: http://www.agesandstages.com/training/.

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to administer or complete the developmental screener?

The original ASQ and ASQ-SE were developed as parent-completed screening tools, and it is best that parents or caregivers complete the screeners. However, child care providers, teachers, and early interventionists can also complete the ASQ-SE. Parents, caregivers, and teachers do not need to have technical training to complete the ASQ-SE.

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to score the developmental screener?

A parent, caregiver, or teacher can score the ASQ-SE without technical training.

Are regular checks on faithful administration required or recommended? If so, when and by whom?

The developer does not provide information regarding the performance of regular checks on faithful administration.

Information Reporting System for the Developmental Screener

Electronic Data Entry. *Does the developmental screener come with a process for entering information from the screener electronically?*

Yes, both the ASQ-SE and the ASQ-3 can be used with online systems called the ASQ Pro (for single sites) and the ASQ Enterprise (for multisite programs). These online management systems help with screening administration, automated scoring, and information storage. An annual subscription to the ASQ Pro costs $149.95. An annual subscription to the ASQ Enterprise costs $499.95

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

The ASQ Pro and the ASQ Enterprise create both individual child reports and program-level reports. The ASQ Enterprise can also generate multisite reports to show trends across programs.

Approaches to Family/Parent Input

Tools for Family Input. *Does the developmental screener include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

The ASQ-SE is designed to be completed by parents or caregivers.

Sharing Results. *Does the developmental screener include recommendations on how to share developmental screener results with a child's family?*

Yes, the ASQ-SE does include some recommendations on how to share the screening results with the child's family.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this a developmental screener with developmental norms?*

Yes, the ASQ-SE is a screener with developmental norms. The sample on which the norms are based included 3,014 preschool-age children and their families, and is representative of the U.S. population in terms of ethnicity, geographic region, parent education, income, and gender of children (based on 2000 U.S. Census data).

Which populations are included in the norming sample?

The ASQ-SE norming sample included 2,633 children whose families contributed at least one completed questionnaire and 381 whose families contributed two or more questionnaires at different age intervals (e.g., at 6 and 12 months). The children in the sample were between the ages of 3 and 66 months. See the table on the next page for more information about these children.

Availability of Versions in Languages Other than English. *Is the developmental screener available in languages other than English? Which languages?*

Yes, the ASQ-SE is available in Spanish. The reliability and validity of the Spanish questionnaires have not been examined.

How were versions in languages other than English developed?

The final English version of the ASQ-SE was translated into Spanish by Spanish-speaking staff from the Migrant Head Start program in Oregon. The Spanish translation was used with 153 children whose families were non-English speakers. These translated questionnaires were not used included in ASQ: SE reliability and validity tests.

What are the findings on the reliability and validity of versions of the developmental screener in languages other than English?

The developers have not examined the reliability and validity of the Spanish-language questionnaires.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

The ASQ-SE is based on parent observation; therefore, accommodations for children with identified or suspected special needs are not needed.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

Items for the ASQ-SE were assembled into a preliminary version called the Behavior-Ages and Stages Questionnaires (B-ASQ). Practitioners in approximately 50 programs across the United States used the B-ASQ with a diverse population of young children and parents. Practitioners and parents then completed questionnaires to provide feedback on the clarity of the meaning of the items and the appropriateness of the items, and suggestions for revisions and additions of items. This input was included in the final revisions of the B-ASQ, which was renamed the Ages and Stages Questionnaires-Social-Emotional (ASQ-SE).

Risk Levels. *What terminology is used to describe risk levels? (e.g., delay, no delay, at risk, caution, rescreen, okay, etc.)?*

Children are classified as "okay" (no further evaluation of social-emotional competence is indicated) or "at risk" (further evaluation of their social-emotional status is indicated).

Characteristics of the Norming Sample
Number of children in the sample: 3,014

	Percentage of Children
Race/Ethnicity	
White	58.9
African American	8.9
Hispanic	8.6
Asian/Pacific Islander	6.3
Native American	1.1
Mixed Ethnicity	16.0
Maternal Education	
Less than High School Graduation	13.0
High School Graduation or Equivalent	47.4
Associate's Degree	11.9
4-Year College or Above	25.3
Unknown	2.4
Family Income Level	
$0-$12,000	20.6
$12,001-$24,000	19.9
$24,001-$40,000	22.8
More than $40,000	29.9
Unknown	6.8

Reliability and Validity Information

What is known about the reliability and validity of the developmental screener…

In English?

There is information about reliability, validity, sensitivity, and specificity of the measure in English. This information is outlined in responses to later questions in this profile.

In other languages?

While the ASQ-SE has been translated into Spanish, the developers have not examined the reliability, validity, sensitivity, and specificity of the Spanish translation.

For dual language learners?

The developers do not provide any information about dual language learners and have not examined the reliability, validity, sensitivity, and specificity of the ASQ-SE for this population.

For children with special needs?

The developers do not provide any information about children with special needs and have not examined the reliability, validity, sensitivity, and specificity of the ASQ-SE for this population.

For American Indian/Alaskan Native children?

While American Indian/Alaskan Native children were included in the sample on which the screener was tested (1.1 percent of children), the developers have not examined the reliability, validity, sensitivity, and specificity of the ASQ-SE for American Indian/Alaskan Native children.

For children of migrant and seasonal farm workers?

The developers do not provide any information about children of migrant and seasonal farm workers and have not examined the reliability, validity, sensitivity, and specificity of the ASQ-SE for this population.

Reliability: Does the developmental screener obtain the same results, consistently, under the same conditions with the same children?

Interrater Reliability. *Do different raters agree when they are assessing the same children?*

The developers have not examined the agreement between raters when they are screening the same children.

Test-Retest Reliability. *How consistent are scores if the developmental screener is administered once and then administered again soon?*

The consistency of scores on the ASQ-SE if the screener is administered once and then again soon is acceptable. This was tested with a sample of 367 parents by comparing two questionnaires completed one to three weeks apart. No additional information about this sample of parents or their children is provided. The results showed that the ASQ-SE scores were consistent across time intervals.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

The relationships between items that are intended to reflect the same sets of skills or behaviors are acceptable. These relationships were examined for each ASQ-SE age interval described in the table on a previous page.

Validity: Does the developmental screener measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the developmental screener do a good job of reflecting what the developmental screener is supposed to be assessing?*

Yes, an interdisciplinary group of experts helped develop the items for the ASQ-SE.

Construct Validity. *How closely related to each other are sets of items within the developmental screener that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

The developers have not examined relationships between sets of items on the ASQ-SE that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors.

Information about whether scores on sets of items relate to children's age as expected is not provided.

Convergent Validity. *How strongly do the scores of this developmental screener show a relationship to the scores of other developmental screeners of similar domains?*

Please see response below to see how accurately the developmental screener correctly identifies children at-risk for developmental delays.

Scores for Further Evaluation. *Are specific scores used to identify whether further evaluation is needed? How are these scores determined by the developer?*

Yes, the developers used specific scores (called cutoff scores) to identify whether further evaluation is needed. To develop the cutoff scores, a method of analysis was used that compares the probability of getting an accurate result (indicating that the child is either "okay" or "at risk") for a range of cutoff scores.

A sample of 1,041 children with completed ASQ-SE questionnaires were then assessed with either the Child Behavior Checklist (CBCL), the Vineland Social-Emotional Early Childhood Scale (SEEC), or they had a professionally diagnosed social-emotional disability. The results of these screenings were then compared for this sample of children in order to determine appropriate cutoff points for the ASQ-SE.

Sensitivity. *How accurately does the developmental screener correctly identify children at risk for developmental problems?*

The ASQ-SE is moderately accurate at correctly identifying children at risk for developmental problems. To test this, children in the norm sample were classified as either "okay" or "at risk" based on their ASQ-SE scores, and classified as either "okay" or "at risk/disabled" using either the CBCL, the SEEC, or based on professional diagnosis. The results showed that the ASQ-SE and the CBCL (or SEEC) classified children the same way 78 percent of the time.

Specificity. *How accurately does the developmental screener correctly identify children who are not at risk for developmental problems?*

The ASQ-SE is highly accurate at correctly identifying children who are not at risk for developmental problems. This was determined using the comparisons between the ASQ-SE and the CBCL, the SEEC, or a professional diagnosis, described in the previous question. The results showed that the ASQ-SE and the CBCL (or SEEC) classified children the same way 94.5 percent of the time.

Follow-Up Guidance

Program Follow-Up Steps. *Does the developmental screener come with guidance about follow-up steps based on the results?*

Yes, the manual contains criteria that provide program staff with guidelines for how to interpret ASQ-SE scores and what types of follow-up are recommended. For example, if a child scores above the cutoff (indicating that there is the potential for a developmental delay or social/emotional concern), possible follow-up steps include:
1) Refer the child for diagnostic social-emotional or mental health assessment or
2) Provide the parent with information and support, and monitor the child using the ASQ-SE.

The manual also recommends that program staff look at other factors that may have influenced the results of the screening (e.g., setting/time of screening, the child's health, developmental factors, and family/cultural factors) and gather additional information before making a referral decision.

Family Follow-Up Steps. *Does the developmental screener include recommendations on how families might follow up on the results of the screening?*

The manual suggests providing parents with information or referrals to appropriate agencies for areas of concern. There is no additional information in the manual on how families might follow up on the results of the screening.

References

Squires, J., Bricker, D., & Twombly, E. (2003). The ASQ: SE User's Guide for the Ages and Stages Questionnaires: Social Emotional. Baltimore, MD: Paul H. Brookes Publishing.

Battelle Developmental Inventory, 2nd Edition- Screening Test

Developers: U.S. Office of Education, Columbus Laboratories of Battelle Memorial Institute
Publisher: Riverside Publishing http://www.riversidepublishing.com/products/bdi2/index.html

Developmental domains addressed in the developmental screener, as stated by the publisher:

- Adaptive
- Personal-social
- Communication
- Motor
- Cognitive

Intended age range: Birth through age 7

Number of items: The BDI-2 Screening Test contains 100 items. However, not all 100 items are used for every child. There are 2 items for each of the five developmental areas for 10 age groups. Thus, a child will most likely complete only 10 items. The Screening Test can be used to determine whether or not to administer the full BDI-2.

In what settings can this developmental screener be used (e.g., centers, homes, medical facilities, other)? The BDI-2 Screening Test can be used in a variety of settings, including Head Start Centers, programs funded for children with special needs, pre-schools, child care centers, homes, and medical facilities.

Background

Purpose:

The Battelle Developmental Inventory, 2nd Edition Screening Test (BDI-2 Screening Test) is a developmental screener that can be administered to get an initial snapshot of a child's development. The BDI-2 Screening Test is made up of items from the Battelle Developmental Inventory-2nd edition, which is a 450-item standardized assessment. The full assessment can be administered after the Screening Test if the administrator believes the child may be at risk for developmental delay. This profile will focus only on the Screening Test, not the full BDI-2 Assessment.

What is the appropriate time period between administering, recording, or reviewing the data?

Rescreening with the BDI-2 Screening Test could be done in as little as six months, especially if interventions or services have been put in place for a child.

How long does it take to administer the developmental screener?

The BDI-2 Screening Test can take 10 to 30 minutes, depending on the age of the child.

Language(s) developed for:

The Screening Test was developed for English, but there are also materials available in Spanish (more information on this is provided later in this profile).

Availability and Cost of the Developmental Screener

Is the developmental screener available to programs without restrictions?

In order to purchase the developmental screener, the person purchasing it must have the following (these restrictions do not necessarily apply to the person administering the developmental screener):

1. Certification as an occupational therapist, physical therapist, or another medical profession. Other medical professions include pediatricians, nurse practitioners, office nurses, visiting nurses, home health care workers for infants and young children, and Head Start specialists. Further information about these restrictions can be found on the publisher's website (see 1st page) or by calling the publisher.
2. Specific undergraduate-level training in one or more of the following: intelligence/cognitive testing, basic tests and measurements, speech, hearing, language assessments, education diagnostics, and developmental milestone assessment.

What is the cost of the developmental screener?

As of 2010, the BDI-2 Screening Test could be purchased separately from the BDI-2 for $324. This includes materials for 30 children. There are several different electronic packages that can also be purchased for use of the BDI-2 Screening Test on a computer or a hand-held device.

Training and Other Requirements for Assessors

Is training available on how to administer and score the developmental screener? Who offers the training?

Yes, training is available on how to administer and score the BDI-2 Screening Test. Training is available through the publishing company, as well as through independent trainers across the country. Detailed information is available on the publishing company's website: http://www.riversidepublishing.com/products/bdi2/training.html.

Is it necessary to have a professional background or technical training over and above the training on the developmental screener to administer or complete the developmental screener?

Yes, the developers recommend that people administering the Screening Test have college-level training. The primary user groups include preschool, kindergarten, and primary school teachers, special educators and early intervention providers. Additionally, the BDI-2 developmental screener is appropriate for use by speech-language pathologists, psychologists, and diagnosticians. Users should have significant understanding of the purpose of the measure and familiarity with child development.

Is it necessary to have a professional background or technical training over and above the training on the developmental screener to score the developmental screener?

No, however, people who interpret and report the results of the screener should have a higher level of training and supervised experience. The manual suggests that they have college-level training.

Are regular checks on administration required or recommended to ensure appropriate administration? If so, when and by whom?

Yes, the developers recommend that a professional train and be available to those who are administering the BDI-2 Screening Test for consultation and to make sure the data are being collected accurately.

Information Reporting System for the Developmental Screener

Electronic Data Entry. *Does the developmental screener come with a process for entering information from the developmental screener electronically?*

Yes, the BDI-2 developmental screener can be entered electronically. Using the electronic score pad replaces the need for a paper record pad. The Spanish-language version can also be entered electronically.

Electronic Reports. *Can programs generate electronic reports of individual children's data?*

Yes, electronic reports can be generated. Reports are available for the child level, for program monitoring by program directors or administrators. There are also special reports that are available for use in Head Start monitoring and for Individualized Education Plan development.

Approaches to Family/Parent Input

Tools for Family Input. *Does the developmental screener include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

No, the BDI-2 Screening Test does not include tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development.

Sharing Results. *Does the developmental screener include recommendations on how to share developmental screener results with the child's family?*

Yes, the manual includes some recommendations on how to share the screening results with a child's family. The manual emphasizes that families have a right to be informed of the results in simple and clear language so they understand them. Thus, the people who administer the developmental screener have an ethical responsibility to communicate the results, explain the meaning of the scores, and provide possible implications or recommendations based on the results. There are several charts that can help parents visualize the results and how their child is developing in comparison to other children of the same age. Additionally, there is guidance for talking with families of children with disabilities.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this a developmental screener with developmental norms?*

Yes, the BDI-2 does have developmental norms that were created using a sample of 2,500 children ages 0-7 years old. However, the BDI-2 Screening Test does not have separate norms from the full BDI-2. There are no developmental norms for the Spanish-language version of the BDI-2 Screening Test.

Which populations were included in the norming sample?

There were 2,500 children in the norming sample; however, the table on the next page shows just the characteristics of children who were of preschool age (750 children) in 2003. The BDI-2 Screening Test has not been examined separately.

Availability of Versions in Languages Other than English. *Is the developmental screener available in languages other than English?*

Yes, the BDI-2 Screening Test is available in Spanish. The developers state that the Spanish version can be administered after a child is given the English version if it becomes apparent that the child does not know enough English to complete the English version. Or, the items that the child scored incorrectly on the English version can be administered from the Spanish version. Additionally, they state that the scoring process for the Spanish version is not different from the English version. This means that the scores for Spanish-speaking children are compared to the norms and developmental abilities of the English-speaking children with whom the BDI-2 Screening Test was developed.

How were versions in languages other than English developed?

The Spanish version of the BDI-2 Screening Test is not a complete translation of the BDI-2 Screening Test in English. The visual materials needed for administering the measure were translated, as well as the record forms and score reports.

Ninety-six percent of the full BDI-2 assessment English items were translated into Spanish. Twenty of the full BDI-2 assessment items needed significant modification when translated from the communication, motor, and cognitive domains. Three of the items were a part of the BDI-2 Screening Test. The modifications occurred when there was not a comparable word in Spanish for the original English word, or when the item had to do with rhyming words that did not rhyme when translated into Spanish. Translation occurred through a consensus process to determine what would be appropriate for many different groups of Spanish speakers in the United States. After items were translated the first time, they were reviewed for grammar issues and cultural biases. Items were revised and reviewed twice more before a final set was published.

What are the findings on the reliability and validity of versions of the developmental screener other than English?

The developers have not examined the reliability and validity of Spanish versions of the BDI-2 Screening Test.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

Yes, there are accommodations for screening children with identified or suspected special needs. The person administering the BDI-2 Screening Test should be familiar with behaviors that may interfere with a child's ability to respond, limitations based on the disability of the child, and relevant information about the child, such as medication and assistive technology. There is particular guidance for children with motor, vision, hearing, or speech impairments or deafness, emotional or behavioral disturbance, and multiple disabilities. For example, a child with a motor impairment might take longer to make small movements, so allowing more time for the child to complete the task might be necessary.

Options for Use with Special and Diverse Populations

Appropriateness. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

During the development of the BDI-2 Screening Test, individuals from five racial/ethnic/linguistic groups (African American, American Indian and Alaskan Native, Asian, Hispanic, and White) and representing both sexes reviewed items from the original BDI. These groups compiled the information and used it to select, revise, or delete items for the final version of the full BDI-2 assessment and the BDI-2 Screening Test.

Risk Levels. *What terminology is used to describe risk levels (e.g., delay, no delay, at risk, caution, rescreen, okay, etc.)?*

If the BDI-2 Screening Test is administered first, the scores indicate "pass" or "refer." If the scores indicate that the child should be referred, then the full BDI-2 can be administered.

Characteristics of 2003 Norming Sample
Number of children in the sample: 2,500

Note, only children between 36-72 months are included in this table (750 children). The total sample includes 2,500 children from birth through 95 months.

	Percentage of Children					
Gender	**36-41 Months**	**42-47 Months**	**48-53 Months**	**54-59 Months**	**60-65 Months**	**66-72 Months**
Male	48.0	51.2	47.2	50.4	48.8	52.8
Female	52.0	48.8	52.8	49.6	51.2	47.2
Race/Ethnicity						
White	61.6	12.8	17.6	3.2	4.8	61.6
African-American	58.4	13.6	19.2	4.0	4.8	12.8
Hispanic/Latino	60.0	14.4	18.4	2.4	4.8	16.8
Asian	56.8	14.4	20.0	3.2	5.6	3.2
Other	59.2	14.4	20.0	2.4	4.0	5.6
Parental Education						
Less than High School	17.6	18.4	19.2	19.2	17.6	19.2
High School or GED	32.0	33.6	21.2	35.2	45.6	30.4
Beyond High School	50.4	48.0	49.6	45.6	50.4	50.4

Reliability and Validity Information

What is known about the reliability and validity of the developmental screener…

In English?

There is information about reliability, validity, sensitivity, and specificity of the BDI-2 Screening Test in English. This information is provided in response to later questions of this profile.

In other languages?

The developers have not examined the reliability, validity, sensitivity, and specificity of the BDI-2 Screening Test in other languages.

For dual language learners?

While the developers discuss use of the BDI-2 Screening Test with dual language learners, they have not examined the reliability, validity, sensitivity, and specificity for this population.

For children with special needs?

The developers have examined the sensitivity and specificity of the BDI-2 Screening Test for children with special needs; however, they have not examined other aspects of reliability and validity.

For American Indian/Alaskan Native children?

While American Indian/Alaskan Native children were included in the sample with which the BDI-2 Screening Test was tested (these children were included in the Other category and thus a specific percentage cannot be extracted), the developers have not examined the reliability, validity, sensitivity, and specificity of the BDI-2 Screening Test for this population.

For children of migrant and seasonal farm workers?

The developers do not provide information about children of migrant and seasonal farm workers, and have not examined the reliability, validity, sensitivity, and specificity for this population.

Reliability: Does the developmental screener obtain the same results, consistently, under the same conditions with the same children?

Interrater Reliability. *Do different raters agree when they are assessing the same children?*

The developers did not examine agreement between raters.

Test-Retest Reliability. *How consistent are scores if the developmental screener is administered once and then administered again soon?*

The developers did not examine the consistency of scores.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

The relationships between items that are intended to reflect the same set of skills on the BDI-2 Screening Test meet the criteria for acceptable.

Validity: Does the developmental screener measure what it is supposed to?

Content Validity. *Do experts agree that the items in the developmental screener do a good job of reflecting what the developmental screener is supposed to be assessing?*

Yes, experts agree that the items in the BDI-2 Screening Test do a good job of reflecting what it is supposed to be measuring. A national task force was created to make sure that important issues of development were included in the measure. When considering the areas to include in the measure, the task force also focused specifically on family, economic, demographic, and cultural issues that might impact a child's development. It is important to note that the task force examined the full BDI-2 assessment, not just the Screening Test.

Construct Validity. *How closely related to each other are sets of items within the developmental screener that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

The developers have not examined the relationships between sets of items that aim to address similar skills compared to those that aim to address different skills and behaviors.

Information about whether scores on sets of items relate to children's age as expected is not provided.

Convergent and Divergent Validity. *How strongly do the scores of this developmental screener show a relationship to the scores of other developmental screeners of similar domains?*

The developers have not compared the BDI-2 Screening Test to other developmental screeners.

Scores for Further Evaluation. *Are specific scores used to identify whether further evaluation is needed? How are these scores determined by the developer?*

Yes, there are specific scores, called cutoff scores, used with the BDI-2 Screening Test to decide whether further evaluation is needed.
There are cutoff scores for each of the five domains, as well as the total screening test. In order to develop these scores, the developers used data from the group of children described in the earlier table. There are cutoff scores for each age in months from birth to 7 years. A score at or below the cutoff score indicates that the child needs to be referred for further testing. A score above the cutoff indicates that the child passed that domain for his or her age.

Validity: Does the developmental screener measure what it is supposed to?

Sensitivity. *How accurately does the developmental screener correctly identify children at risk for developmental problems?*

The BDI-2 Screening Test is moderately accurate at correctly identifying children who are at risk for developmental delay. In order to test this, 512 children divided into five groups completed the BDI-2 Screening Test. In each group, some children had a previously diagnosed developmental delay, including autistic delay, developmental delay, cognitive delay, motor delay, and speech and language delay. The remaining children were developing typically. The BDI-2 Screening Test accurately identified the children who are at risk in the autistic and cognitive delay group better than in the other delay groups. The developmental and speech and language delay groups had many fewer children who were correctly identified as being at risk for developmental delay.

Specificity. *How accurately does the developmental screener correctly identify children who are not at risk for developmental problems?*

The BDI-2 Screening Test is moderately accurate at correctly identifying children who are not at risk for developmental problems. In order to test this, 512 children divided into five groups completed the BDI Screening Test. In each group, some children had a previously diagnosed developmental delay, including autistic delay, developmental delay, cognitive delay, motor delay, and speech and language delay; the remainder of the children were developing typically. For all groups, 79% or more children who were not at risk for developmental problems were identified as not having developmental problems. The BDI-2 Screening Test accurately identified the children who were not in a risk group in the motor and developmental delay group better than in the other delay groups.

BRIGANCE® Screens

Developers: Albert Brigance and Frances Page Glascoe
Publisher: Curriculum Associates, Inc.

http://www.curriculumassociates.com/products/detail.asp?title=BrigEC-Screens

Developmental domains addressed in the developmental screener, as stated by the publisher:
See below

Intended age range:
Birth through first grade

Number of items:
Each BRIGANCE® Screen has at least 32 domain- and skill-specific sections. Each section contains between 2 and 24 items.

In what settings can this developmental screener be used (e.g., centers, homes, medical facilities, other)?
The BRIGANCE® Screens can be used in early childhood program settings, pediatric clinics, and at screening fairs, which are often offered in communities in collaboration with health care providers.

Background

Purpose:

The BRIGANCE® Screens are developmental screeners used to quickly and accurately identify those children who may have developmental problems such as language impairments, learning disabilities, or cognitive delays, or who may be academically talented or gifted. The BRIGANCE® Screens include the Early Childhood Screen II (0-35 months), the Early Childhood Screen II (3-5 years), the K & 1 Screen II (kindergarten and first grade), the Early Head Start Screen (0-35 months), and the Head Start Screen (3-5 years). The Head Start editions contain the same assessments as the early childhood editions, but the introduction is specific to Head Start and relates the content of the assessments to the Head Start domains. The technical information profiled here pertains to all of the screens that are appropriate for use with 3- to 5-year-olds.

What is the appropriate time period between administering, recording, or reviewing the data?

According to the *BRIGANCE® Screens* manual, screening can be scheduled at any time during the year and should be offered in response to concerns by parents and teachers. In addition, children at psychosocial risk should be rescreened within six months of initial screening to review progress and make any needed referral decisions.

How long does it take to administer the developmental screener?

The *BRIGANCE® Screens* can be administered and scored in about 15 minutes.

Language(s) developed for:

The *BRIGANCE® Screens* were developed in English. Some materials are available in Spanish.

Developmental domains addressed in the developmental screener, as stated by the publisher:

- Communication
  - Expressive vocabulary
  - Syntax
  - Articulation
  - Fluency
  - Receptive language
- Motor
  - Gross motor
  - Fine motor
  - Graphomotor (handwriting skills)
- Academics/preacademic
  - Knowledge of colors
  - Knowledge of letters
  - Knowledge of letter sounds
  - Knowledge of numbers

Availability and Cost of the Developmental Screener

Is the developmental screener available to programs without restrictions?

Yes, the BRIGANCE® Screens are available to programs without restrictions.

What is the cost of the developmental screener*?*

As of 2010, the complete BRIGANCE® Screening Kit for Early Childhood Screen II (3-5 years) or the Head Start Screen cost $219.00. This kit includes: The Early Childhood Screen II (3-5 years) or the Head Start Screen, 60 assorted data sheets, screen accessories, tote bag, free 24/7 online training, and free online scoring.

Training and Other Requirements for Assessors

Is training available on how to administer and score the developmental screener? Who offers the training?

Yes, training is available on how to administer and score the BRIGANCE® Screens. The screener's publisher, Curriculum Associates, offers free online inservice training on the BRIGANCE® Screens. Please see www.CAtraining.com for more information.

Is it necessary to have professional a background or technical training (over and above training on the developmental screener) to administer or complete the developmental screener?

The BRIGANCE® Screens can be used by teachers, paraprofessionals, special educators, psychologists, occupational and physical therapists, day care and early childhood teachers, and speech-language pathologists. The BRIGANCE® Screens manual suggests that all BRIGANCE® Screens administrators become familiar with the directions and that they administer the screens in accordance with the instructions. The manual also suggests that those administering the screens, especially to the youngest age groups, have experience and a background in child development.

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to score the developmental screener?

Administrators of the BRIGANCE® Screens do not need a technical training to score the measure. However, the manual suggests that all administrators become familiar with the directions and scoring procedures, and that they score the screens in accordance with the instructions.

Are regular checks on faithful administration required or recommended? If so, when and by whom?

No information is provided regarding the performance of regular checks on faithful administration of the screens.

Information Reporting System for the Developmental Screener

Electronic Data Entry. *Does the developmental screener come with a process for entering information from the developmental screener electronically?*

Yes, the BRIGANCE® Screens can be scored by hand or with the BRIGANCE® Online Management System. The software for the BRIGANCE® Online Management System must be purchased. A year's license to use the *Online Management System* costs $8.00 per child.

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

Yes, the BRIGANCE® Online Management System can generate reports of screening data for individual children, the classroom, and the whole program or school system.

Approaches to Family/Parent Input

Tools for Family Input. *Does the developmental screener include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

Yes, the BRIGANCE® Screens contain a Parent's Rating Form that asks parents a series of questions about their child that address motor skills and health status, fine-motor and visual skills, self-help skills, speech and language, general knowledge and comprehension, and social and emotional skills. Parents respond to questions by checking the appropriate box (no, uncertain, yes). Parents are also asked for additional information that would help school staff in working with the child.

Sharing Results. *Does the developmental screener include recommendations on how to share developmental screener results with a child's family?*

Yes, the BRIGANCE® Screens include several recommendations on how to share the screening results with a child's family.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this a developmental screener with developmental norms?*

Yes, the BRIGANCE® Screens are screeners with developmental norms. The items for the BRIGANCE® Screens were selected from the BRIGANCE® Diagnostic Inventory of Early Development (IED), created in 1979. The original IED was normed with a sample of 1,156 children ranging in age from 1 year, 1 month to beyond 6 years. The group was 50 percent male, 73 percent White, 15 percent African American, and 12 percent Hispanic. In 2005, the BRIGANCE® Screens were renormed using both new and existing data. Existing data included: 1) children assessed as part of the norming for the BRIGANCE® Inventory of Early Development II, 2) the BRIGANCE® Comprehensive Inventory of Basic Skills-Revised, 3) the BRIGANCE Infant and Toddler Basic Assessments, and 4) the 1995 norming and 2005 renorming of the BRIGANCE® Screens.

Which populations are included in the norming sample?

The BRIGANCE® Screens were tested on a nationally representative sample of children from 29 U.S. states and included African Americans, Hispanics, Asian and Pacific Islanders, and Native Americans in proportion to their prevalence in the US population according to the U.S. Census Bureau. Socioeconomic variables such as parents' level of education and income were also considered, again in proportion to prevalence in the U.S. population. Children in the sample whose primary language was Spanish were tested in Spanish using standardized Spanish directions. The demographic information is reported in the manual by geographic location for the 2005 study. See the table on the next page for more information about these children.

Availability of Versions in Languages Other than English. *Is the developmental screener available in languages other than English? Which languages?*

Spanish-language directions booklets are available for administering the screens in Spanish, but there is no separate Spanish version of the screens.

How were versions in languages other than English developed?

The BRIGANCE® Screens are not available in languages other than English.

What are the findings on the reliability and validity of versions of the developmental screener in languages other than English?

The BRIGANCE® Screens are not available in languages other than English.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

Yes, the BRIGANCE® Screens manual includes several accommodations and adaptations for children with motor impairment, hearing impairment or deafness, vision impairment or blindness, severe speech impairments, emotional disturbance and behavioral problems, significant health problems, autism and developmental disorders, and traumatic brain injury.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

The developers do not provide information about whether the appropriateness of the *BRIGANCE® Screens* for diverse populations was addressed in this way.

Risk Levels. *What terminology is used to describe risk levels (e.g., delay, no delay, at risk, caution, rescreen, okay, etc.)?*

The BRIGANCE® Screens have a range of results, including: below cutoff with risk factors, below cutoff without risk factors, above cutoffs (i.e., within normal limits), and above cutoffs for gifted/talented.

Characteristics of 2005 Norming Sample
Number of children in the sample: 1,366

	Percentage of Children			
Age (in years and months)	South	West	North	Central
3-0 to 3-11	1.9	1.7	1.4	2.0
4-0 to 4-11	2.0	0.9	1.4	2.0
5-0 to 5-11	4.4	1.7	3.3	3.8
Race/Ethnicity				
White	17.5	7.3	17.8	22.5
African American	4.8	5.0	3.0	1.0
Hispanic	5.6	0.0	5.5	3.2
Asian/Other	1.3	1.1	2.4	2.0
Gender				
Male	14.1	6.4	13.3	15.6
Female	15.0	6.9	15.4	13.1
Parental Education				
Less than High School	7.5	4.5	5.8	2.8
High School	9.2	3.9	6.7	6.9
High School +	6.7	3.1	7.7	9.9
College +	5.8	2.0	8.3	9.0

Reliability and Validity Information

What is known about the reliability and validity of the developmental screener…

In English?

There is reliability, validity, sensitivity, and specificity information for the BRIGANCE® Screens in English. This information is outlined in later questions in this profile.

In other languages?

The BRIGANCE® Screens are not available in other languages.

For dual language learners?

While dual language learners were included in the sample with which the screener was tested[1], the developers have not examined the reliability, validity, sensitivity, and specificity of the BRIGANCE® Screens for this population.

For children with special needs?

The developers do not provide any information about children with special needs, and have not examined the reliability, validity, sensitivity, and specificity of the BRIGANCE® Screens for this population.

For American Indian/Alaskan Native children?

While American Indian/Alaskan Native children were included in the sample with which the screener was tested, the developers did not examine the reliability, validity, sensitivity, and specificity for this population.

For children of migrant and seasonal farm workers?

The developers do not provide any information about children of migrant and seasonal farm workers, and have not examined the reliability, validity, sensitivity, and specificity of the BRIGANCE® Screens for this population.

[1]The developers state that children whose primary language was Spanish were tested in that language, but do not indicate what percentage of the norming sample were dual language learners.
[2] Native Americans and Pacific Islanders are included under "Asian/other" in the demographic table. The developers do not report what percentage of the norming sample is Native American or Pacific Islander.

Reliability: Does the developmental screener obtain the same results, consistently, under the same conditions with the same children?

Interrater Reliability. *Do different raters agree when they are assessing the same children?*

There is acceptable agreement between raters when they are screening the same children using the BRIGANCE® Screens. Agreement between raters was examined across numerous sites. The process involved examiners who were paraprofessionals, teachers, and health care providers working with a range of children, including those with and without risk factors and special needs.

Test-Retest Reliability. *How consistent are scores if the developmental screener is administered once and then administered again soon?*

The consistency of scores when the BRIGANCE® Screens are administered once and then administered again soon is acceptable. To examine this, Enright (1991) administered the Inventory of Early Development (IED) twice to 1,156 students (14 percent were African American and 11 percent were Hispanic). Additional evidence for the consistency of scores comes from the norming of the Inventory of Early Development II (IED II) and the Comprehensive Inventory of Basic Skills-Revised (CIBS-R), which contain all items of the BRIGANCE® Screens. The results of the norming studies showed that the scores on the IED, IED II, and the CIBS-R are very consistent over short periods of time.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

Relationships between items on the BRIGANCE® Screens that are intended to reflect the same set of skills or behaviors are acceptable. This was examined with the 2005 standardization study population described above.

Validity: Does the developmental screener measure what it is supposed to?

Content Validity. *Do experts agree that the items in the developmental screener do a good job of reflecting what the developmental screener is supposed to be assessing?*

Yes, the development of the BRIGANCE® Screens was based on collaboration with other educators who helped with item selection.

Construct Validity. *How closely related to each other are sets of items within the developmental screener that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

There are moderate to high relationships between related subtests on the BRIGANCE® Screens that aim to address similar skills and behaviors, for example, between expressive and receptive language and between gross and fine motor skills.

Yes, the developers examine whether scores on sets of items relate to children's age as expected.

Convergent Validity. *How strongly do the scores of this developmental screener show a relationship to the scores of other developmental screeners of similar domains?*

The BRIGANCE® Screens are strongly related to other well-established measures aimed at measuring the same skills and behaviors. This was examined with the sample described in the earlier table by comparing the total scores on the BRIGANCE® Screens to scores from the IED II or the CIBS-R during the 2005 norming studies, as well as to a range of other diagnostic and screening tools. There are strong relationships between similar domains across these measures. It should be noted that the items on the BRIGANCE® Screens are taken from the IED II, so there is overlap between the two tools.

Scores for Further Evaluation. *Are specific scores used to identify whether further evaluation is needed? How are these scores determined by the developer?*

Yes, specific scores called cutoffs are used to identify children who may need further evaluation. Statistical analyses were done to determine which cutoff scores for each of the screens best identify children with disabilities or those who are at risk for academic difficulties.

Sensitivity. *How accurately does the developmental screener correctly identify children at risk for developmental problems?*

The BRIGANCE® Screens for 3-year-olds and 4-year olds are moderately accurate at correctly identifying children at risk for developmental delays. The screens for 5-year-olds are highly accurate at correctly identifying children at risk for developmental delays.

Specificity. *How accurately does the developmental screener correctly identify children who are not at risk for developmental problems?*

The BRIGANCE® Screens are moderately accurate at correctly identifying children who are not at risk for developmental problems for children 3 through 5 years old.

Follow-Up Guidance

Program Follow-Up Steps. *Does the developmental screener come with guidance about follow-up steps based on the results?*

Yes, the manual includes thorough guidance about follow-up steps based on the results of the screening.

Family Follow-Up Steps. *Does the developmental screener include recommendations on how families might follow up on the results of the screening?*

Yes, the manual includes many recommendations on how families might follow-up on the results of the screening.

References

Enright, B. (1991). *BRIGANCE® Diagnostic Inventory of Early Development-Revised*: A technical report. North Billerica, MA: Curriculum Associates, Inc.

Glascoe, F.P. (2010). *Technical Report for the BRIGANCE® Screens*. North Billerica, MA: Curriculum Associates, Inc.

Denver II

Developers: William Frankenberg and Josiah Dodds
Publisher: Denver Developmental Materials, Inc.

http://www.denverii.com/

Developmental domains addressed in the developmental screener, as stated by the publisher:
There are four developmental areas in the Denver II:

- Personal-social
- Fine motor-adaptive
- Language
- Gross motor

Intended age range:
0 months to 6 years

Number of items:
The Denver II includes 125 items; there are 5 additional behavior items that are administered at the end. However, all 125 items are not administered to each child. The number of items administered depends on how much time is available and whether the goal of the screening is to determine only if the child is at risk or also the child's relative strengths.

In what settings can this developmental screener be used (e.g., centers, homes, medical facilities, other)?
The Denver II can be used in many settings, including schools, early childhood programs, doctors' offices, public health clinics, and home visiting programs.

Background

Purpose:
The Denver II is a developmental screener that examines children's ability on age-appropriate activities to see if there might be a delay. The Denver II is meant to compare the child's ability to other children of the same age.

What is the appropriate time period between administering, recording, or reviewing the data?
If the person administering the measure thinks that there might be any concerns with a child, the child should be rescreened using the Denver II one to two weeks after the initial screening. This can rule out whether the child was showing his or her true ability, or whether the screening results were influenced by other factors such as fatigue, fear, or illness.

How long does it take to administer the developmental screener?
The amount of time it takes to administer the Denver II depends on the age and developmental level of the child. Infants may take 10 minutes; 5 year-olds may take 30 minutes.

Language(s) developed for:
The Denver II was developed in English, but the materials have been translated into Spanish.

Availability and Cost of the Developmental Screener

Is the developmental screener available to programs without restrictions?

Yes, the Denver II is available without restrictions.

What is the cost of the developmental screener?

As of 2010, the materials and the test kit cost $110. These materials can screen 100 children. The Spanish-language kit can be purchased for $150. Another resource for parents and families to fill out, the Prescreening Developmental Questionnaire (PDQ II), is available in English and Spanish and costs about $32 for 100 sheets.

Training and Other Requirements for Assessors

Is training available on how to administer and score the developmental screener? Who offers the training?

Yes, the publishers offer in person training on how to administer and score the Denver II. Training is available in Colorado several times a year. Onsite training is also available upon request. Those who administer the Denver II can become certified trainers to train teachers or professionals within their programs. Training may also be available through videotapes; contact the publisher for more information (http://www.denverii.com/training.html).

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to administer or complete the developmental screener?

No, the Denver II can be administered by many different types of people and they do not need to professional background of technical training over and above training on the screener. However, they must be trained to administer the screener in the standardized manner. They must also pass a proficiency test before administering the Denver II. The proficiency test is included in the Denver II technical manual and can be photocopied.

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to score the developmental screener?

No, if the person administering the Denver II has successfully completed the training and passed the proficiency test, he or she can also score the developmental screening tool.

Are regular checks on administration required or recommended to ensure appropriate administration? If so, when and by whom?

Master instructors are required to recertify every three years. It is recommended that their screener-trainees be recertified yearly or at most, within three years.

Information Reporting System for the Developmental Screener

Electronic Data Entry. *Does the developmental screener come with a process for entering information from the developmental screener electronically?*

Yes, the Denver II comes with a process for entering information electronically. It is linked to an internet site where those administering the Denver II can create a secure account with a log-in ID. All of the scores from the Denver II can be saved in the online account (see http://www.denveriionline.com/).

Electronic Reports. *Can programs generate electronic reports of individual children's data?*

Yes, programs can generate electronic reports; however the manual does not provide information about what type of reports can be generated.

Approaches to Family/Parent Input

Tools for Family Input. *Does the developmental screener include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

Yes, the Denver II includes parent and family input on a child's skills and development. The Denver II is often administered with an adult (e.g., parent, teacher, someone who knows the child well) in the room. This allows for the administrator to ask the parents or another adult who knows the child questions about the child's behavior that cannot be examined while the child is being screened—for example, can the child dress without help. Many of the questions asked about younger children need more parental or familial input. If the Denver II is administered without a parent or family member present, the person administrating can ask for input at a later time.

Additionally, the Prescreening Developmental Questionnaire (PDQ II), available in English and Spanish, can be filled out by parents or another family member. This is a 105-item questionnaire, but parents or another family member complete only a handful of questions based on the child's age. It takes about 10 minutes to complete. When this is filled out, the Denver II administrator has some initial information about the child and can use the information from the PDQ II when talking with the family of the child being screened.

Sharing Results. *Does the developmental screener include recommendations on how to share developmental screener results with a child's family?*

No information is provided about sharing the results with a child's family.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this a developmental screener with developmental norms?*

Yes, the Denver II has developmental norms. The norms for the Denver II are based on a sample of 2,096 children from Colorado from 1990.

Which populations are included in the norming sample?

This group of 2,096 children were from either Denver County (1,039 total) or another county in Colorado (1,057 total). They were from three types of areas: urban (50,000+ residents), semi-rural (2,500-50,000 residents), or rural (not fitting into either of the other categories). The information in the table that follows is for children in the 3 to 5 year old age range. However, the manual includes information for children from birth to 78 months. Children in this group were recruited from health care settings, child care centers, preschools, Head Start programs, churches, and social services agencies. See the table on the next page for more information about these children.

Availability of Versions Other than English. *Is the developmental screener available in languages other than English? Which languages?*

The Denver II has been translated into Spanish.

How were versions in languages other than English developed?

The Spanish-language version is translated from the English version. No additional information is provided about the development of the Spanish version of the Denver II.

What are the findings on the reliability and validity of versions of the developmental screener in languages other than English?

The developers have not provided findings on reliability and validity of the Spanish version.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

The developers do not provide information about specific accommodations for screening children with special needs.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

The developers have not provided information about whether the appropriateness of the Denver II for diverse populations was addressed through cognitive testing or focus groups.

Risk Levels. *What terminology is used to describe risk levels (e.g., delay, no delay, at risk, caution, rescreen, okay, etc.)?*

The Denver II uses several terms for the scores on each item. Children can *"pass"* an item if they do it correctly. Children can *"fail"* an item if they do it incorrectly. Children can score a *"normal"* on an item if they fail or refuse to do it correctly and it is an item that is indicated to be much above their age level (25-75 percent). Children can score a *"caution"* on an item if they fail or refuse to do it correctly and 75-90 percent of children their age can do the item. These percentages are based on the developmental norms that are mentioned earlier in this profile. Finally, children can score a *"delay"* on an item if they fail or refuse to do an item that is at or below their age level.

There are also several terms used to describe the overall score on the Denver II and the child's risk level. The test result is considered *"normal"* if there are no delays on any items and only one caution. If the test has two or more cautions and/or one or more delay, then the test result is considered *"suspect."* Lastly, if a child refuses to complete one or more items that are at or below age level or more than one item that 75-90 percent of children of the same age can do, then the test result is considered "*untestable."*

Demographic Information (2008) about the Colorado Sample of Children
Number of children in the sample: 2,096

Note, only children between 36 and 72 months are included in this table (393 children). The total sample includes 2,096 children from birth through 78 months.

	Percentage of Children					
	36-42 Months	**42-48 Months**	**48-54 Months**	**54-60 Months**	**60-66 Months**	**66-72 Months**
Race/Ethnicity						
White	60	58	54	55	58	38
Hispanic	25	26	27	32	26	38
Black	15	14	19	12	14	24
Other[16]	0	2	0	1	2	0
Gender						
Male	44	48	41	54	49	41
Female	56	52	59	46	51	50
Maternal Education (in years)						
<12	20	26	22	37	32	24
12	31	32	32	20	36	44
>12	49	42	46	39	32	32
Residence						
Urban	62	74	76	75	66	64
Rural	21	15	9	15	20	18
Semi-rural	17	11	15	10	14	18

[16] Other includes "American Indian", "Asian", etc. (as stated in manual)

Reliability and Validity Information

What is known about the reliability and validity of the developmental screener…

In English?

There is information about reliability and validity of the developmental screener in English. To test the reliability and validity of the Denver II in English, the screener was tested on a sample of children from Colorado, mentioned earlier in the profile.

In other languages?

The developers have not examined the reliability, validity, sensitivity, and specificity of the Denver II in other languages.

For dual language learners?

The developers do not provide any information about dual language learners, and have not examined the reliability, validity, sensitivity, and specificity of the Denver II for this population.

For children with special needs?

The developers have examined the sensitivity and specificity of the Denver II for children with special needs; however, they have not examined other aspects of reliability and validity.

For American Indian/Alaskan Native children?

While American Indian/Alaskan Native children were included in the sample with which the developmental screener was tested (these children were included in the Other category, so a specific percentage cannot be extracted), the developers have not examined the reliability, validity, sensitivity, and specificity of the Denver II for this population.

For children of migrant and seasonal farm workers?

The developers do not provide any information about children of migrant and seasonal farm workers, and have not examined the reliability, validity, sensitivity, and specificity of the Denver II for this population.

Reliability: Does the developmental screener obtain the same results, consistently, under the same conditions with the same children?

Interrater Reliability. *Do different raters agree when they are screening the same children?*

There is acceptable agreement between different raters when they screen the same children with the Denver II. This was examined with 38 children ages birth to 78 months with about 3 children per each 3-month age group from the Colorado sample. The maternal education of these children ranged from 12 to 20 years with an average of 15.5 years. Seventeen trained raters administered the Denver II. The developers did not provide additional information about these raters.

Test-Retest Reliability. *How consistent are scores if the developmental screener is administered once and then administered again soon?*

Consistency of scores for the Denver II was examined in two different ways. First, consistency of scores was examined with 5 to 10 minutes between administering the Denver II. For most of the 125 items, the consistency met the criteria for acceptable, but for 18 items, the consistency met the criteria for weak. Next, consistency of scores was examined with 7 to 10 days between administering the Denver II. Again, most items showed acceptable consistency, but 24 items had moderate to weak consistency. This was examined with 38 children ages birth to 78 months with about 3 children per 3-month age group from the Colorado sample. The maternal education of these children ranged from 12 to 20 years with an average of 15.5. Seventeen trained raters administered the Denver II. There is no additional information about these raters.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

The developers have not examined the relationship between items that are intended to reflect the same set of skills or behaviors.

Validity: Does the developmental screener measure what it is supposed to?

Content Validity. *Do experts agree that the items in the developmental screener do a good job of reflecting what the developmental screener is supposed to be assessing?*

The developers have not provided any information regarding whether experts agree that items on the Denver II do a good job of reflecting what the developmental screener is supposed to be measuring.

Construct Validity. *How closely related to each other are sets of items within the developmental screener that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

The developers do not provide information regarding the relationships between sets of items on the Denver II that aim to address similar skills and behaviors in comparison with sets of items that aim to address different skills and behaviors.

Information about whether scores on sets of items relate to children's age as expected is not provided.

Convergent Validity. *How strongly do the scores of this developmental screener show a relationship to the scores of other developmental screeners of similar domains?*

The developers do not provide information regarding the relationships between the results of the Denver II and the results of other developmental screeners for similar domains or for those aimed at measuring different skills or behaviors.

Scores for Further Evaluation. *Are specific scores used to identify whether further evaluation is needed? How did the developers determine these scores?*

Yes, scores are used to identify whether further evaluation is needed. Please see the question on risk level terminology for more information.

Sensitivity. *How accurately does the developmental screener correctly identify children at risk for developmental problems?*

The Denver II is moderately accurate at correctly identifying children who are at risk for developmental problems. In order to examine this, 104 children were screened with the Denver II. Eighteen of the children had identified disabilities. Thus, the Denver II is only moderately accurate in correctly identifying children at risk. This information is not presented in the manual, but appears in a separate article (Glascoe, et al. 1992).

Specificity. *How accurately does the developmental screener correctly identify children who are not at risk for developmental problems?*

The Denver II meets the criteria for low accuracy in identifying children who are not at risk for developmental problems. The Denver II was used with the same group of 104 children mentioned in the previous question; 86 of the children did not have a developmental problem. However, the Denver II identified many of these 86 children as being at risk for a developmental problem. Further, experts think the Denver II was designed to over identify children and as a result is not as accurate as other developmental screeners (Glascoe, et al. 1992).

Follow-Up Guidance

Program Follow-Up Steps. *Does the developmental screener come with guidance about follow-up steps based on the results?*

Yes, the Denver II has guidance about follow-up steps based on the results. If is the test result is *"normal,"* then the child should be screened when he or she next goes to the doctor for a well-visit or at a comparable time. If the test result is *"suspect,"* the child should be rescreened in one or two weeks to rule out factors such as fatigue, fear, or illness. If the test result is *"untestable,"* then he or she should be rescreened in one or two weeks.

If the rescreening results indicate the child is *"suspect"* or "*untestable,"* then the child should see a professional. The professional may want to take into consideration items on which the child received cautions or delays, as well as the total number of cautions or delays. Additionally, if there is other information available, such as rate of past development, other clinical considerations, and availability of referral resources, the professional should take them into account.

Family Follow-Up Steps. *Does the developmental screener include recommendations on how families might follow up on the results of the screening?*

If the program in which the child is enrolled does not offer recommendations, there are handouts that can be purchased. These Denver Developmental Activities can guide parents in activities to help in their child's development.

References

Frankenburg, W.K., Dodds, J., Archer, P., Bresnick, B., Maschka, P., Edelman, N., & Shapiro, H. (1996). Technical manual: Denver II. Denver, CO: Denver Developmental Materials.

Glasco, F.P., Bryne, K.E., Ashford, L.G., Johnson, K.L., Chang, B., & Strickland, B. (1992). Accuracy of the Denver-II in developmental screening. *Pediatrics 89(6)*, 1221-1225.

http://www.denverii.com/DenverII.html

http://www.denverii.com/denveriionline.html

Developmental Indicators for the Assessment of Learning-3rd Edition (DIAL-3)*

Developers: Carol Mardell and Dorothea Goldenberg
Publisher: Pearson Assessments

http://www.pearsonassessments.com/pai/ca/RelatedInfo/DIAL-3TechnicalInformation.htm

Developmental domains addressed in the developmental screener, as stated by the publisher:

- Motor concepts
- Language
- Self-help development
- Social development

Intended age range: 3 years, 0 months to 6 years, 11 months.

Number of items: The DIAL-3 contains 21 items.

In what settings can this developmental screener be used (e.g., centers, homes, medical facilities, other)? The DIAL-3 can be used in settings where developmental screening often takes place, including centers, school districts, homes, pediatric offices, and health fairs.

Background

Purpose:

The DIAL-3 is an individually administered developmental screener designed to identify young children in need of further diagnostic assessment for potential developmental delays. A shorter version of the DIAL-3, called Speed DIAL, is also available. The Speed DIAL consists of 10 DIAL-3 items and takes approximately half the time of the full DIAL-3 to administer.

What is the appropriate time period between administering, recording, or reviewing the data?

The developers do not provide information regarding the appropriate time between initial screening and rescreening.

How long does it take to administer the developmental screener?

The DIAL-3 takes 30 minutes to administer. The Speed DIAL takes approximately 15-20 minutes to administer.

Language(s) developed for:

The DIAL-3 and the Speed DIAL are available in both English and Spanish.

Availability and Cost of the Developmental Screener

Is the developmental screener available to programs without restrictions?

Yes, the DIAL-3 is available to programs without restrictions.

What is the cost of the developmental screener?

As of 2010, the cost of the DIAL-3 complete kit was $540.95. The complete kit contains materials in both English and Spanish, including the manual, 50 record forms (in English), 1 record form (in Spanish), 50 cutting cards, 50 parent questionnaires (in English), manipulatives, dials, operator's handbooks (in English and Spanish for motor, concepts, and language areas) plus the Speed DIAL and training packet. Additional DIAL-3 record forms (in packages of 50, available in both English and Spanish) can be purchased for $52.50. Additional parent questionnaires (in packages of 50 in English and Spanish) can be purchased for $28.75.

Training and Other Requirements for Assessors

Is training available on how to administer and score the developmental screener? Who offers the training?

Yes, some training is available on how to administer and score the DIAL-3 through Psych Corps, a division of Pearson, the publisher of the DIAL-3. There is a training video and access to a webinar through the Pearson website (http://psychcorp.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=PAa13700) and further information is available by contacting Pearson.

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to administer or complete the screener?

Teachers, professionals, or paraprofessionals can administer the DIAL-3 or Speed DIAL if they have been trained in the use of the test materials. The manual suggests that the screener be administered by a team of adults. This team is composed of a professional coordinator and three other adults called operators, each of whom administers the items in one of the three performance areas: motor, concepts, and language. The DIAL-3 coordinator is responsible for making sure that each operator is adequately trained to administer the six or seven items in the performance area they have been assigned to screen. Three children can be assessed at the same time by the team of three administrators, each working with individual children on a different performance area. The DIAL-3 kit contains all the necessary materials for conducting a DIAL-3 training workshop. A workshop to train one screening team can be conducted in four hours.

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to score the developmental screener?

The DIAL-3 or Speed DIAL can be scored by a professional or paraprofessional who has been trained in the use of the test materials.

Are regular checks on faithful administration required or recommended? If so, when and by whom?

The developers do not provide information about the performance of regular checks on faithful administration.

Information Reporting System for the Developmental Screener

Electronic Data Entry. *Does the developmental screener come with a process for entering information from the developmental screener electronically?*

Yes, the Automated System for Scoring and Interpreting Standardized Tests (ASSIST) for the DIAL-3 is a computer program that provides the user with an electronic way to score the DIAL-3. The ASSIST software costs $264.95.

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

Yes, the DIAL-3 ASSIST generates a report of the individual child's score that can be kept on file for future reference. The developers do not provide information about whether DIAL-3 ASSIST generates reports at the classroom or institution level.

Approaches to Family/Parent Input

Tools for Family Input. *Does the developmental screener include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

The DIAL-3 includes a parent questionnaire that concentrates on the child's self-help and social development. It also requests other information such as medical history, family background, and general concerns about development.

Sharing Results. *Does the developmental screener include recommendations on how to share developmental screener results with a child's family?*

Yes, the DIAL-3 manual provides information on how to hold a conference with parents about the results of the screening. This information includes guidelines for talking to parents about concerns, suggestions about how to explain the purpose of the DIAL-3, and how to go about scheduling further evaluations, if necessary.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this a developmental screener with developmental norms?*

The DIAL-3 is a screener with developmental norms. The norming of the DIAL-3 was completed between November 1995 and June 1997, with a sample of 1,560 children aged 3 years, 0 months through 6 years, 11 months.

Which populations are included in the norming sample?

The DIAL-3 norming sample was selected to match U.S. Census data taken from the March 1994 Current Population Survey. Additionally, while separate norms were not created, the DIAL-3 Spanish-language version was equated with the English-language version and tested on a sample of 605 Spanish-speaking children aged 3 years, 0 months to 6 years, 11 months. The Spanish-speaking sample included children from all regions of the United States, as well as from Puerto Rico and Panama. This sample consisted of all children at the testing sites who were monolingual Spanish speakers. See the tables on the next pages for more information about these children.

Availability of Versions in Languages Other than English. *Is the developmental screener available in languages other than English? Which languages?*

The DIAL-3 is also available in Spanish.

How were versions in languages other than English developed?

To develop the Spanish-language version of the DIAL-3, the administration instructions for the motor, concepts, and language areas and the parent questionnaire were translated from English to Spanish. The adapted Spanish version was reviewed by people with knowledge of various Spanish dialects, many of whom were early childhood professionals. It was then decided that the English and Spanish versions of the DIAL-3 should be equated so that children are compared to the same set of norms, regardless of whether they are tested in English or Spanish.

What are the findings on the reliability and validity of versions of the developmental screener in languages other than English?

The developers do not provide information about the reliability and validity of the Spanish-language version.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

The developers do not provide information about suggested accommodations for screening children with identified or suspected special needs.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

Following a review of the items on the DIAL-R, new items and tasks were written in the motor, concepts, and language areas. The authors administered all new items to a sample of 54 children. This sample was evenly divided by 6-month age segments, race, gender, and nationality (U.S./Mexican). Items that were difficult to administer, were biased, or did not discriminate between age groups were eliminated from this trial version of the DIAL-3. No additional information regarding the use of cognitive testing or focus groups with diverse populations is provided.

Risk Levels. *What terminology is used to describe risk levels (e.g., delay, no delay, at risk, caution, re-screen, okay, etc.)?*

The risk levels are described in the DIAL-3 as "potential delay" and "OK."

Characteristics of 1997 Norming Sample
Number of children in the sample: 1,560

	Percentage of Children							
	3-0 to 3-5	**3-6 to 3-11**	**4-0 to 4-5**	**4-6 to 4-11**	**5-0 to 5-11**	**5-5 to 5-11**	**6-0 to 6-5**	**6-6 to 6-11**
Race/Ethnicity								
White	5.2	8.6	9.5	11.1	12.7	8.6	9.0	5.4
African American	1.2	2.6	3.3	3.6	1.3	.70	.80	1.2
Hispanic	.40	.80	1.3	3.0	2.1	.90	1.1	.60
Other	.20	.64	1.0	1.2	1.5	.40	.12	.51
Gender								
Male	4.0	6.1	7.7	10.8	9.3	5.6	5.4	4.1
Female	3.1	6.5	7.3	8.3	8.3	4.9	4.9	3.6
Geographic Region								
Northeast	.40	1.2	1.6	3.0	1.8	1.7	1.2	1.0
North Central	2.9	5.2	4.7	3.7	5.6	2.3	3.1	2.1
South	3.3	5.1	6.6	8.0	7.0	5.1	5.4	4.0
West	.40	1.1	2.1	3.9	3.1	1.5	.64	.70
Parent Education (in years)								
11	.51	1.3	2.7	3.6	3.1	2.1	1.6	.90
12	1.9	4.2	5.3	6.8	6.4	3.9	3.2	2.7
13-15	3.1	4.6	.50	6.8	5.8	3.4	3.7	2.9
16	1.5	2.5	2.0	1.8	2.3	1.1	1.8	1.2

Characteristics of 1997 Spanish-speaking Norming Sample
Number of children in the sample: 605

	Percentage of Children							
	3-0 to 3-5	**3-6 to 3-11**	**4-0 to 4-5**	**4-6 to 4-11**	**5-0 to 5-11**	**5-5 to 5-11**	**6-0 to 6-5**	**6-6 to 6-11**
Gender								
Male	1.5	4.3	6.3	16.2	12.2	6.6	3.5	3.0
Female	2.6	4.1	7.1	12.9	7.6	6.0	4.0	2.1
Parental Education Level	**3-0 to 3-11**		**4-0 to 4-11**		**5-0 to 5-11**		**6-0 to 6-11**	
Grade 11 or Less	8.1		31.4		23.8		6.8	
High School Graduate	1.5		5.3		2.8		2.0	
1-3 Years of College	1.8		4.0		3.6		2.5	
4 or More Years of College	1.1		1.8		2.1		1.3	
Geographic Region								
Northeast	0.0		0.0		1.1		.50	
North Central	.33		2.3		1.8		.16	
South	4.0		6.4		5.1		2.3	
West	8.2		33.7		24.3		9.6	

Reliability and Validity Information

What is known about the reliability and validity of the developmental screener…

In English?

There is information about reliability, validity, sensitivity, and specificity of the DIAL-3 in English. This information is outlined in responses to later questions in this profile.

In other languages?

The developers do not provide information about reliability, validity, sensitivity, or specificity of the Spanish version of the DIAL-3.

For dual language learners?

The developers do not provide any information about dual language learners, and have not examined the reliability, validity, sensitivity, and specificity of the screener for this population.

For children with special needs?

The developers have examined the sensitivity and specificity for children with special needs, but not other aspects of reliability and validity with this population.

For American Indian/Alaskan Native children?

The developers do not provide any information about American Indian/Alaskan Native children, and have not examined the reliability, validity, sensitivity, and specificity of the DIAL-3 for this population.

For children of migrant and seasonal farm workers?

The developers do not provide any information about children of migrant and seasonal farm workers, and have not examined the reliability, validity, sensitivity, and specificity of the DIAL-3 for this population.

Reliability: Does the instrument obtain the same results, consistently, under the same conditions with the same children?

Interrater Reliability. *Do different raters agree when they are assessing the same children?*

The developers do not provide information about whether different raters agree when they are screening the same children.

Test-Retest Reliability. *How consistent are scores if the developmental screener is administered once and then administered again soon?*

The consistency of DIAL-3 scores is acceptable when the developmental screener is administered once and then administered again to the same children. To examine this, the DIAL-3 was administered twice to 158 children drawn from the norming sample. The time between administrations ranged from 12 to 65 days, with an average of 28 days. The 158 children were drawn from the norming sample and divided into two groups based on age. One group consisted of 80 children aged 3 years, 6 months to 4 years, 5 months, and an older group consisted of 78 children aged 4 years, 6 months to 5 years, 10 months. The test-retest sample was 45.6 percent female and 54.4 percent male. There were 10 African American children, 3 Hispanic children, and 145 White children. Twelve children had parents with 11 years of education or less, 46 children had parents who were high school graduates, 65 children had parents with 1 to 3 years of college or technical school, and 35 children had parents with 4 or more years of college. Eighty-two children were from the North Central region of the United States, 74 were from the South, and 2 were from the West. The developers do not provide information about the teachers in this sample.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

The developers do not provide information about the relationships among items that are intended to reflect the same set of skills or behaviors.

Validity: Does the instrument measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the developmental screener do a good job of reflecting what the developmental screener is supposed to be assessing?*

The developers do not provide information regarding consultation with experts on this issue.

Construct Validity. *How closely related to each other are sets of items within the developmental screener that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

The developers do not provide information about relationships between sets of items on the DIAL-3 that aim to address similar skills and behaviors.

Information about whether scores on sets of items relate to children's age as expected is not provided.

Convergent and Divergent Validity. *How strongly do the scores of this developmental screener show a relationship to the scores of other developmental screeners of similar domains?*

Seventy-six children who were part of the norming sample were also administered the Early Screening Profiles (ESP). They were given the DIAL-3 and the ESP between 0 and 154 days apart. Results showed weak to moderate relationships between scores on similar domains of the DIAL-3 and the ESP. The majority of the relationships between similar domains, for example the ESP verbal concepts subtest and the DIAL-3 language domain, fall in the moderate range. The DIAL-3 had the strongest relationships with the ESP language subscale and with the ESP verbal concepts and cognitive/language profile.

The relationships between domains aimed at measuring different skills or behaviors are weak, as expected. For example, there are weak relationships between the ESP logical relations subtest and the DIAL-3 motor domain.

The age range of the sample of 76 children used to examine relationships between the DIAL-3 and other developmental screeners was 3 years, 8 months to 5 years, 8 months. There were 34 females and 42 males. Eighteen children were African American, 6 were Hispanic, and 52 were White. Eighteen children had parents with 11 years of education or less, 26 had parents who were high school graduates, 27 had parents with 1 to 3 years of college or technical school, and 5 children had parents with 4 or more years of college. Three children were from the North Central region of the United States and 73 children were from the South.

Scores for Further Evaluation. *Are specific scores used to identify whether further evaluation is needed? How are these scores determined by the developer?*

Yes, the DIAL-3 Total, Speed DIAL, and the five performance areas (motor, concepts, language, self-help, and social development) offer a range of cutoff scores to decide whether further evaluation is needed. These cutoffs give programs the option to identify lower or higher proportions of children for referral (ranging from 2 percent to 16 percent). The cutoffs are designed to identify children who, when compared with children their own age, score at the lower end of a range of scores. The cutoff level chosen corresponds to the approximate percentage of children nationally, based on the DIAL-3 norming sample described earlier, who would be identified as having "potential delay" using that cutoff score. Both the English- and the Spanish-language versions of the DIAL-3 use the same cutoffs. The five cutoff levels will identify approximately 16, 10, 7, 5, or 2 percent of the total screening population as "potential delay."

Validity: Does the instrument measure what it is supposed to?

Sensitivity. *How accurately does the developmental screener correctly identify children at risk for developmental problems?*

The DIAL-3 manual contains two examples of how well the developmental screener correctly identifies children who are at risk for developmental problems. In first example the DIAL-3 motor subscale was administered to 26 children from the norming sample identified as receiving services for physical problems. Each child in this sample (called the clinical sample) was matched with another child who was not reported as having special needs (called the normal sample). The children in the sample were matched on age, sex, race, and parents' education level. Each child's score on the DIAL-3 motor was categorized as "below 75" or "75 and above." The DIAL-3 showed low levels of accuracy in identifying children at risk for physical problems. According to the manual, the most plausible explanation for this finding is that the children in the sample may have had a wide variety of physical problems, some of which may not have a negative effect on a child's performance on the motor area tasks.

The second example is taken from a sample of 37 children who were receiving special education services for cognitive problems. While there is no further information regarding specific diagnoses for these children, the manual states that the sample was mixed in terms of types of cognitive disabilities. Each child in the clinical sample was matched with a child from the normal sample on age, sex, race, and parents' education level. Each child's score on the DIAL-3 total was categorized as "below 75" or "75 and above." Again, the manual suggests that the accuracy of identifying children at risk for cognitive problems was low due to the fact that children with cognitive problems may show them in many ways.

Specificity. *How accurately does the developmental screener correctly identify children who are not at risk for developmental problems?*

The DIAL-3 manual contains two examples of how well the developmental screener correctly identifies children who are not at risk for developmental problems. In the first example, the DIAL-3 motor subscale was administered to 26 children from the norming sample identified as receiving services for physical problems. Each child in this sample (called the clinical sample) was matched with another child who was not reported as having special needs (called the normal sample). The children in the sample were matched on age, sex, race, and parents' education level. Each child's score on the DIAL-3 motor was categorized as "below 75" or "75 and above." The DIAL-3 motor scores showed high levels of accuracy in identifying children who were not at risk for physical problems. All except one of the children in the normal sample had scores above 75.

The second example is taken from a sample of 37 children who were receiving special education services for cognitive problems. There is no further information regarding specific diagnoses for these children, but the manual states that the sample was mixed in terms of types of cognitive disabilities. Each child in the clinical sample was matched with a child from the "normal" sample on age, sex, race, and parent education level. Each child's score on the DIAL-3 total was categorized as "below 75" or "75 and above." The DIAL-3 total scores showed a high level of accuracy in identifying children not at risk for cognitive problems. All of the children in the normal sample had scores above 75.

Follow-Up Guidance

Program Follow-Up Steps. *Does the developmental screener come along with guidance about follow-up steps based on the results?*

The manual briefly discusses follow-up steps based on whether the child scores as "potential delay" or "OK." If a child's overall screening score on the DIAL-3 falls within the potential delay score range for his or her age, the child should be referred for a diagnostic case study evaluation. It should be noted, however, that due to the range of cutoff scores (from 2 to 16 percent of the sample), the potential differences in rates of referral for further diagnostic evaluation can be rather large.

If a child's overall screening score falls within the "OK" range for his or her age, the screening administrator may still wish to give the child's parents or teachers suggested activities that will allow the child to practice specific skills. The activities will depend on the age of the child and developmental appropriateness. Some children score "OK" on a developmental screener at one age and show evidence a year later that warrants further evaluation. For this reason, development should be assessed on a yearly basis. Since children grow and develop at different rates, it is important to offer developmental evaluation on a continuing time frame rather than just once.

Family Follow-Up Steps. *Does the developmental screener include recommendations on how families might follow up on the results of the screening?*

The manual suggests that the administrator of the screener discuss the results with the child's parent or caregiver privately and in person, and request permission if further evaluation is needed. There is no additional information in the manual regarding follow-up steps the family might take based on the results of the screening.

Reference

Mardell-Czudnowski, C. & Goldenberg, D.S. (1998). Developmental Indicators for the Assessment of Learning-3rd Edition. Circle Pines, MN: American Guidance Service, Inc.

Early Screening Inventory-Revised (ESI-R)

Developers: Samuel J. Meisels, Dorothea B. Marsden, Martha Stone Wiske, and Laura W. Henderson
Publisher: Pearson Assessments

http://www.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=PAaESI&Mode=summary

Developmental domains addressed in the developmental screener, as stated by the publisher:

- Visual motor/adaptive
- Language and cognition
- Gross motor

Intended age range: The Early Screening Inventory-Preschool (ESI-P) is used with children ages 3 years, 0 months to 4 years, 5 months, and the Early Screening Inventory-Kindergarten (ESI-K) is used with children ages 4 years, 6 months to 5 years, 11 months.

Number of items: The ESI-R contains 25 items.

In what settings can this developmental screener be used (e.g., centers, homes, medical facilities, other)? According to the developer, the ESI-R has been successfully used in schools, clinics, and medical facilities.

Background

Purpose:

The Early Screening Inventory-Revised (ESI-R) is a brief developmental screener designed to identify children who may need further evaluation in order to determine if they require special educational services. The ESI-R is divided into two separate screeners: the Early Screening Inventory-Preschool (ESI-P) and the Early Screening Inventory-Kindergarten (ESI-K). This profile includes information about both the ESI-P and the ESI-K.

What is the appropriate time period between administering, recording, or reviewing the data?

A child's score on the ESI-R determines whether the child should be rescreened. If so, the ESI-R should be readministered in 8 to 10 weeks.

How long does it take to administer the developmental screener?

The ESI-R takes approximately 15-20 minutes to administer, although this may vary depending on the age of the child.

Language(s) developed for:

The ESI-R was developed in English and Spanish.

Availability and Cost of the Developmental Screener

Is the developmental screener available to programs without restrictions?

Yes, the ESI-R is available to programs without restrictions.

What is the cost of the developmental screener?

As of 2010, the cost of the ESI-R screening kit cost $137.50. The screening kit contains the ESI-R examiner's manual, screening materials, ESI-P score sheets or ESI-K score sheets (in English or Spanish) for 30 children, and parent questionnaires (in English or Spanish). Each of these items can also be purchased separately. The examiner's manual cost $59.15. The screening materials cost $22.95. The score sheets (30 per package) cost $31.95. The parent questionnaires (package of 30) cost $27.

Training and Other Requirements for Assessors

Is training available on how to administer and score the developmental screener? Who offers the training?

Yes, training videos, DVDs, and a training manual for the ESI-R are available through Pearson, the publisher of the screener. Some information about these materials is available on the Pearson website (http://www.pearsonassessments.com/pai/ca/training/training.htm), but readers should contact Pearson directly for more specific training information.

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to administer or complete the developmental screener?

The manual states that proper use of the ESI-R requires an understanding of the basic principles of standardized assessment and knowledge in early childhood behavior and development. Therefore, the person administering the developmental screener should have some formal background in early childhood assessment. The manual also says that individuals with less training and experience can administer the ESI-R under the supervision of a person with the qualifications mentioned above.

Is it necessary to have a professional background or technical training (over and above training on the developmental screener) to score the screener?

As mentioned above, the manual states that proper use of the ESI-R requires an understanding of the basic principles of standardized assessment and knowledge of early childhood behavior and development. Therefore, the person scoring the developmental screener should have some formal background in early childhood assessment. This may include teachers, paraprofessionals, social workers, and psychologists.

Are regular checks on faithful administration required or recommended? If so, when and by whom?

While regular supervision of a screener administrator during the process of learning to administer the ESI is suggested, the developers do not provide information about the performance of regular checks on faithful administration.

Information Reporting System for the Developmental Screener

Electronic Data Entry. *Does the developmental screener come with a process for entering information from the developmental screener electronically?*

Yes, ESI-R Online is the online scoring and training system for the ESI-R. ESI-Online is licensed annually and priced according to the number of children to be entered online and screened. ESI-Online costs $2.95 per child.

Electronic Reports. *Can programs generate electronic reports of their data and if so, at what level can those reports be made available (at the level of the individual child, classroom, or institution)?*

ESI-Online can generate reports for individual children and can provide summarized screening results for an entire class. ESI-Online does not generate reports at the institutional level.

Approaches to Family/Parent Input

Tools for Family Input. *Does the developmental screener include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

The ESI-R contains a parent questionnaire that consists of five sections. The first three provide basic information about the child, the child's family, and his or her school history. The fourth section contains the child's medical history, which includes an overview of the child's illnesses, hospitalizations, and health conditions. The fifth section deals with the child's overall development and addresses issues other than medical concerns that may pose problems for the child in a school setting.

Sharing Results. *Does the developmental screener include recommendations on how to share developmental screener results with the child's family?*

The manual suggests that every parent should receive feedback about screening results either verbally after the administrator has had time to determine the total score or in writing after a team review. The manual also says that screening gives only tentative conclusions and this should be communicated clearly to parents. Parents should also be given the opportunity to ask questions both before and after screening, especially when screening indicates that further evaluation is necessary.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this a developmental screener with developmental norms?*

Yes, the ESI-R is a screener with developmental norms. The ESI-P was first normed between 1993 and 1996 with a sample of 977 children divided among three groups of 6-month age spans between 3 years, 0 months and 4 years, 5 months. Data for the ESI-K norms were collected between 1986 and 1990, and then again between 1992 and 1994, with a sample of 5,034 children ages 4 years, 6 months through 5 years, 11 months. Both the ESI-P and the ESI-K were renormed in 2007.

What characteristics of the sample are the norms are based on?

The 2007 ESI-P and ESI-K norming samples included 1,200 children from 89 sites (including Head Start centers, public and private preschools, and elementary schools) in all four geographical regions in the United States. Additional data were collected from individual examiners (school psychologists, speech-language pathologists, and special-education teachers) to ensure that the norming sample matched the U.S. population on various demographic characteristics. See the tables on the next pages for more information about these children.

Children who speak both English and Spanish were screened in the language judged by their parents and the program to be their primary language. The Spanish-language versions of the ESI-P and ESI-K were used with 13 percent of the norming sample.

Availability of Versions in Languages Other than English. *Is the developmental screener available in languages other than English? Which languages?*

Both the ESI-P and the ESI-K are available in Spanish.

How were versions in languages other than English developed?

To develop the Spanish-language version of the ESI-P and ESI-K, scores on the Spanish-language version were calibrated to be comparable to scores on the English version so that equal scores on both versions represent the same level of ability.

What are the findings on the reliability and validity of versions of the developmental screener in languages other than English?

The developers have not examined the reliability and validity of the Spanish-language versions.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

The developers do not provide information regarding accommodations for screening children with special needs.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

The developers do not provide information about whether the appropriateness of the ESI-R for diverse populations was examined in this way.

Risk Levels. *What terminology is used to describe risk levels (e.g., delay, no delay, at risk, caution, rescreen, okay, etc.)?*

The risk levels on the ESI-R are labeled "OK," "rescreen," and "refer."

Characteristics of ESI-P 2008 Norming Sample
Number of children in the sample: 600

	Percentage of Children
Age (in years and months)	
3-0 to 3-5	33.3
3-6 to 3-11	33.3
4-0 to 4-5	33.3
Gender	
Male	50.0
Female	50.0
Race/Ethnicity	
White	57.2
African American	15.2
Hispanic	21.8
Other	.06
Maternal Education	
Grade 11 or Less	15.8
High School Graduate	28.5
1-3 Years of College	28.5
4 or More Years of College	27.2
Region	
Northeast	18.2
North Central	39.3
South	21.3
West	21.2
Language Version	
English	86.6
Spanish	13.3

Characteristics of ESI-K 2008 Norming Sample
Number of children in the sample: 600

	Percentage of Children
Age	
4 years, 6 months to 4 years, 11 months	33.3
5 years to 5 years, 5 months	33.3
5:6-5:11	33.3
Gender	
Male	50.0
Female	50.0
Race/Ethnicity	
White	57.5
African American	15.3
Hispanic	21.6
Other	.05
Maternal Education	
Grade 11 or less	16.2
High School graduate	27.2
1-3 years of college	28.3
4 or more years of college	28.3
Region	
Northeast	19.6
North Central	29.8
South	26.5
West	24.0
Language Version	
English	88.2
Spanish	11.8

Reliability and Validity Information

What is known about the reliability and validity of the developmental screener…

In English?

There is information about the reliability, validity, sensitivity, and specificity of the ESI-R in English. This information is outlined in responses to later questions in this profile.

In other languages?

The developers have not examined the reliability, validity, sensitivity, and specificity of the Spanish-language version of the ESI-R.

For dual language learners?

The developer does not provide any information about this population, and the reliability, validity, sensitivity, and specificity of the ESI-R for dual language learners have not been examined.

For children with special needs?

The developer does not provide any information about this population, and the reliability, validity, sensitivity, and specificity of the ESI-R for children with special needs have not been examined.

For American Indian/Alaskan Native children?

The developer does not provide any information about this population, and the reliability, validity, sensitivity, and specificity of the ESI-R for American Indian/Alaskan Native children have not been examined.

For children of migrant and seasonal farm workers?

The developer does not provide any information about this population and the reliability, validity, sensitivity, and specificity of the ESI-R for children of migrant and seasonal farm workers have not been examined.

Reliability: Does the instrument obtain the same results, consistently, under the same conditions with the same children?

Interrater reliability. *Do different raters agree when they are assessing the same children?*

Agreement between raters when they are screening the same children was tested during the first standardization of the ESI-P and ESI-K. The initial ESI-P sample included 977 children. Approximately 53 percent of the children in this sample were White, 21 percent were African American, and 25 percent were included in an Other category for race/ethnicity. The parents of more than 25 percent of the sample had not completed high school. The majority of the children attended Head Start programs, and the remaining children attended either public or private preschools or child care. The initial ESI-K sample included 5,034 children. Approximately 70 percent of this sample was White (non-Hispanic) and 30 percent were non-White children. The mothers of 20 percent of the children had not completed high school.

For the ESI-P, both an administrator and an observer independently scored the ESI-P as it was being administered. Results from 35 administrator-observer pairs showed that agreement between the two raters was acceptable when screening the same child. Agreement between two raters was also tested during the standardization of the ESI-K. Results from 586 administrator-observer pairs who scored the same child simultaneously showed acceptable agreement between the two raters.

Test-Retest Reliability. *How consistent are scores if the developmental screener is administered once and then administered again soon?*

Scores on the ESI-K were studied to determine how consistent they are if the screener is administered once and then administered again soon. Two different administrators used the ESI-K with the same child 7 to 10 days apart. The results showed acceptable consistency from the first to the second administration. One hundred seventy four children from the initial ESI-K norming sample were used in this analysis.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

The developers have not examined relationships between items that are intended to reflect the same set of skills or behaviors.

Validity: Does the instrument measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the developmental screener do a good job of reflecting what the screener is supposed to be assessing?*

The developers do not provide information about whether experts agree that the items in the ESI-R do a good job of reflecting what the ESI-R is supposed to be measuring.

Construct Validity. *How closely related to each other are sets of items within the developmental screener that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

The developers do not provide information about relationships between sets of items on the ESI-R.

Yes, the developers examine whether scores on sets of items relate to children's age as expected.

Convergent Validity. *How strongly do the scores of this developmental screener show a relationship to the scores of other developmental screeners of similar domains?*

The developers have not examined relationships between a child's scores on the ESI-R and his or her scores on other developmental screeners of similar domains.

Scores for Further Evaluation. *Are specific scores used to identify whether further evaluation is needed? How are these scores determined by the developer?*

Yes, specific scores (called cutoff scores) are used to identify whether further evaluation is needed. The cutoff scores were developed with the norming sample of the original version of the ESI. To develop the cutoff scores, a method of analysis was used that compares the probability of getting an accurate result (indicating that the child is either "okay" or "at risk") for a range of cutoff scores. For this analysis, ESI scores were compared with scores on the General Cognitive Index of the McCarthy Scales of Children's Abilities for 251 children. The results of these screenings, performed 7 to 9 months apart, were then compared for this sample of children in order to determine appropriate cutoff points for the ESI. These initial cutoff scores were reexamined during the standardization of the 2008 version of the ESI-R. The cutoff scores identified approximately the same percentage of children in both samples as at risk for developmental problems, indicating that the cutoffs can continue to be used with the 2008 edition.

Sensitivity. *How accurately does the developmental screener correctly identify children at risk for developmental problems?*

Both the ESI-P and the ESI-K are highly accurate in correctly identifying children at risk for developmental problems.

Specificity. *How accurately does the developmental screener correctly identify children who are not at risk for developmental problems?*

The ESI-P and the ESI-K are moderately accurate in identifying children who are not at risk for developmental problems.

Follow-Up Guidance

Program Follow-Up Steps. *Does the developmental screener come along with guidance about follow-up steps based on the results?*

Children who score in the "OK" category are considered to be developing normally and are not in need of further assessment. Children who score in the "rescreen" category have borderline ESI-R scores. The manual suggests that the ESI-R should be readministered to these children in 8 to 10 weeks. If a child's score is in the "refer" category, he or she should be evaluated by an assessment team and, if the problems identified in the screening are confirmed, a definitive plan of action or individualized education plan should be developed.

Family Follow-Up Steps. *Does the developmental screener include recommendations on how families might follow up on the results of the screening?*

The developers do not provide information regarding recommendations for how families might follow-up on the results of the screening.

References

Meisels, S.J., Marsden, D.B., Wiske, M.S. & Henderson, L.W. (2008). Early Screening Inventory-Revised (ESI-R) examiner's manual, 2008 Edition. Minneapolis, MN: NCS Pearson, Inc.

Learning Accomplishment Profile-Diagnostic Screens (LAP-D Screens)

Developer: The Chapel Hill Training-Outreach Project
Publisher: Kaplan Early Learning Company

http://chtop.org/Products/The-LAP-D-Screens.html

Developmental domains addressed in the developmental screener, as stated by the publisher:

- Gross motor
- Fine motor
- Cognitive
- Language

Intended age range: 3 to 5 years. There are three versions of the LAP-D Screens: one each for 3-, 4-, and 5-year-olds. The 5-year-old version is meant for children in kindergarten.

Number of items: The number of items depends on the version. The 3-year-old version has 18 items. The 4-year-old version has 55 items. The 5-year-old version has 25 items. However, not all items are administered to all children.

In what settings can this developmental screener be used (e.g., centers, homes, medical facilities, other)?

The LAP-D Screens can be used in early childhood programs, universities, research laboratories, hospitals, rehabilitation centers, and other medical practices.

Background

Purpose:

The Learning Accomplishment Profile-Diagnostic Screeners (LAP-D Screens) is a brief developmental screener that provides an initial snapshot of whether a child might be at risk for a developmental delay. Four of the tools that are included in this document are from the Learning Accomplishment System (LAP). The four tools are distinct from each other, but are from a comprehensive system of assessment and developmental screening. The Learning Accomplishment System-3rd Edition (LAP-3) is a criterion-referenced assessment, too, meaning that a child's scores on the assessment are compared to developmental benchmarks. The Learning Accomplishment System-Diagnostic (LAP-D) is not a diagnostic tool, but is a norm-referenced assessment, meaning that a child's scores on the assessment are compared to the scores of a group of children with which the assessment was developed and on which it was tested. There is a separate profile for the LAP-D assessment in Spanish. Finally, there is a profile for the Learning Accomplishment System-Diagnostic Screener (LAP-D Screen), a shorter version of the LAP-D assessment that is used for screening for potential developmental delays.

What is the appropriate time period between administering, recording, or reviewing the data?

The developer does not provide information regarding the appropriate time period between initial screening and rescreening.

How long does it take to administer the developmental screener?

The LAP-D Screens takes about 10-15 minutes to administer; however, administration time depends on the child's age and ability.

Language(s) developed for:

The LAP-D Screens was developed for English-speaking children. The materials are also available in Spanish.

Availability and Cost of the Developmental Screener

Is the developmental screener available to programs without restrictions?

Yes, the developmental screener is available to programs without restrictions.

What is the cost of the developmental screener?

As of 2010, each version of the LAP-D Screens cost $124.95. Additional records sheets can be purchased for an additional $30. A complete kit that includes all three screening levels (ages 3 to 5 years) costs $349.95.

Training and Other Requirements for Assessors

Is training available on how to administer and score the developmental screener? Who offers the training?

Yes, Kaplan Early Learning Company offers training on the LAP System. Information is available on the Kaplan website (http://www.kaplanco.com/services/profDev_onSiteTraining.asp), however the website does not detail which LAP tools are covered in the training.

Is it necessary to have a professional background or technical training over and above training on the developmental screener to administer or complete the developmental screener?

Yes, it is necessary to have a professional background to administer and complete the LAP-D Screens. Teachers can administer the LAP-D Screens, but they must have at least a Child Development Associate (CDA) credential. Additionally, the LAP-D Screens can be administered by clinical psychologists, school psychologists, occupational and physical therapists, physicians, nurses, and social workers.

Is it necessary to have a professional background or technical training over and above training on the developmental screener to score the developmental screener?

Anyone who can administer the LAP-D Screens can score it.

Are regular checks on faithful administration required or recommended to ensure appropriate administration? If so, when and by whom?

Regular checks on faithful administration are recommended but not required. The developers do not provide additional information regarding when to perform regular checks on administration or who should perform these checks.

Information Reporting System for the Developmental Screener

Electronic Data Entry. *Does the developmental screener come with a process for entering information from the developmental screener electronically?*

Yes, the LAP-D Screens information can be entered electronically, but the software must be purchased in addition to the materials needed to administer the measure. The information can be entered on a computer or on a hand-held electronic scoring pad.

Electronic Reports. *Can programs generate electronic reports of individual children's data?*

Yes, programs can generate electronic reports of the LAP-Screens information at the child level. There is also an electronic parent report.

Approaches to Family/Parent Input

Tools for Family Input. *Does the developmental screener include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

No, there is no specific information about gathering information from parents or family members about the child.

Sharing Results. *Does the developmental screener include recommendations on how to share developmental screener results with a child's family?*

No, there are no recommendations on how to share the results with a child's family.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this a developmental screener with developmental norms?*

The LAP-D Screens is a screener with developmental norms. A program director can choose to use the norms presented by the developers or can establish local norms, which would be centered around the type of children the program serves and who is being screened with the LAP-D Screens. However, the manual suggests that a program director consult a measurement specialist if local norms will be established.

Which populations were included in the norming sample?

The LAP-D Screens norms were developed with a group of 907 children ages 3 to 5. The children were from the Northeast (29 percent), North Central (13 percent), West (13 percent), and South (45 percent) regions of the United States. See the table on the next page for more information about these children.

Availability of Versions in Languages Other than English. *Is the developmental screener available in languages other than English? Which languages?*

Yes, the LAP-D Screens have been translated into Spanish.

How were versions in languages other than English developed?

The developer does not provide information about how the Spanish-language version was developed.

What are the findings on the reliability and validity of versions other than English?

The developer does not provide information about the development of the Spanish-language version of the LAP-D Screens.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

No, there are no suggested accommodations for screening children with identified or suspected special needs.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

The developer does not provide information regarding whether the appropriateness of the LAP-D Screens for diverse populations has been examined in this way.

Risk Levels. *What terminology is used to describe risk levels (e.g., delay, no delay, at risk, caution, rescreen, okay, etc.)?*

The only terminology used by the LAP-D Screens is "pass" and "refer." If a child passes the screener, it indicates that at the time, he or she is not at risk for developmental delay. If a child is given a "refer" on a certain number of items, which depends on age and the cutoff score being used, then the child should be evaluated further.

Characteristics of 1996 Norming Sample
Number of children in the sample: 907

	Percentage of Children		
	3 year olds	**4 year olds**	**5 year olds**
Gender			
Male	44.4	51.6	47.4
Female	55.6	48.4	52.6
Race/Ethnicity			
White	46.2	55.3	53.7
African American	36.4	24.4	25.8
Hispanic	15.3	14.0	14.3
Other	2.1	6.3	6.2

Reliability and Validity Information

What is known about the reliability and validity of the developmental screener…

In English?

There is reliability, validity, and sensitivity information for the LAP-D Screens in English. This information is outlined in responses to later questions in this profile.

In other languages?

The LAP-D Screens is available in Spanish; however, the developers have not examined the reliability, validity, sensitivity, and specificity for the Spanish-language version.

For dual language learners?

The developer does not provide information about dual language learners, and have not examined the reliability, validity, sensitivity, and specificity for this population.

For children with special needs?

The developers have not examined the reliability, validity, sensitivity, and specificity for children with special needs.

For American Indian/Alaskan Native children?

While American Indian/Alaskan Native children were included in the sample (1-2 percent of children), the developers have not examined the reliability, validity, sensitivity, and specificity for this population.

For children of migrant and seasonal farm workers?

The developers do not provide information about migrant and seasonal farm workers, and have not examined the reliability, validity, sensitivity, and specificity for this population.

Reliability: Does the instrument obtain the same results, consistently, under the same conditions with the same children?

Interrater Reliability. *Do different raters agree when they are assessing the same children?*

Yes, different raters agree when they are screening the same children and the relationships meet the criteria for acceptable. The raters agreed most strongly when using the LAP-D Screens with 5-year-old children. This was examined with 18 3-year-olds, 21 4-year-olds, and 13 5-year-olds. There is no information about the teachers who administered the LAP-D Screens.

Test-Retest Reliability. *How consistent are scores if the developmental screener is administered once and then administered again soon?*

The LAP-D Screens meets the criteria for acceptable when it is administered once and then administered again soon. In order to examine this, the LAP-D Screens was administered twice within a two- to three- week period (with an average of 14 days between the screener administrations). The scores were all very consistent, but the 4-year-old developmental screener was the most consistent. No information is provided about the teachers and children with whom this was examined.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

There are acceptable relationships among items on the LAP-D Screens that are intended to reflect the same set of skills or behaviors. The relationships were stronger with the 4- and 5-year-old versions of the LAP-D Screens than with the 3-year-old version. No information is provided about the teachers and children with whom this was examined.

Validity: Does the instrument measure what it is supposed to?

Content Validity. *Were experts consulted regarding whether the items in the developmental screener do a good job reflecting what the developmental screener is supposed to be assessing?*

The items that are in the LAP-D Screens are taken from the LAP-D Assessment, also profiled in this document. Experts agreed upon the items in the LAP-D Assessment and agree that they reflect what the tool is supposed to measure. However, experts were not consulted separately about the items on the LAP-D Screens.

Construct Validity. *How closely related to each other are sets of items within the developmental screener that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do the developers examine whether scores on sets of items relate to children's age as expected?*

While the manual states that sets of items within the LAP-D Screens are related, specific information about how closely they are related is not provided.

Information about whether scores on sets of items relate to children's age as expected is not provided.

Convergent Validity. *How strongly do the scores of this developmental screener show a relationship to the scores of other developmental screeners of similar domains?*

The LAP-D Screens shows a strong relationship when compared to the LAP-D Standardized Assessment, which is a comprehensive assessment for children between the ages of 30 and 72 months. However, it should be noted that many of the items on the LAP-D Screens are taken from the LAP-D Standardized Assessment, so there is overlap between the two tools.

Scores for Further Evaluation. *Are specific scores used to identify whether further evaluation is needed? How are these scores determined by the developer?*

Yes, specific scores are used to identify whether a child may need further evaluation. There are different levels from which a program may choose. If the program chooses a higher level cutoff score, then more children will be recommended for further evaluation. If the program chooses a lower cutoff score, then fewer children will be recommended for further evaluation. The cutoff scores are determined by looking at the average score of the children in the appropriate age range (3, 4, or 5 years old). These averages are given in the manual and are based on the distributions of scores in the norming sample. Then, there are certain levels below this average score that can be used for the cutoff. These scores vary by the age of the child.

Sensitivity. *How accurately does the developmental screener correctly identify children at risk for developmental problems?*

The LAP-D Screens is moderate to highly accurate at correctly identifying children at risk for developmental delay. To test this, the LAP-D Screens was compared to the Early Screening Profile (ESP), which is a comprehensive developmental screener used with children from 2 to 7 years of age. The ESP was administered to 84 children from the larger study group. The LAP-D Screens and the ESP identified children in the same way (either passed or referred) 83 percent of the time.

Specificity. *How accurately does the developmental screener correctly identify children who are not at risk for developmental problems?*

The developers do not provide information on how accurate the LAP-D Screens is at identifying children who are not at risk for developmental delay.

Follow-Up Guidance

Program Follow-Up Steps. *Does the developmental screener come with guidance about follow-up steps based on the results?*

No, the developmental screener does not come with guidance about follow-up steps based on the results.

Family Follow-Up Steps. *Does the developmental screener include recommendations on how families might follow up on the results of the screening?*

No, the developmental screener does not include recommendations on how families might follow up on the results of the screening.

References

Lockhart, C.S. (1997). Technical and Examiner's Manual for LAP-D Screens: 5 Year Old (Kindergarten), 3 Year Old and 4 Year Old Children. Chapel Hill, NC: Chapel Hill Training Outreach Project, Inc. Kaplan Early Learning Company.

Parents' Evaluation of Developmental Status (PEDS)

Developer: Frances P. Glascoe
Publisher: Ellsworth & Vandermeer Press LLC

http://www.pedstest.com/

Developmental domains addressed in the developmental screener, as stated by the publisher:

- Global/cognitive
- Expressive language and articulation
- Receptive language
- Fine motor
- Gross motor
- Behavior
- Social-emotional
- Self help
- School

Intended age range: Birth through age 8

Number of items: PEDS includes 10 items that are the same for all children.

In what settings can this developmental screener be used (e.g., centers, homes, medical facilities, other)?

PEDS can be used in many settings, including medical practices, clinics and other primary care facilities, public health departments, Child Find programs, Head Start or other early childhood programs, pediatric and other professional training programs, and research projects.

Background

Purpose:

PEDS is a developmental screener used to help detect early developmental and behavioral problems. PEDS relies on parent-completed questionnaires to gather information about how a child is developing. It is used to gather information about specific areas of child development and to see if further evaluation may be needed. PEDS can be used with a related measure called PEDS-Developmental Milestones (PEDS-DM), which has a separate profile in this document, but will be referred to in this profile.

What is the appropriate time period between administering, recording, or reviewing the data?

PEDS follows the guidelines of the American Academy of Pediatrics, which recommends setting up a regular screening schedule with a child's pediatrician.

How long does it take to administer the developmental screener?

PEDS takes under 30 minutes for parents to complete.

Language(s) developed for:

The developmental screener was developed for English-speaking families, but there are forms available in 14 different languages.

Availability and Cost of the Developmental Screener

Is the developmental screener available to programs without restrictions?

Yes, the developmental screener is available to programs without restrictions.

What is the cost of the developmental screener?

As of 2010, a starter kit for PEDS cost $36 and includes 50 PEDS response forms, 50 reusable score/interpretation forms, and a 12-page brief guide to scoring and interpreting results. PEDS in print is available in English, Spanish, and Vietnamese. Additional translations into Arabic, Chinese, Farsi, French, Galician, Haitian-Creole, Hmong, Indonesian, Malaysian, Portuguese, Russian, Somali, Swahili, Thai, and Taiwanese have been requested by programs and completed through a contract with PEDS publishers.

An optional comprehensive manual, Collaborating with Parents, includes information on brief approaches to parent intervention, background research on relying on the parent report, and PEDS' psychometrics. It is available for $79.95.

Training and Other Requirements for Assessors

Is training available on how to administer and score the developmental screener? Who offers the training?

Self-training for those who ask parents to complete the PEDS is available on the PEDS website (http://www.pedstest.com/default.aspx) in the form of videos, slide shows, and case examples. Live training or contacts with local professionals are often available.

Is it necessary to have a professional background or technical training over and above the training on the developmental screener to administer or complete the screener?

No, it is not necessary. The PEDS response form is usually completed by a parent rather than a teacher. Teachers or examiners score the PEDS and are encouraged to add their own observations before scoring.

Is it necessary to have a professional background or technical training over and above the training on the developmental screener score the screener?

No, a teacher, administrator, or other professional familiar with the PEDS can score the developmental screener without a technical background or training as long as they adhere to the PEDS brief guide when scoring or make use of PEDS Online.

Are regular checks on faithful administration required or recommended to ensure appropriate administration? If so, when and by whom?

Since the PEDS is usually completed by a parent or family member, regular checks of faithful administration are not necessary. However, teachers and examiners must faithfully use the PEDS brief guide to scoring and administration if they are scoring the screener by hand. PEDS Online corrects for common errors that may arise during administration (e.g., it prompts users if nothing is written on the PEDS response form for an item suggesting parents may not have understood the questions, skipped items, etc.).

Information Reporting System for the Developmental Screener

Electronic Data Entry. *Does the developmental screener come with a process for entering information from the developmental screener electronically?*

Yes, information from the PEDS can be entered and scored online. There is a parent portal on the website that allows parents to complete the forms on their own. The results are then sent to the doctor or other professional who will speak with the parents about the results. Additionally, there are other features for administrators to enter data, and view, export, and sort results (e.g., by name, school/clinic, teacher/examiner, birthdate, etc.).

Electronic Reports. *Can programs generate electronic reports of individual children's data?*

Yes, reports can be generated electronically using PEDS Online. Reports can be generated by child or by risk group based on the results of the developmental screener. (More information about risk groups is provided below.) A database of all results can be exported for use with EXCEL or other statistical software.

Approaches to Family/Parent Input

Tools for Family Input. *Does the developmental screener include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

The developmental screener mainly comprises input from parents or other caregivers on various developmental skills. Teachers and examiners are encouraged to add their own observations (but these observations cannot detract from or override those from families).

Sharing Results. *Does the developmental screener include recommendations on how to share developmental screener results with a child's family?*

Yes, the developmental screener comes with very extensive recommendations on how to share the screening results with a child's family.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this a developmental screener with developmental norms?*

Yes, the PEDS is a screener with developmental norms. The norms were created based on a sample of families from five sites selected to represent the broad geographic regions of the U.S. According to the developer, the characteristics of this sample were comparable to U.S. Census data from 1996.

Which populations were included in this norming sample?

The development norms were developed with 771 families from five cities across the United States. Families were recruited from education programs and pediatric practices, but the majority were from education programs. About half (53.7 percent) of the children were male, and 69.8 percent of the children had parents who were married. Children ranged in age from birth to age 8. See the table on the page after next for more information about these children.

Availability of Versions in Languages Other than English. *Is the developmental screener available in languages other than English?*

The developmental screener was developed in English, but has been translated into 17 other languages.

Options for Use with Special and Diverse Populations

How were versions in languages other than English developed?

All translations were developed with a group of bilingual professionals with a background in child development. The developers do not provide additional information regarding the development of the PEDS in other languages.

What are the findings on the reliability and validity of versions of the developmental screener in languages other than English?

The reliability and validity for versions other than English have not been examined.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

Because the PEDS relies on parents' concerns and observations, accommodations for children with identified or suspected special needs are not needed.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations?*

The developer does not provide information regarding whether cognitive testing or focus groups have been conducted with diverse populations to determine the appropriateness of the screener.

Risk Levels. *What terminology is used to describe risk levels (e.g., delay, no delay, at risk, caution, rescreen, okay, etc.)?*

There are five categories used to describe risk levels based on the results of the PEDS (follow-up steps based on risk levels are described in the next question).

Path A
Children who receive two or more predictive concerns on the PEDS fall into Path A. This is the high risk group. Path A is also divided into two subgroups depending on patterns of concerns which will indicate whether speech-language, developmental psychology or autism specialists are needed. Teachers and examiners are encouraged to use their observations to add to referral recommendations. The manual indicates that about 1 in every 10 children screened will fall into Path A, although the rate will vary depending upon the population being screened.

Path B
Children who receive one predictive concern on the PEDS fall into Path B. This is the moderate risk group. Path B is also divided into two subgroups depending on whether the concerns are mainly health related (for which a referral for medical care is needed) or non-health related (for which follow-up screening is recommended, such as with the 6- to 8- question PEDS-Developmental Milestones). If additional screening is passed, developmental promotion–i.e., teaching parents how to teach their children well–and careful monitoring is recommended. About 2 in every 10 children screened will fall into Path B, although the rate will vary depending upon the population being screened.

Path C
Children who have nonpredictive concerns on the PEDS fall into Path C. This is the elevated risk group for behavioral and mental health problems, but these children are often at low risk of a developmental disability. Path C is divided into two subgroups based on the child's age (younger or older than 4 ½ years). For younger children, the PEDS recommends that parenting guidance is needed, along with careful monitoring of progress. For older children, mental health risks are higher and so mental health screening or referrals for services and evaluations are needed. The manual indicates that about 2 in every 10 children screened will fall into Path C, although the rate will vary depending upon the population being screened.

Path D
Children whose parents or family members have difficulty communicating their concerns on the PEDS forms fall into Path D. The recommendation here is either to repeat the PEDS via interview or to use a measure like PEDS: Developmental Milestones. About 3 percent of families fall into Path D. This problem occurs less often with online administration of the PEDS because there are prompts asking for written responses and when a parent has missed an item, although the rate will vary depending upon the population being screened.

Path E
Children with no concerns fall into Path E. The manual indicates that about 5 in every 10 children screened will fall into Path E, although the rate will vary depending upon the population being screened.

Characteristics of Norming Sample
Number of children in the sample: 771

	Percentage of children
Race	
White	64.5
African American	21.5
Hispanic/Other	14.0
Parental Education	
Less than High School	18.0
High School	31.5
High School and Some College	22.6
College	27.9
Family Income Level	
Low Income	25.4
Not Low Income[17]	74.6
Parental Employment Status	
Full-Time	48.5
Part-Time	18.5
Unemployed	33.1

[17] Low income is defined by meeting one of the following criteria: child participated in free or reduced meals at school, child was enrolled in a federally subsidized child care program, or the characteristics of the child's family are consistent with the characteristics of families falling into the first two categories.

Reliability and Validity Information

What is known about the reliability and validity of the developmental screener…

In English?

There is reliability, validity, sensitivity, and specificity information for the PEDS in English. This information is outlined in later questions of this profile.

In other languages?

The developers have not examined the reliability, validity, sensitivity, and specificity information for the PEDS in languages other than English.

For dual language learners?

The developers do not provide information about dual language learners and have not examined the reliability, validity, sensitivity, and specificity of the PEDS for this population.

For children with special needs?

While children with special needs were included in the norming sample, the developers have not examined the reliability, validity, sensitivity, and specificity of the PEDS for children with special needs.

For American Indian/Alaskan Native children?

The developers do not provide information about American Indian/Alaskan Native children and have not examined the reliability, validity, sensitivity, and specificity of the PEDS for this population.

For children of migrant and seasonal farm workers?

The developers do not provide information about children of migrant and seasonal farm workers and have not examined the reliability, validity, sensitivity, and specificity of the PEDS for this population.

Reliability: Does the instrument obtain the same results, consistently, under the same conditions with the same children?

Interrater Reliability. *Do different raters agree when they are assessing the same children?*

Yes, different raters agree when they are screening the same children. In order to test this, the PEDS was completed by parents or family members and then interpreted by a trained rater. The raters interpreted the information the same way an average of 95 percent of the time. Additionally, the developers looked at whether parents give the same information based on who interviewed them, if the developmental screener was administered orally. Parents gave the same information 88 percent of the time. The PEDS evaluations for 68 percent of children in the PEDS standardization sample were examined and summarized by pairs of trained raters. These children were enrolled in education programs including Head Start, subsidized day care, and private preschools; however, the developers do not provide further detail.

Test-Retest Reliability. *How consistent are scores if the developmental screener is administered once and then administered again soon? What about much later?*

When the developmental screener was given two times, with a two-week period in between administrations, the scores met the criteria for adequate consistency. The scores were the same an average of 88 percent of the time. This was examined with a subsample of 20 percent of the parents from the group described in the previous table. Parents were first given the PEDS during a pediatric encounter, such as a well-child visit, and then were given the PEDS over the phone the second time.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

Overall, the items that are intended to reflect the same set of skills and behaviors meet the criteria for acceptable relationships. The items reflecting fine motor skills and gross motor skills have strong relationships. The self-help and motor skills items also have strong relationships. This was examined with the population described in the table. The developers do not provide any additional information about the population.

Validity: Does the instrument measure what it is supposed to?

Content Validity. *Do experts agree that the items in the developmental screener do a good job of reflecting what the developmental screener is supposed to be assessing?*

Yes, experts agree the PEDS does a good job at reflecting what it is supposed to be measuring.

Construct Validity. *How closely related to each other are sets of items within the developmental screener that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do scores on sets of items relate to children's age as expected?*

The developers have not examined the relationships among sets of items that address the same skills and behaviors in comparison with different skills and behaviors.

Information about whether scores on sets of items relate to children's age as expected is not provided.

Convergent and Divergent Validity. *How strongly do the scores of this developmental screener show a relationship to the scores of other developmental screeners of similar domains?*

The PEDS was compared with 14 other developmental assessments and screeners. There were strong relationships between many of the developmental areas of the PEDS and developmental areas of the comparison tools aimed at measuring the same skills and behaviors. Developmental areas were most strongly related on the following tools: Child Development Inventory (including socialization self-help, gross motor, fine motor, expressive language, and listening comprehension), Kaufman Assessment Battery for Children (diagnostic measure of intelligence), Bayley Scales of Infant Development (mental development index), Stanford-Binet Intelligence Scale, 4th Edition (diagnostic measure of intelligence), Test of Language Development (expressive and receptive language skills), Developmental Profile-II (parent report measure of socialization, communication, academic self-help, and motor development), Brigance Screens (short screening test), and Batelle Developmental Inventory Screening Test.

Several developmental areas of the PEDS were not strongly related to other developmental assessments or screeners aimed at measuring different skills and behaviors, providing evidence of divergent validity. For example, the gross and fine motor developmental areas of the PEDS were not strongly related to the Articulation Screening Test, which is a screener aimed at measuring speech production.

Scores for Further Evaluation. *Are specific scores used to identify whether further evaluation is needed? How are these scores determined by the developer?*

Yes, specific scores are used to identify whether further evaluation is needed. In order to develop these specific scores, the PEDS was completed by 711 parents as described in the table earlier in this profile. The developers looked at the trends among the responses from these parents and examined the outcome of the screener based on the parents' responses. This created five distinct cutoff scores that are used to identify whether further evaluation is need. See the question on the terminology used to describe risk levels (below) for more information about these cutoff scores and what they indicate about a child's development.

Sensitivity. *How accurately does the developmental screener correctly identify children at risk for developmental problems?*

To determine how accurately PEDS identifies children at risk for developmental delays, the results of children's diagnostic tests were compared to the concerns that parents identified on PEDS. Results showed that PEDS is moderately accurate at correctly identifying children who are at risk for developmental delays.

Specificity. *How accurately does the developmental screener correctly identify children who are not at risk for developmental problems?*

To determine how accurately PEDS identifies children who are not at risk for developmental delays, results of children's diagnostic tests were compared to the absence of parental concerns on PEDS. Results showed that PEDS is moderately accurate at correctly identifying children who are not at risk for developmental delays.

Follow-Up Guidance

Program Follow-Up Steps. *Does the developmental screener come with guidance about follow-up steps based on the results?*

Yes, the developmental screener comes with guidance about follow-up steps based on the path on which the child is placed, as explained in a previous question.

When a child is on Path A, multiple concerns are present and the child should be referred for further evaluation. This may include, for example, audiological (speech and language) testing or another form of educational evaluation that is deemed necessary by a professional. If a child is placed on Path B, one main concern is present. These children should be further evaluated using a health screener and/or the PEDS-DM. Follow up for a Child on Path C includes screening in which areas parents raised concerns and counseling parents about their concerns since issues for these children are nonpredictive and not as severe. For children on Path D, the PEDS-DM should be administered since the parents had difficulty communicating their concerns or lack of concerns. Finally, for children on Path E, screening with PEDS should take place at the next doctor's visit or during regular yearly screenings since there are no concerns.

Family Follow-Up Steps. *Does the developmental screener include recommendations on how families might follow up on the results of the screening?*

Yes, the developmental screener includes extensive recommendations on how parents might follow up on the results of the screening.

References

Glascoe, F.P. (2002). Collaborating with parents: Using parents' evaluation of developmental status (PEDS) to detect and address developmental and behavioral problems. Nashville, TN: Ellsworth & Vandermeer Press LLC.

Glascoe, F.P. (2007). Using parents' evaluation of developmental status (PEDS) and PEDS-Developmental milestones (PEDS-DM): A case example. Retrieved November 17, 2009, from http://pedstest.com/dm/casestudy-1.php

Glascoe, F.P. (2008). Parents' evaluation of developmental status: Guide to using PEDS. Retrieved November, 2009, from http://www.forepath.org/custom.php

Glascoe, F.P. (2009). Parents' evaluation of developmental status: Brief administration and scoring guide. Nashville, TN: Ellsworth and Vandermeer Press, LLC.

Wake, M., Gerner, B., & Gallagher, S. (2005). Does parent's evaluation of developmental status at school entry predict language, achievement and quality of life two years later? *Ambulatory Pediatrics, 5*(3), 143-149.

What is PEDS-Developmental Milestones (PEDS-DM©)? Retrieved November 17, 2009, from http://pedstest.com/dm/dm-whatisit.php

Parents' Evaluation of Developmental Status-Developmental Milestones (PEDS-DM)

Developers: Frances Page Glascoe and Nicholas Robertshaw
Publisher: PEDSTest.com (formerly Ellsworth & Vandermeer Press LLC)

http://pedstest.com

Developmental domains addressed in the developmental screener, as stated by the publisher:

- Expressive language
- Receptive language
- Fine motor
- Gross motor
- Social-emotional
- Self help
- Academic
- Prereading
- Premath
- Written language

Intended age range: Birth through age 7 years, 11 months

Number of items: The PEDS-DM screen has 6 to 8 items per age. The PEDS-DM Assessment Level involves about 45 items per age.

In what settings can this developmental screener be used (e.g., centers, homes, medical facilities, other)? The PEDS-DM can be used in many settings, including medical practices, subspecialty health clinics, primary care services including public health departments, Child Find programs, Head Start or other early childhood programs, pediatric and other professional training programs, and research projects.

Background

Purpose:
PEDS-DM is a 6- to 8- item screener that tracks a child's development in several domains. The PEDS-DM screener can be administered by parent report, parent-child interview, or direct administration with the child. It tracks progress over time on a recording form with multiple time periods, through which strengths and weakness in various domains become apparent. The PEDS-DM can be used with the PEDS developmental screener (to capture parents' concerns) or separately, but the developers recommend using them together to get a full picture of a child's development. There is a separate profile of PEDS in this document.

What is the appropriate time period between administering, recording, or reviewing the data?
PEDS-DM follows the guidelines of the American Academy of Pediatrics, which recommends setting up a regular screening schedule with a child's pediatrician.

How long does it take to administer the developmental screener?
The PEDS-DM screen takes about five minutes for families to complete.

Language(s) developed for:
The PEDS-DM was developed with English- and Spanish-speaking families and the screener is available in both languages. PEDSTest.com offers research/translation support and financial assistance for translations into other languages. For example, a Taiwanese translation was requested by programs and was completed through a contract with the PEDS-DM publisher. Arabic and Portuguese translations are under way.

Availability and Cost of the Developmental Screener

Is the developmental screener available to programs without restrictions?

Yes, the developmental screener is available to programs without restrictions.

What is the cost of the developmental screener?

As of 2010. the PEDS-DM Screen Starter Kit cost $275; this includes the manual, materials needed to screen children, and 100 reusable record sheets. Additional packs of 100 forms are available for $32 each. The PEDS-DM Screener with PEDS cost $315. The PEDS-DM Assessment Level cost $318 alone, and with the PEDS $399. The Starter Kit is also available in Spanish.

Training and Other Requirements for Assessors

Is training available on how to administer and score the developmental screener? Who offers the training?

The PEDS-DM website, Pedstest.com, offers self-training through videos and slide shows. Live training may also be available, but there is not information in the manual.

Is it necessary to have a professional background or technical training over and above training on the assessment to administer or complete the developmental screener?

No, the PEDS-DM is best completed by a parent rather than a teacher or professional. If the PEDS-DM is completed by a parent, it may be necessary for a professional to give parents guidance and directions on completing the forms. This professional could be anyone from the list of applicable settings mentioned earlier. The developmental screener can also be completed by a professional, if necessary, through observations of the child and the child's behavior. If the PEDS-DM is completed through direct observation, some training needs to be completed. This training material is available from the PEDS and PEDS-DM websites.

Is it necessary to have a professional background or technical training over and above training on the assessment to score the developmental screener?

No, a teacher, administrator, or other professional can score the developmental screener without a technical background or training.

Are regular checks on faithful administration required or recommended to ensure appropriate administration? If so, when and by whom?

Since the PEDS-DM is usually completed by a parent or family member, regular checks of faithful administration are not necessary.

Information Reporting System for the Developmental Screener

Electronic Data Entry. *Does the developmental screener come with a process for entering information from the developmental screener electronically?*

The PEDS-DM is available online. The site provides automated scoring, summary reports for parents, referral letters, billing and procedure codes for optimizing reimbursement, and a searchable administration database (e.g., by birth date, date of test, type of result, etc.).

Electronic Reports. *Can programs generate electronic reports of individual child's data?*

Yes, child-level reports can be generated electronically.

Approaches to Family/Parent Input

Tools for Family Input. *Does the developmental screener include specific tools or guidance for gathering and incorporating parental/family input on an individual child's skills and development?*

Yes, the PEDS-DM is a parent/family report developmental screener.

Sharing Results. *Does the developmental screener include recommendations on how to share developmental screener results with a child's family?*

Yes, the developmental screener comes with extensive recommendations on how to share the screening results with a child's family.

Options for Use with Special and Diverse Populations

Developmental Norms. *Is this a developmental screener with developmental norms?*

Yes, the PEDS-DM is a screener with developmental norms. However, the items for the PED-DM were selected from the BRIGANCE® Inventory of Early Development-II (IED-II), created in 2004, and the BRIGANCE® Comprehensive Inventory of Basic Skills-Revised (CIBS-R), developed in 1999. Thus, the norms for the PEDS-DM are based on the norms for these two other tools.

Which populations were included in the norming sample?

Data from all children who participated in the IED-II norming study and all children between 5 and 8 years of age in the CIBS-R norming study were used in the norming sample for the PEDS-DM. In total, there were 1,619 children. This PEDS-DM norming sample was compared to U.S. demographics using 2006 data from the U.S. Census Bureau and is considered representative of the U.S. population as a whole. More information is provided in the table on the next page.

Availability of Versions in Languages Other than English. *Is the developmental screener available in languages other than English? Which languages?*

The PEDS-DM is available in English; some of the forms are translated into Spanish.

How were versions in languages other than English developed?

The developers do not provide information about the development of the PEDS-DM in other languages.

What are the findings on the reliability and validity of versions of the developmental screener in languages other than English?

The reliability and validity in languages other than English have not been examined.

Accommodations for Children with Special Needs. *Are there suggested accommodations for assessing children with special needs?*

Yes, there are suggested accommodations for screening children who have identified or suspected special needs. While the PEDS-DM is usually completed by parents, when a hands-on administration is needed, guidelines are provided for establishing rapport, managing children with behavioral problems, and making accommodations for children with autism spectrum disorders as well as visual, hearing, and motor impairment.

Consultation with Diverse Populations. *Have cognitive testing or focus groups been conducted to determine whether this developmental screener is appropriate for use with diverse populations*

The developers do not provide information regarding whether the appropriateness of the PEDS-DM for diverse populations was examined in this way.

Risk Levels. *What terminology is used to describe risk levels (e.g., delay, no delay, at risk, caution, rescreen, okay, etc.)?*

The PEDS-DM screener describes milestones in each domain as "met" or "unmet." Guidance is provided on how to explain results to families using appropriate language.

Characteristics of 2006 Norming Sample
Number of children in the sample: 1,619

	Percentage of Children
Gender	
Male	51
Female	49
Ethnicity	
White	66
African American	15
Hispanic	16
Asian/other	3
Geographic Region	
West	32
South	26
Central	23
Northern	19
Site Location	
Pediatrician's Office	37
Day Care Center/Preschool	22
Child Find Program	14
Public School	27

Reliability and Validity Information

What is known about the reliability and validity of the developmental screener…

In English?

There is reliability, validity, sensitivity, and specificity information for the PEDS-DM in English. This information is outlined in response to later questions in this profile.

In other languages?

While Spanish-speaking children and children from the Spirit Spirit (Dakota) tribe who speak the native tribal language were included in the sample, the developers have not examined the reliability, validity, sensitivity, and specificity for this population.

For dual language learners?

The developers do not provide information about dual language learners and have not examined the reliability, validity, sensitivity, and specificity for this population.

For children with special needs?

The developers have examined the sensitivity and specificity of the PEDS-DM for children with special needs; however, they have not examined other aspects of reliability and validity for children without special needs.

For American Indian/Alaskan Native children?

While American Indian children and Hawaiian/Pacific Islander children were included in the sample, reliability, validity, sensitivity, and specificity have not been examined separately for these groups.

For children of migrant and seasonal farm workers?

The developers do not provide information about children of migrant and seasonal farm workers and they have not examined the reliability, validity, sensitivity, and specificity for the PEDS-DM for this population.

Reliability: Does the developmental screener obtain the same results, consistently, under the same conditions with the same children?

Interrater Reliability. *Do different raters agree when they are assessing the same children?*

Yes, different raters meet the criteria for acceptable agreement when they are screening the same children. Raters agreed between 82 percent and 96 percent of the time. Additionally, parents and professionals agreed 81 percent of the time when the screeners were directly administered to the children. Agreement between raters was examined with a sample of 77 children; however the developers do not provide further information about the children or adults involved.

Test-Retest Reliability. *How consistent are scores if the developmental screener is administered once and then administered again soon?*

There is acceptable consistency of scores when the developmental screener was administered and then administered again within one week. This was examined with a sample of 153 children from the larger group previously described.

Internal Consistency Reliability. *How strongly related are items that are intended to reflect the same set of skills or behaviors?*

There are acceptable relationships between items that are intended to reflect the same set of skills and behaviors. This was examined with all of the children in the sample described in the table.

Validity: Does the developmental screener do what it is supposed to?

Content Validity. *Do experts agree that the items in the developmental screener do a good job of reflecting what the developmental screener is supposed to be assessing?*

Items on the IED-II and CIBS-R from which the PEDS-DM was drawn were generated with the help of teachers, pediatricians, and others. Additionally, a panel of experts helped refine the unique item set for the PEDS-DM.

Construct Validity. *How closely related to each other are sets of items within the developmental screener that aim to address similar skills and behaviors, compared to sets of items that aim to address different skills and behaviors? Do scores on sets of items relate to children's age as expected?*

Because the PEDS-DM is very short, the developers have not examined this question.

Convergent and Divergent Validity. *How strongly do the scores of this developmental screener show a relationship to the scores of other developmental screeners of similar domains?*

To examine the relationships between the PEDS-DM and other developmental screeners, children were screened using the PEDS-DM and either the IED-II or CIBS-R. Overall, the results of the PEDS-DM and the two measures with which it was compared meet the criteria for strong relationships. For example, children who score highly on the IED-II or CIBS-R are likely to "pass" the PEDS-DM, which would suggest that both tools agree that the children are not at risk for delay. It should be noted that the items on the PEDS-DM are taken IED II and the CIBS-R, so there is inherent overlap between the tools.

Scores for Further Evaluation. *Are specific scores used to identify whether further evaluation is needed? How are these scores determined by the developer?*

Yes, specific scores are used to identify whether further evaluation is needed. If a child scores at or below the 16th percentile on an item, then he or she failed the item. At this level, 84 percent or more of typically developing children can complete that item.

Sensitivity. *How accurately does the developmental screener correctly identify children at risk for developmental delays?*

To determine how accurately PEDS-DM identifies children at risk for developmental delays, children's scores on PEDS: DM were compared to scores on similar domains of the IED-II and CIBS-R. Results showed that PEDS: DM meets the criteria for moderately accurate at correctly identifying children at risk for developmental problems.

Specificity. *How accurately does the developmental screener correctly identify children who are not at risk for developmental problems?*

To determine how accurately PEDS-DM identifies children not at risk for developmental delays, children's scores on PEDS: DM were compared to scores on similar domains of the IED-II and CIBS-R. The PEDS-DM meets the criteria for moderately accurate at correctly identifying children who are not at risk for developmental problems.

Follow-Up Guidance

Program Follow-Up Steps. *Does the developmental screener come with guidance about follow-up steps based on the results?*

Yes, the PEDS-DM comes with guidance and follow-up steps based on the results, including information about additional developmental screeners or assessments that can be used for further evaluation.

Family Follow-Up Steps. *Does the developmental screener include recommendations on how families might follow up on the results of the screening?*

Yes, the PEDS-DM manual includes many recommendations for families, including a resource guide, informational handouts, and parent education information.

References

Glascoe, F.P., & Robertshaw, N.S. (2007). PEDS: Developmental milestones: A tool for surveillance and screening, Professionals Manual. Ellsworth & Vandermeer Press, LLC.

Glascoe, F.P. (2007). Using Parents' Evaluation of Developmental Status (PEDS) and PEDS-Developmental Milestones (PEDS-DM): A case example. Retrieved November 17, 2009, from http://pedstest.com/dm/casestudy-1.php..

Parents' Evaluation of Developmental Status--Developmental Milestones (PEDS-DM). Retrieved December 7, 2009, from http://www.pedstest.com/dm/.

What are the components of the PEDS: DM? Retrieved December 7, 2009, from http://www.pedstest.com/dm/dm-components.php.

Appendix A: Glossary of Terms

Adaptation or Accommodation – A change in the way assessments and screeners are presented or in how the child is allowed to respond so that children with disabilities or limited English proficiency can be assessed or screened. For example, one might include Braille forms for blind children (adaptation) or allow more time for children whose primary language is not English (accommodation). This term generally refers to changes that do not substantially alter what is being measured.

Assessment – A tool used to measure skills and abilities which helps determine progress over time.

Battery – An array of similar tools intended for use together, such as "a battery of assessments" for different developmental areas.

Concurrent validity – This term describes the relationship between two separate measures of similar constructs which, when administered at the same time, provide results that are consistent with one another. Note: Sometimes manuals refer to this as convergent criterion validity, which could be interpreted to mean that the two tools concur or agree in the measurement of a particular construct.

Construct – The concept, idea, or theory that an assessment or screener is designed to measure.

Construct validity – The extent to which a tool measures a clearly defined theoretical concept. The instrument should be based on a theory, and scores from the instrument should reflect what would be expected based on that theory.

Content validity – The extent to which a tool reflects the range of possible skills or behaviors that make up the domain or construct being assessed. This is often determined through expert review.

Convergent validity – A subtype of criterion-related validity. This term indicates the degree to which a tool correlates with other tools assessing the same construct.

Correlation – A statistic that tells the strength of the relationship between different variables, items, constructs, or responses. When two measures correlate highly, one cannot necessarily be used as a substitute for the other. For example, students' reading test scores may correlate highly with their math test scores, but giving the students extra help and practice in math is not likely to improve their reading skills. Although a correlation tells how strongly two measurements tend to agree, it cannot tell why they agree. A *positive* correlation means that when one variable increases, the other increases as well, such as when language skills increase as a child gets older. A *negative* correlation means that as one variable increases, the other decreases, such as when children with more advanced language skills are less likely to show aggressive behaviors.

Criterion-referenced test – A test that is intended for comparing each child's score with one or more fixed standards of performance. The purpose of the test is not to determine each child's relative position in a group, but to determine whether each child is above or below a particular level of the knowledge or skill being measured. Scores have meaning relative to a child's own progress, rather than in relation to the scores made by some external reference (or norm) group. See also norm-referenced test.

Criterion-related validity – The degree to which the scores of one tool are related to the scores of another existing tool which measures the same construct. This other well-established tool is referred to as the criterion. The comparison between the tool and the criterion can be done either concurrently (i.e., concurrent validity), or later in time (i.e., predictive validity).

Cutoff scores – Minimum scores used to decide whether further evaluation is needed, usually differentiated by age in months and years. A score at or below the cutoff score indicates that the child needs to be referred for further testing. A child's score above the cutoff indicates that the child has demonstrated mastery of the skills and abilities in that domain for his/her age.

Developmental delay – A delay in the appearance of some steps or phases of growth and development. NOTE: Programs serving at-risk populations may expect to find higher rates of children being identified as at risk for developmental delay than typically found when looking at the total population of both at-risk and not-at-risk children.

Developmental norms – Standards by which the progress of a child's development can be measured relative to the development of a representative cross section of children, i.e. the norm. For example, the average age at which a child walks, learns to talk, or achieves toileting independence would be a standard used to judge whether the child is progressing normally. While norms are usually thought of as age-related, norms can also be tied to other developmental variables such as race, ethnicity, and gender. Norms can inform teachers, parents, and others in judging the appropriateness of certain types of activities for different children.

Direct assessment – An assessment that is administered by an assessor and requires a child to perform tasks in response to a set of standardized, direct instructions.

Discriminant or divergent validity – A subtype of criterion-related validity that indicates the degree to which the tool is less closely related to measures of theoretically different constructs.

Domain – A set of related skills, behaviors, or information that is classified as a single area of study or development. Domains typically cover multiple, related constructs within a broad area of study or development, such as fine motor development or approaches toward learning.

Factor analysis – A procedure used to examine the relationships among items or questions to see whether the items group together, or are distinct, in expected ways. Researchers sometimes describe this as how well items being measured "hang together."

Faithful administration – Individuals demonstrate consistency in the skill and accuracy with which they administer an assessment or screening tool to children. Such accuracy is verified through regular checks on faithful administration, using training materials or guidance from the developer of that tool.

Indicators – Questions included in the tool that are related to the developmental skill or ability being measured.

Internal consistency reliability – How closely items or indicators within a construct are interrelated.

Interrater reliability – How similar the results of an assessment are when different individuals administer the same assessment with the same child.

Learning objective – A short but clear statement of an expectation of a skill a child can achieve, and the behavior that will be observed in order to determine whether this expectation is met. For assessment to be effective, objectives must be clearly articulated before deciding upon methods and measures.

Norm-referenced test – A tool in which a child's score is compared with the scores of a norming group, which is a representative cross-section of all those being assessed or screened. Scores from all children who take the norm-referenced test in the future are compared to the performance of the initial norm group. See also criterion-referenced test. NOTE: Depending upon the population served, a program may or may not expect scores from its assessments to

reflect those in the norming sample. For instance, programs serving more at-risk populations may expect to see a larger proportion of children falling below the norm.

Ongoing observational tool – A tool used at multiple intervals during a program year to observe and record a child's skills and abilities.

Population – The total number of all possible subjects or elements which could be included in a study. If the data are valid, the results of research on a sample of individuals drawn from a much larger population can then be generalized to the population.

Portfolio assessment – A collection of work, usually drawn from students' classroom work. A portfolio becomes a portfolio assessment when (1) the assessment purpose is defined; (2) criteria or methods are made clear for determining what is put into the portfolio, by whom, and when; and (3) criteria for assessing either the collection or individual pieces of work are identified and used to make judgments about performance.

Predictive validity – The degree to which items relate to an established measurement tool that assesses the same, or a related, outcome at a later point in time.

Psychometrics – The science concerned with evaluating the attributes of tests used to measure various skills and abilities. Three of these attributes of particular interest include (1) the type of data (scores) generated by the application of such tests, (2) the reliability of data from such tests, and (3) issues concerning the validity of data obtained from such tests.

Reliability – A term which describes whether a tool produces consistent information across different circumstances. Scores will be stable regardless of *when* the tool is administered, *where* it is administered, and *who* is administering it. Therefore, reliability is an indication of the consistency of scores across raters, over time, or across different tasks or items that measure the same thing. An unreliable assessment or screener cannot be valid.

Sample – A subset of a population. Samples are collected and statistics are calculated from the samples so that one can draw conclusions about the total population. A *representative sample* refers to a carefully chosen number of representatives of a specific group, such as children of a certain age, race/ethnicity, or income status, whose characteristics represent as accurately as possible the entire population of children with these characteristics.

Screener – A tool used to evaluate whether a child may be at risk for a developmental delay.

Sensitivity – A term which describes the degree to which children who are at risk for developmental delay are accurately identified as needing further evaluation by a screening tool.

Specificity – A term which describes the degree to which children who are not at risk for developmental delay are accurately identified as typically developing by a screening tool.

Subscales – A set of items within a domain that capture a particular aspect of the domain. For example, the domain of language development might have the following subscales: receptive communication, expressive communication, and alphabet knowledge.

Test-retest reliability – An indicator of whether the tool will yield the same score across two administrations of the tool within a short period of time. This tells us whether the tool provides a consistent assessment of a skill, regardless of other factors, such as the child's mood or health, the time of day, or the time of year that the child was assessed. A child should score similarly (within a defined range) if tested within a short period of time, usually defined as within three months.

Typically developing – Children who pass a set of predictable milestones at expected times as they grow and develop.

Validity – A term which describes whether a tool assesses what it is supposed to assess and indicates that scores are accurately capturing what the tool is meant to measure in terms of content. For example, if a child performs well on a vocabulary test, a valid measure would mean there is confidence that the child is good at word comprehension. An assessment or screener cannot be valid if it is not reliable.

Variable – A quality, characteristic, or attribute that may change depending on the sample being studied. For example, commonly used variables include age, gender, race/ethnicity, poverty status, or levels of education.

References

Corekin, K. (2009). Anecdotal notes in early childhood education: Use objective record keeping to evaluate children's development. Retrieved from http://preschool.suite101.com/article.cfm/anecdotal_notes_in_early_childhood_education

Correlation. (2004,). In North Western University's Holmgren Lab glossary. Retrieved from http://www.biochem.northwestern.edu/holmgren/Glossary/Definitions/Def-C/correlation.html

Educational Testing Service. (2004). *Definitions of statistical terms relating to tests*. Retrieved from http://www.ets.org/Media/Log_On_Lets_Talk/stat_terms_english.pdf

Educational Testing Service (n.d.). *Glossary of assessment terms*. Retrieved from http://www.ets.org/portal/site/ets/menuitem.c988ba0e5dd572bada20bc47c3921509/?vgnextoid=3f1caf5e44df4010VgnVCM10000022f95190RCRD&vgnextchannel=8b0d253b164f4010VgnVCM10000022f95190RCRD#NormReferenced

Furr, R.M., & Bachrach, V.R. (2008). Psychometrics: An introduction. Thousand Oaks, CA: Sage Publications.

Leach, P. (1997). *Your baby and child: From birth to age five*. New York: Knopf.

Maryland State Department of Education. (2010). *Assessment literacy glossary*. Retrieved from http://mdk12.org/process/cfip/Assessment_Literacy_Glossary.html

National Center for Research on Evaluation, Standards, and Student Testing (CRESST) at UCLA. (2008). *Products and resources: Glossary*. Retrieved from http://www.cse.ucla.edu/products/glossary_2.html/

National Research Council of the National Academies (2008). *Standards for educational and psychological testing*. Washington, DC: National Academy Press.

Riley Child Development Center. (2010). *Terminology*. Retrieved from http://www.child-dev.com/page.cfm?id=23

Schaefer, C.E. & Digeronimo, T.F. (2000). *Ages and stages: A parent's guide to normal childhood development*. New York: Wiley.

The National Assessment of Education Progress (NAEP). (2010). *Processing assessment materials*. Retrieved from http://nces.ed.gov/nationsreportcard/tdw/process_materials/

University of St. Andrews. (n.d.). Statistics: Glossary. Retrieved from http://psy.st-andrews.ac.uk/resources/glossary.shtml

White, S. (n.d.) *Pre-K assessment*. Retrieved from http://www.preknow.org/policy/assessment.cfm

Appendix B: Psychometric Documentation and Rationale

In order to describe reliability and validity in these profiles, the information presented in each technical manual was analyzed against a range of values, or cutpoints, that represent varying levels of evidence for each type of reliability and validity. For each type of reliability and validity, statistical indicators representing the strength of the relationship between two variables or items were examined. These scores can range from 0 to 1. A set of criteria or cutpoints were established for each type of reliability and validity. Wherever possible, these criteria were based on generally accepted standards in the field. Where there is no generally accepted standard in the field, the cutpoints were established by consulting research literature on early childhood assessment, statistical texts related to measurement development, criteria used in the *Resources for Measuring Services and Outcomes in Head Start Programs Serving Infants and Toddlers* (published by the U.S. Department of Health and Human Services), and recommendations made by professional organizations such as the American Academy of Pediatrics. The criterion and terminology used in the profiles to describe each type of reliability and validity are outlined in the table below. Please refer to the Glossary in Appendix A for more details about each type of reliability and validity.

Type of Reliability or Validity	**Description and Source of Evidence Used to Establish Criteria**	**Criterion and Terminology Used**
Construct Validity	Measured by examining associations between subscales within the assessment or screener instrument. Also measured by examining associations between subscale scores and child characteristics, such as age. No established standard in the field	0.50 or higher=strong/high 0.30 – 0.49=moderate 0.29 or below=weak/low
Content Validity	Measured by whether tool was reviewed by experts to determine if content reflects what the assessment or developmental screener is supposed to be measuring	Content was or was not reviewed by experts
Convergent/Concurrent Validity	Measured by correlating the scores of the assessment/ developmental screener with scores on other assessments/ developmental screeners of similar content to determine the strength of relationships between the two Source: Administration for Children and Families (2003)	0.50 or higher=strong/high 0.30 – 0.49=moderate 0.29 or below=weak/low

Type of Reliability or Validity	Description and Source of Evidence Used to Establish Criteria	Criterion and Terminology Used
Predictive Validity	Measured by correlating the scores of the assessment/ developmental screener with scores on later assessments to determine how well the assessment/developmental screener predicts to later achievement or development Source: Administration for Children and Families (2003)	0.40 or higher=provides evidence that measure may predict to later achievement 0.39 or below=does not provide evidence that measure predicts to later achievement
Sensitivity	Measured by how often the developmental screener correctly identifies children at risk for developmental delays Source: Council on Children with Disabilities (2006)	0.90 or higher=high 0.70 – 0.89=moderate 0.69 or below=low
Specificity	Measured by how often the developmental screener correctly identifies children not at risk for developmental delays Source: Council on Children with Disabilities (2006)	0.90 or higher=high 0.70 – 0.89=moderate 0.69 or below=low
Internal Consistency Reliability	Measured by correlating items within a construct to determine the interrelatedness of the items No established standard in the field	0.70 or higher=acceptable 0.69 or below=low/weak
Interrater Reliability	Measured by the level of agreement between two raters when assessing the same children No established standard in the field	0.80 or higher=acceptable 0.79 or below=low/weak

Type of Reliability or Validity	Description and Source of Evidence Used to Establish Criteria	Criterion and Terminology Used
Test-Retest Reliability	Measured by correlating the scores on two administrations of the same assessment/ developmental screener given to the same child within a short period of time to determine consistency No established standard in the field	0.70 or higher=acceptable (across a period of three months or less) 0.69 or below=low/weak

Sources Consulted in Determining Cutpoints

Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.

Brown, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models*. Newbury Park, CA: Sage.

Carmines, E.G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Sage University Series on Quantitative Applications in the Social Sciences, 07-001. Beverly Hills, CA, and London: Sage.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum.

Council on Children with Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee, Medical Home Initiatives for Child with Special Needs Project Advisory Committee. (2006). Indentifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics 118*(1), 405-420.

Lewis-Beck, M.S. (1995). *Data analysis: An introduction*. Thousand Oaks, CA: Sage.

Robinson, J.P., & Shaver, P.R., Wrightsman, S. (1991). *Measures of personality and social psychological attitudes.* San Diego, CA: Academic Press.

Snow, C.E., & Van Hemel, S.B. (Eds.). (2008). *Early childhood assessment: What, why, and how*. Washington, DC: National Research Council of the National Academies.

Administration for Children and Families (2003). *Resources for measuring services and outcomes in Head Start Programs Serving Infants and Toddlers.* E. Kisker, K. Boller, C. Nagatashi, C. Sciarrino, V. Jethwani, T. Zavitsky, M. Ford, J. Love, &Mathematica Policy Research, Inc. Washington, DC: U.S. Department of Health and Human Services. Retrieved February 15, 2009 from http://www.acf.hhs.gov/programs/opre/ehs/perf_measures/reports/resources_measuring/res_meas_cdi.html