

# **Achieving Student Progress With Scientifically Based Formative Assessment**

**A white paper from Pearson Education**



**June 2005**

**Table of Contents**

<b>Introduction</b>	<b>3</b>
<b>The Basis of PASeries (Progress Assessment Series)</b>	<b>4</b>
<b>Scientifically Designed Test Items</b>	<b>5</b>
<b>Secure, Accurate Administration</b>	<b>8</b>
<b>Timely Scoring and Reporting</b>	<b>8</b>
<b>Conclusion</b>	<b>10</b>
<b>About Pearson Educational Measurement</b>	<b>11</b>
<b>About Pearson Education</b>	<b>11</b>
<b>Endnotes</b>	<b>12</b>

## Introduction

With the acceleration of education reform over the past several years, schools, districts, and states have been deluged with myriad tools and resources purporting to:

- Guide teachers in improving instruction
- Aid administrators in demonstrating annual yearly progress and increased student achievement
- Allow state officials to meet reporting requirements

These materials range from traditional black line masters to complex student management systems. One area, however, is gaining increased focus and attention: formative assessment. Formative assessments – tests that are given throughout the year – are a necessary tool in the current educational environment. Formative assessments provide educators with insight into student performance, inform instruction, and provide administrators with valuable reporting data.

However, not all formative assessment products are equal. While many claim to help meet education reform requirements, few are able to actually do so as they lack a solid scientific basis.<sup>1</sup> Assessments without this foundation may be poorly formed, provide incomplete or flawed performance data, and produce suspect reports.

For educators and administrators to receive the highest quality data from assessments, the assessments must be scientifically based. Only then can educators and administrators gain an accurate picture of student achievement with actionable results.

## The Basis of PASeries™ (Progress Assessment Series™)

To meet the need for valid, scientifically based formative assessment, Pearson Education developed PASeries™ (Progress Assessment Series). Designed to provide educators with a reliable measure of student progress, PASeries was built from the ground up to address the challenges that education professionals encounter.

PASeries was developed by experts from throughout Pearson Education, ranging from editorial content experts to industry-leading psychometricians. This collaboration allowed the PASeries team to develop the highest caliber test items. These items were based on the same principles and standards that guide high-stakes, large-scale assessment development to ensure scientific validity and reliability.

One challenge for educators is matching the results of a formative assessment to the summative state test. Standardized measures of performance in core instructional areas help teachers to complete this task more easily and effectively. The first two products in the PASeries product line, PASeries Reading and PASeries Mathematics, both use standardized measures of performance to evaluate student progress: the widely adopted Lexile<sup>®2</sup> scale for reading measurement and its companion Quantile<sup>®3</sup> scale for mathematics.

The Lexile scale, part of the Lexile Framework for Reading and used in PASeries Reading, is a developmental scale for reading ability ranging from 200L for beginning readers to above 1700L for advanced readers. Nearly 20 million students receive a Lexile measure from a state or school district test each year, making it the most widely adopted reading measure in use today.

Similarly, results for PASeries Mathematics are calculated using The Quantile Framework for Mathematics. The Quantile scale is a numeric scale of mathematics achievement (with scores followed by the letter "Q," as Lexile measures are followed by the letter "L") ranging from below 200Q to 1700Q. The Quantile scale assesses three competencies of mathematics: conceptual competence, procedural competence, and utilization competence.

PASeries is the first formative assessment product line designed to forecast student growth toward state performance standards. Using the Lexile and Quantile scales, PASeries allows educators to use formative assessments to monitor student progress toward state assessment goals, adjust instruction to accommodate learning needs, and view projected student achievement on state summative assessments.

## Scientifically Designed Test Items

PASeries Reading and PASeries Mathematics were developed by measurement professionals using the standards for high-stakes test development described in the 1999 guidelines, “Standards for Educational and Psychological Assessments.” PASeries test items were developed based on an extensive review of national and state standards, national and state tests, and curriculum and instructional practices.

Pearson Education leveraged the strength of its Pearson Educational Measurement (PEM) unit for the development of the PASeries formative assessments. PEM is recognized as an industry leader in academic assessment and is the largest comprehensive provider of academic assessments.<sup>4</sup> With more than 50 years of experience in academic testing, PEM had the expertise to develop a formative assessment that would meet today’s rigorous requirements for scientific validity and reliability.

Content for the test items for PASeries Reading and PASeries Mathematics was carefully developed for age and grade level appropriateness. Test items underwent review by content area and testing and assessment experts. Reviews included curricular review, assessment review, grammar review, sensitivity review, universal design principles review, and fact-checking review.

Test items were then subjected to field tests spanning a variety of socio-economic groups, again in keeping with the same rigor as would apply to the development of statewide assessment items. More than 40,000 students from districts such as Miami-Dade County School District, Paterson (N.J.) Public Schools, Chicago Public Schools, Long Beach Unified School District, and Houston Independent School District participated in 2003-2004 school year field tests.

Professional item development, rigorous reviews, and field testing mean few, if any, products in the market can claim a similar level of scientific validity and reliability as PASeries.

## Mathematics and Reading Content Validity

**Validity:** The validity of a test is the degree to which the test actually measures what it purports to measure. Validity provides a direct check on how well the test fulfills its function. “The process of ascribing meaning to scores produced by a measurement procedure is generally recognized as the most important task in developing an educational or psychological measure, be it an achievement test, interest inventory, or personality scale.”<sup>5</sup>

The content validity of a test relates to how important content has been acquired and how well it is used in the test. Content validity was built into the PASeries Mathematics assessments during the development process. All items are aligned with state (California, Florida, North Carolina, and Texas) and national (National Council of Teachers of Mathematics and the National Assessment of Educational Progress) mathematical content standards. Educators and mathematics specialists wrote and reviewed the test items.

As with PASeries Mathematics, content validity was built into the PASeries Reading assessments during the development process. All texts are authentic and developmentally appropriate, and the student is asked to respond to the text in ways that are appropriate for the genre (for example, with nonfiction texts, the student is asked specific questions related to the content rather than asked to make inferences about what will happen in the text). Again, educators selected and reviewed the texts and wrote and reviewed the questions.

PASeries Reading is designed to measure how well readers read literature and expository texts. PASeries Reading measures reading ability by focusing on skills readers use when studying written materials sampled from various content areas. These skills include referring to details in the passage, drawing conclusions, and making comparisons and generalizations. PASeries Reading does not require prior knowledge of ideas outside of the passage, vocabulary taken out of context, or formal logic. PASeries Reading is built from authentic passages that are typical of the materials students read both in and out of school, including topics in prose fiction, the humanities, social studies, science, and everyday texts such as magazines and newspapers.

The passage native-Lexile item format used on all PASeries Reading assessments is an extension of the “embedded completion” item format that has been shown to measure the same core reading competency that is measured by norm-referenced, criterion-referenced, and individually administered reading tests.<sup>6</sup>

More than a decade of research went into defining the rules for sampling text and writing passage native-Lexile and embedded completion items. These rules were precisely followed in developing the PASeries Reading items. A multi-stage review process was used to ensure conformance with the item-writing specifications.

## Mathematics and Reading Content Reliability

**Reliability:** If use is to be made of some piece of information, then the information should be reliable – stable, consistent, and dependable. In reality, all test scores have some error (or level of uncertainty). This uncertainty in the measurement process is related to three factors: (1) the statistical model that was used to compute the score, (2) the questions that were used to determine the score, and (3) the condition of the reader when the questions used to determine the score were collected. Once the level of uncertainty in a test score is known, it can be taken into account when using the test results. Reliability, or the consistency of scores obtained from an assessment, is a major consideration in evaluating any assessment procedure.<sup>7</sup>

Reliability, or the consistency of scores obtained from an assessment, is a major consideration in evaluating any assessment. Uncertainty has been examined with PASeries Mathematics in terms of examinee error.

PASeries Mathematics was developed using the Rasch one-parameter item response theory model to relate a student's ability to the difficulty of the items. A unique amount of measurement error is attributable to model misspecification (violation of model assumptions) associated with each score on the assessment. The computer algorithm that controls the administration of the assessment uses a Bayesian procedure to estimate each student's ability. This procedure uses prior information about students to control the selection of questions and the recalculation of each student's reading ability after responding to each question.

Two sources of uncertainty have been examined with PASeries Reading: text error and reader error. To determine a Lexile measure for a text, the standard procedure is to process the entire text. All pages in the work are concatenated into an electronic file that is processed by a software program called The Lexile Analyzer<sup>®</sup>, developed by MetaMetrics Inc. The analyzer "slices" the text file into as many 125-word passages as possible, analyzes the set of slices, and then calibrates each slice in terms of the logit metric. That set of calibrations is then processed to determine the Lexile measure corresponding to a 75-percent comprehension rate. The analyzer uses the slice calibrations as test item calibrations and then solves for the measure corresponding to a raw score of 75 percent (e.g., 30 out of 40 correct, as if the slices were test items). Obviously, the measure corresponding to a raw score of 75 percent on *Goodnight Moon* (300L) slices would be lower than the measure corresponding to a comparable raw score on "USA Today" (1200L) slices.

## Secure, Accurate Administration

PASeries tests are designed for completion in one class period and for delivery either on paper or online. Skilled psychometricians conducted thorough comparability studies to ensure that PASeries items delivered online are comparable to the printed form.

First, screener tests are administered to determine students' initial ability or achievement levels. Screener tests are designed at the difficulty level of the reading or mathematics content at the end of the previous school year and cover content that was introduced in the prior grade.

Throughout the school year, PASeries progress monitoring tests are administered. With each progress assessment, teachers can administer the most appropriate test for individual students. PASeries Reading includes three progress monitoring tests for each grade, at three different levels of difficulty: on level, above level, and below level. PASeries Mathematics includes six progress assessments for each grade designed to be administered approximately every six weeks.

PASeries assessments are delivered online through TestNav, a proprietary online test delivery system used to deliver high-stakes summative tests in many states. TestNav is a secure delivery system that can "lock down" the student desktop to prevent the use of outside resources. A full-featured system, TestNav contains a suite of tools and resources such as a calculator, ruler, and protractor, for use by students as appropriate within an assessment.

## Timely Scoring and Reporting

PASeries assessments are easily scored with results available to teachers in a timely manner. If administered online, results are quickly reported to schools, districts, and states electronically. If the tests are administered using paper and pencil, the assessments are accurately and efficiently scored using equipment such as any Pearson OpScan® scanner. With throughput of 1,500 to 8,000 forms per hour, these scanners also offer timely scoring of PASeries tests.

PASeries results provide schools and districts with growth forecasting and progress monitoring in reading and math. The formative assessment tool provides a reliable indication of each student's future performance on statewide outcome assessments, allowing educators to identify students at risk of failing and then develop appropriate interventions.

PASeries is the only formative assessment that is able to reliably forecast student achievement on these statewide assessments. PASeries utilizes a proprietary algorithm that takes into account past performance, time between tests, and a host of other factors to determine potential performance.

In addition to individual student progress monitoring and forecasting reports, a variety of reports is available at the class, school, district, and state levels. PASeries provides a number of disaggregate reports as well to allow for subgroup reporting to help meet educational reform requirements. PASeries reports are available to educators on the Web immediately or, for more complex reports, within 24 hours of their request.

As mentioned previously, PASeries assessments use the Lexile and Quantile scales as standard reporting measures. PASeries Mathematics is the first product to implement the Quantile scale. These scales are linked to an individual state scale to provide an accurate representation of a student's estimated performance on a statewide summative assessment and, as longitudinal scales, also measure a student's progress across grades, schools, and districts.

Lexiles and Quantiles are also used for instructional purposes outside of test scoring. For example, more than 100,000 trade and textbooks have Lexile measures of text difficulty (a database is available at [www.Lexile.com](http://www.Lexile.com)). A student's Lexile measure of reading ability, when matched with the Lexile measure of a text, generally means a 75-percent expected comprehension rate – difficult enough to be challenging and promote reading progress but not so difficult that it is frustrating. This helps educators, parents, and students choose books and periodical articles for reading both in and out of the classroom, as well as helps students understand curriculum topics at an appropriate level of comprehension.

## Conclusion

Pearson Education developed PASeries, the first formative assessment product designed to forecast student growth toward state performance standards, to help schools and districts meet the increasing demands for high levels of achievement on statewide tests. Test items in PASeries assessments were developed by experts in testing and assessment and rigorously field-tested to ensure the highest quality.

The first two products in the PASeries product line, PASeries Reading and PASeries Mathematics, report in standard measures of performance – Lexiles for reading and Quantiles for mathematics – to ensure accurate and easy mapping to state assessment goals. Using the valuable student performance data reported from PASeries assessments, schools can monitor student progress and meet the ultimate goal of testing: improving student learning.

PASeries formative assessments represent a high level of quality. The product line leads the industry in scientific validity and reliability. Additionally, PASeries assessments are currently the only formative assessment that provides a scientifically based forecast of student achievement toward state performance goals.

By employing reliable, valid assessment instruments, educators can ensure that the quality of student data they receive is of the highest caliber. Only when reliable data are available can informed instructional decisions be made and action taken. Using PASeries formative assessments, educators have access to the data they need to inform instruction, assess intervention strategies, and help comply with state and national reporting requirements.

For more information on PASeries, visit [www.PearsonPASeries.com](http://www.PearsonPASeries.com).<sup>9</sup>

## **About Pearson Educational Measurement**

Pearson Educational Measurement (PEM), the developer of PASeries, is the largest comprehensive provider of educational assessment products, services, and solutions, helping states and large school districts meet the requirements of education reform while using testing and assessment to promote learning. As a pioneer, PEM has been a trusted partner in district, state, and national assessments for more than 50 years. PEM's full-service offerings for K-12 and other assessment organizations include Pearson Educational Measurement Solutions for custom assessments, both online and on paper; PASeries for formative assessments; Perspective for performance reporting; and a wide range of data warehousing and other educational assessment products and services. Pearson Educational Measurement operates as a business of Pearson Education, the world's largest education company.

## **About Pearson Education**

Educating 100 million people worldwide, Pearson Education is the global leader in educational publishing. With brands such as Pearson Prentice Hall, Pearson Scott Foresman, Pearson Educational Measurement, Pearson NCS, Pearson Learning Group, LessonLab, and many others, Pearson Education provides quality content, assessment tools, and educational services in all available media, spanning the learning curve from birth through college and beyond. Pearson Education is part of Pearson, the international media company, whose other primary operations include the Financial Times Group and the Penguin Group.

## Endnotes

- <sup>1</sup> U.S. Department of Education, *Proven Methods: Scientifically Based Research*. Washington, D.C.: U.S. Department of Education Web site, Feb. 6, 2002, [www.ed.gov/nclb/methods/whatworks/research/index.html](http://www.ed.gov/nclb/methods/whatworks/research/index.html)
- <sup>2</sup> The Lexile Framework for Reading Web site, [www.Lexile.com](http://www.Lexile.com)
- <sup>3</sup> The Quantile Framework for Mathematics Web site, [www.Quantiles.com](http://www.Quantiles.com)
- <sup>4</sup> Simba Information, *Academic Testing Materials: Market Analysis & Forecast 2004-2005*, (2004).
- <sup>5</sup> Stenner, A.J., Smith M., & Burdick, D.S. (1983). Toward A Theory of Construction Definition. *Journal of Educational Measurement*, 20(4), 305-315.
- <sup>6</sup> Stenner, A.J., Smith, D.R., Horiban I., & Smith, M. (1987a). *Fit of Lexile Theory to Item Difficulties on Fourteen Standardized Reading Comprehension Tests*. Durham, NC: MetaMetrics Inc.
- <sup>7</sup> Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of Item Response Theory*. Newbery Park, CA: Sage Publications, Inc.
- <sup>8</sup> Pearson NCS Web site, [www.pearsonncls.com/scanners/](http://www.pearsonncls.com/scanners/)
- <sup>9</sup> Pearson PASeries Web site, [www.PearsonPASeries.com](http://www.PearsonPASeries.com)

PASeries and Progress Assessment Series are trademarks or U.S. registered trademarks of Pearson Education, Inc., and/or one or more of its direct or indirect affiliates. The names of other companies and products mentioned herein may be the trademarks of their respective owners. Copyright © 2005 Pearson Education, Inc., and/or one or more of its direct or indirect affiliates.