

A

Accountability

The use of assessment data and other information to assure participants in the educational system that schools are moving in desired directions. Elements commonly included are goals, indicators of progress toward meeting those goals, analysis of data, reporting procedures and consequences or sanctions. Accountability often includes the use of assessment results and other data to determine program effectiveness and to make decisions about resources, rewards and consequences.

Alignment

The similarity or match between and among content standards, performance standards, curriculum, instruction, and assessment in terms of knowledge and skill expectations. The inferences made on the basis of assessment results are valid only to the extent that the system components are aligned.

C

Content standards

Statements of knowledge and skills that schools are expected to teach and students are expected to learn. Content standards indicate what students should know and be able to do as a function of schooling.

Criterion-referenced

A kind of assessment in which results are interpreted in relation to the specific knowledge or skills possessed by a student. Such tests usually cover relatively small units and are related to instruction. Performance is measured in reference to the mastery of particular skills. Scores from these tests have meaning in terms of what the student knows or can do, rather than in relation to the scores made by some external reference (or norm) group.

D

Differentiated instruction

The attempt "on the part of classroom teachers to meet students where they are in the learning process and move them along as quickly and as far as possible in the context of a mixed-ability classroom" (Tomlinson, 1999). Differentiated instruction promotes high-level and powerful curriculum for all students, but varies the level of teacher support, task complexity, pacing, and avenues to learning based on student readiness, interest and learning profile.

F

Forecasted comprehension rate

The Lexile Framework for Reading is based on a conjoint measurement model such that a reader and a text are placed on the same scale, the Lexile scale. When the reader measure and the text calibration are the same (difference of 0L) then the forecasted comprehension rate is 75 percent. When the difference between the reader measure and the text calibration is -250L (the text is more

difficult than the reader is able), then forecasted comprehension rate falls to 50 percent. Conversely, when the difference between the reader measure and the text calibration is +250L (the reader is more able than the text is difficult), then forecasted comprehension rate improves to 90 percent.

G

Grade equivalent

A score that represents the typical (mean or median) performance of students tested in a given month of the school year. It is a decimal number that shows performance in terms of grade level (to the left of the decimal) and months (to the right of the decimal).

L

Lexile Measure

A unit of measurement used when determining the difficulty of text and the reading level of readers. A Lexile is equivalent to 1/1000th of the difference between the comprehensibility of basal primers (the midpoint of first-grade text) and the comprehensibility of an electronic encyclopedia (the midpoint of workplace text).

Lexile Framework for Reading

The Lexile Framework for Reading is a system that can help determine the reading level of any written material—from a book to a test item. The Lexile Framework can also be used to assess a readers' reading comprehension level. After test results are converted into Lexile measures, readers can be matched to reading materials on their own level and comprehension rates of readers can be forecasted to determine how well a reader will comprehend thousands of texts that have been measured in the Lexile metric.

Lexile map

The Lexile map provides a common sense context for understanding Lexile as well as reading test results. It is representative of the Lexile Book Database. The map contains representative books and educational levels at all points on the Lexile scale. The educational levels provide a classroom context for Lexile measures. The levels are expressed as ranges because within every classroom there are a range of readers and a range of reading materials. For example, in a fifth-grade classroom students will likely be reading from the third-grade level to the eighth-grade level. Conversely, in a fifth-grade classroom there will be reading materials ranging from the third-grade level to the eighth-grade level.

Lexile range

The range of Lexile text measures that a particular reader is forecast to comprehend while still encountering enough challenge to grow as a reader. The Lexile range for a reader is from 50L above her or his Lexile measure to 100L below. The forecasted comprehension rate for the Lexile range is about 65% to 80%. When a reader reads within his or her Lexile range, this is referred to as a targeted reading experience. When a reader reads outside of his or her Lexile range, this is referred to as a mistargeted reading experience.

Lexile reader measure

A Lexile reader measure is the specific number that describes a student's reading comprehension ability. A student receives his or her Lexile measure through formal methods such as a linking study where the reporting scale of a norm-referenced or criterion-referenced assessment is linked with the Lexile scale, or through informal methods such as listening to a student read a book with a known Lexile measure.

Lexile scale

The Lexile scale ranges from 200L to 1600L, although actual Lexile measures can range from below 0L to above 2000L.

Lexile text measure

A Lexile text measure is the specific number assigned to any text indicating its reading demand in terms of its semantic difficulty (vocabulary) and syntactic complexity (sentence length). A computer program called the Lexile Analyzer computes the measure. The Lexile Analyzer carefully examines the whole text to measure two characteristics: sentence length and word frequency. Research has proven that these characteristics are highly related to overall reading comprehension. The Lexile Analyzer then reports the Lexile measure of the text.

N

Normal curve equivalent (NCE)

Normal curve equivalents allow comparison between different tests for the same student or group of students and between different students on the same test. An NCE is a normalized student score with a mean of 50 and a standard deviation of 21.06. NCEs range from 1 to 99. NCEs have many of the same characteristics as percentile ranks, but have the additional advantage of being based on an interval scale. That is, the difference between two consecutive scores on the scale has the same meaning throughout the scale. NCEs are often required by many categorical funding agencies (for example, Title I).

Norm-referenced

A test where the results are interpreted in relation to the performance of a comparison group. Norm-referenced interpretations tell how the scores of each student or group of students compare to the scores of the original (norm) group that took the test. The scores of the students do not necessarily produce the same distribution of scores as the scores of the norm group.

P

Percentile rank

The percentile rank of a score indicates the percentage of scores less than or equal to that score. Percentile ranks range from 1 to 99. For example, if a student scores at the 65th percentile rank, it means that he or she performed as well as or better on the assessment than 65 percent of the norm group. Real differences in performance are greater at the ends of the percentile range than in the middle. Percentile ranks of scores can be compared across two or more distributions; percentile ranks cannot be used to determine differences in relative rank due to the fact that the interval

between adjacent percentile ranks do not necessarily represent equal raw score intervals. Also, the percentile rank does not refer to the percentage of items answered correctly.

Performance standards

What students must do to demonstrate various levels of proficiency with respect to the specific content. Educators and parents want to know more than just how a student's performance compares with that of other students, asking "What level of performance does a score represent?" and "How good is good enough?" Performance standards consist of four components: (1) performance levels which provide descriptive labels for student performance, e.g., "advanced," "proficient;" (2) descriptions of what students at each performance level must demonstrate relative to the test; (3) examples of student work that illustrate the range of performance for each performance level; and (4) cut scores which separate one level of performance from another. For any test scale a number of points can be identified that correspond to a specified level of performance. For example, on the National Assessment of Educational Progress (NAEP) three levels of performance are identified: basic, proficient and advanced. Proficient achievement on the NAEP reading assessment is defined as "solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations and analytical skills appropriate to the subject matter" (NAGB, 1999).

R

Reading ability

The reading comprehension skills of a reader

Readability

The difficulty level (reading demand) of a text

S

Semantic difficulty

The difficulty of the words in a text. Most operationalizations of semantic difficulty are proxies for the probability that an individual will encounter a word in a familiar context and thus be able to infer its meaning. The Lexile Framework uses a nearly 600-million-word corpus when examining the semantic component of text. This corpus was assembled from more than 44,000 texts measured by MetaMetrics for publishers from 1998 through May 2003.

Stanine

A stanine is a standardized student score with a mean of 5 and a standard deviation of 2. Stanines range from 1 to 9. In general, stanines of 1 to 3 are considered below average, stanines of 4 to 6 are considered average and stanines of 7 to 9 are considered above average. A difference of 2 between the stanines for two measures indicates that the two measures are significantly different. Stanines, like percentiles, indicate a student's relative standing in a norm group.

Syntactic complexity

The complexity of the sentence structure in a text. The Lexile Framework uses sentence length as a proxy for the syntactic component of reading demand. The research of various individuals has provided evidence that sentence length is a good proxy for the demand that syntactic complexity places upon verbal short-term memory.

Standard error of measurement (SEM)

The SEM is an index of test uncertainty. It allows us to estimate the amount of each type of error associated with "true" scores. We can compute standard errors of measurement for raters, occasions and item samples. Statistically, the SEM is the standard deviation of the error scores of a test.

T

Targeted reading experience

An interaction between a reader and a text where the reader is challenged by new vocabulary and varying sentence structures, but is not overwhelmed. In the Lexile Framework, this means that a reader is reading in his or her Lexile range. A 1000L reader reading 1000L text (75-percent comprehension) reports confidence, competence and control over the text. Teachers listening to such a reader report that the reader can sustain the meaning thread of the text and can read with motivation and appropriate emotion and emphasis. In short, such readers appear to comprehend what they are reading. This value, 75 percent, was selected to ensure that texts selected are not so hard that the reader experiences frustration and loses the meaning thread of the text, but, at the same time, are not so easy that the reader does not experience any challenge. When the measure for a text is 250L above the reader's measure, comprehension drops to 50 percent and the reader experiences frustration and inadequacy. Conversely, when the measure of a text is 250L below a reader's measure, comprehension goes to 90 percent and the reader experiences total control and automaticity. When a reader reads outside of his or her Lexile range, this is referred to as a mistargeted reading experience.

U

Uncertainty

All assessments, national norm-referenced tests, state-level criterion-referenced tests, interest inventories and diagnostic surveys have some inherent measurement error, or uncertainty. The source of the measurement error may be the test, the reader, or the interaction between the test and the reader. Every test has some inherent measurement error related to the how the test items are developed and calibrated, and the number of questions that are asked. In addition, the reader always provides some level of measurement error, e.g. prior knowledge, health and motivation. It is expected that some students may not score consistently due to other sources of measurement error internal to the student. It is advisable to go back and review the testing experiences of these students to better understand what is happening. As with any test, teacher judgment as to the

validity of the testing session and the results should be reviewed. There are days when students are not in the mood to take a test and, therefore, do not take the test.