

Glossary

Assessment, Accountability and Data-informed Decisionmaking Terms

Compiled by the Data-informed Decisionmaking
Wallace Foundation Leadership Interest Group

Introduction

This glossary has been created to serve the needs of the Wallace Foundation’s Leadership Issue Group (LIG) on Data-informed Decisionmaking. As the No Child Left Behind Act (NCLB) and data-informed decisionmaking gain the attention of school leaders, the need for a common language and definition of terms has become evident.

The Data-informed Decisionmaking (DID) LIG was charged to create a glossary of data terms. The basis for this glossary was developed by the Governor’s Office of Student Achievement and the Georgia Leadership Institute for School Improvement. Many terms in this glossary are specific to Georgia (and noted as such). Other states may have different definitions of terms and these will be collected and added over time.

Terms from several glossaries were selected to be included in this glossary and their source is documented. This glossary is a “living document” as terms will continue to be added and refined to serve the needs of Wallace Education Leadership Action Network. Contact information appears at the end of the glossary.

The Glossary

Absolute Bars – Or annual measurable objectives (AMOs), were set for the Georgia Criterion Referenced Competency Test (CRCT) and the Enhanced Georgia High School Graduation Test (E-GHSGT). These absolute bars are specified annual targets based on the percent of students identified as proficient in reading/English language arts and mathematics. These bars will be increased periodically so that all students reach 100% proficiency by school year 2013-2014. (Source A – Georgia-specific definition)

Academic Performance – One component of the Adequate Yearly Progress (AYP) process to produce annual determinations of adequacy. This component requires schools, Local Education Agencies (LEAs), and states to achieve specified levels of proficiency in the academic areas of reading/language arts and mathematics. Also known as annual measurable objective (AMO). (Source A)

Accommodations and Adaptations – Modifications in the way assessments are designed or administered so that students with disabilities (SWD) and limited English proficient students can be included in the assessment. Assessment accommodations or adaptations might include Braille forms for blind students or tests in native languages for students whose primary language is other than English. (Source B)

Accountability – The demand by a community (public officials, employers and taxpayers) for school officials to prove that money invested in education has led to measurable learning. "Accountability testing" is an attempt to sample what students have learned, or how well teachers have taught, and/or the effectiveness of a school's principal's performance as an instructional leader. School budgets and personnel promotions, compensation, and awards may be affected. Most school districts make this kind of assessment public; it can affect policy and public perception of the effectiveness of taxpayer-supported schools and be the basis for comparison among schools. (Source D)

Achievement – The demonstration of student performance measured against agreed-upon standards. (Source A)

Action Research – School- and classroom-based studies initiated and conducted by teachers and other school staff. Action research involves teachers, aides, principals, and other school staff as researchers who systematically reflect on their teaching or other work and collect data that will answer their questions. It offers staff an opportunity to explore issues of interest to them in an effort to improve classroom instruction and educational effectiveness. (Source D)

Adequate Yearly Progress (AYP) – The measure used to hold schools, LEAs, and the state responsible for student achievement. Georgia's AYP plan requires schools, LEAs, and the state to demonstrate 95% participation on annual tests, certain specified levels of proficiency in academics (reading/language arts and mathematics) and an additional indicator. AYP is based on a series of performance goals that every school, LEA and state must achieve within specified timeframes in order to meet the 100% proficiency goal established by the federal No Child Left Behind Act of 2001 (NCLB). (Source A)

Alignment – The process of linking content and performance standards to assessment, instruction, and learning in classrooms. One typical alignment strategy is the step-by-step development of (a) content standards, (b) performance standards, (c) assessments, and (d) instruction for classroom learning. Ideally, each step is informed by the previous step or steps, and the sequential process is represented as follows:

Content Standards – Performance Standards – Assessments – Instruction for Learning

In practice, the steps of the alignment process will overlap. The crucial question is whether classroom teaching and learning activities support the standards and assessments. System alignment also includes the link between other school, district and state resources. Alignment supports the goals of the standards, i.e., whether professional development priorities and instructional materials are linked to what is necessary to achieve the standards. (Source B)

Alternative Assessment (also authentic or performance assessment) – An assessment that requires students to generate a response to a question rather than choose from a set of responses provided to them. Exhibitions, investigations, demonstrations, written or oral responses, journals, and portfolios are examples of the assessment alternatives we think of when we use the term "alternative assessment." Ideally, alternative assessment requires students to actively accomplish complex and significant tasks, while bringing to bear prior knowledge, recent learning, and relevant skills to solve realistic or authentic problems. Alternative assessments are usually one key element of an assessment system. (Source B)

Anchor(s) – A sample of student work that exemplifies a specific level of performance. Raters use anchors to score student work, usually comparing the student performance to the anchor. For example, if student work is scored on a scale of 1-5, there would typically be anchors (previously scored student work) exemplifying each point on the scale. (Source B)

Analysis – The process of examining a situation to fully understand its elements and their relationships, with the end goal of differentiating the problem's root cause from its symptoms. *See also Root Cause and Symptom.* (Source A)

Analytic Rubric – Also called analytic scoring rubric. A rubric that applies several distinct criteria to evaluate learner products and performances. In effect, a performance is assessed several times, using the lens of a separate criterion each time. For example, when using an analytic rubric to evaluate essays, a teacher might evaluate five traits: organization, use of detail, attention to audience, persuasiveness, and conventions. *See also Rubric.* (Source A)

Analytic Scoring – A procedure in which performances are evaluated for selected dimensions or traits, with each trait receiving a separate score (i.e., organization, use of details, language usage, mechanics, etc.) Analytic scoring involves a quantitative judgment of whether identified dimensions or traits are more or less present. Analytic scoring reports a profile of characteristics or attributes. Analytic scoring guides specify criteria to be assessed, provide a separate score for each criterion, and may include a composite score for overall performance. In some cases, the composite score is weighted based on the importance of each dimension. (Source A)

Anchor Tests – 1. A common set of items administered with each of two or more different forms of a test for the purpose of equating the scores obtained on these forms. 2. "Anchor papers provide a connection between a rubric narrative and student writing, and an example of what writing at a certain score of the rubric should look like." (Source C)

Annual Measurable Objective (AMO) – Specified targets based on the percent of students identified as proficient in reading/English language arts (Georgia combines reading and English/language arts) and mathematics. (Source A – Georgia-specific definition)

Assessment – The process of quantifying, describing, gathering data about or giving feedback about performance in relation to established standards and criteria. In the area of learner achievement, assessment is gathering and interpreting information about a learner or group of learners, using a variety of tools and techniques, including paper-pencil tests, exhibits, interviews, surveys, observations, etc. Good assessment requires a balance of types because each in isolation is limited and prone to error. It is the act of describing performance, primarily for the purpose of enhancing learning. (Source A)

Assessment FOR Learning – Assessment that typically is formative (before or during the learning). Assessment FOR Learning includes descriptive feedback, peer assessment, self-assessment, etc. and is used for the purpose of helping the learner learn. Assessment that makes learning more possible. (Source A)

Assessment OF Learning – Assessment that typically is summative (after the learning) and looks at learning to decide how much has been learned and report on it. (Source A)

Assessment Literacy – The possession of knowledge about the basic principles of sound assessment practice, including terminology, the development and use of assessment methodologies and techniques, familiarity with standards of quality in assessment. Increasingly, familiarity with alternatives to traditional measurements of learning. (Source D)

Assessment Plan – A set of choices regarding how learner learning will be assessed in relation to identified standards and criteria. (Source A)

Assessment System – The combination of multiple assessments into a comprehensive reporting format that produces comprehensive, credible, dependable information upon which important decisions can be made about students, schools, districts, or states. An assessment system may consist of a norm-referenced or criterion-referenced assessment, an alternative assessment system, and classroom assessments. (Source B)

Attendance – One of the menu of second indicator options for elementary and middle schools. To reflect attendance, Georgia is extracting the number of days absent for any student enrolled for any time during the school year and determining the distribution in the categories of 0-5 days absent, 6-15 days absent, and more than 15 days absent. The criteria for the second indicator is, *No more than 15% of students absent more than 15 days in one school year or show progress from the preceding year.* For purposes of AYP, attendance follows the NCLB grade levels; for example, a K-5 school's attendance is based on grades 3 through 5 while a K-2 school's attendance would have to be based on grades 1 and 2. (Source A – Georgia-specific definition)

AYP Workbook – The document that is officially known as the *Consolidated State Application Accountability Workbook*. Each state annually submits to the United States Department of Education (US ED) its workbook describing how AYP determinations will be calculated and how the state will comply with the No Child Left Behind Act of 2001. Updates to the AYP Workbook reflect the state's response to feedback from Georgia education stakeholders, analysis of academic-related data, and changes in state curriculum and assessments, state laws, and federal legislation and/or guidance/regulations from US ED. (Source A – Georgia-specific definition)

Balanced Scorecard – An operational and strategic management tool, a measurement and monitoring system for organizational performance, and a communication tool. The Balanced Scorecard contains a wide range of performance indicators that monitor student achievement, student and stakeholder loyalty and engagement, and operational effectiveness. (Source A)

Bar Graph – A pictorial representation that uses bars to display data, usually frequencies, percentages and averages. (Source F)

Baseline Data – Data collected to establish the starting point from which leaders can measure improvement toward their target goal. (Source A)

Benchmark – A detailed description of a specific level of student performance expected of students at particular ages, grades, or development levels. Benchmarks are often represented by samples of student work. A set of benchmarks can be used as "checkpoints" to monitor progress toward meeting performance goals within and across grade levels, i.e., benchmarks for expected mathematics capabilities at Grades 3, 7, ten, 10 graduation. (Source B)

Body of Evidence – A collection of data from multiple assessments of a learner’s proficiency as it relates to each standard. The assessments can include formative and summative assessments, anchor performance tasks and others, some required and others optional. In combination, these data guide the leader in making an informed judgment of the learner’s achievement of an established standard or target. (Source A)

Cause and Effect – The relationship and complexities between a problem or effect and the possible causes. (Source F)

Classroom Assessment – An assessment developed, administered and scored by a teacher or set of teachers with the purpose of evaluating individual or classroom student performance on a topic. Classroom assessments may be aligned into an assessment system that includes alternative assessments and either a norm-referenced or criterion-referenced assessment. Ideally, the results of a classroom assessment are used to inform and influence instruction that helps students reach high standards. (Source B)

Cohort – A group whose progress is followed by means of measurements at different points in time. (Source D)

Common Assessments – Assessments that are often developed by school districts, provided by a state item bank, or purchased through vendors, which are administered across the district in specified grades in academic content areas often on four to six week intervals. These common assessments are aligned to state academic standards and provide teachers feedback on how well individual students are meeting state standards – and how well they are teaching students. These assessments are often coordinated with a district’s “curriculum map” which requires teachers to teach state standards on a specific schedule. The timely feedback gives educators information on student progress on standards and thus allows quick interventions and support for students not yet meeting the standards.

Common Core of Data (CCD) – The database of public school and school district data annually collected by the U.S. Department of Education's National Center for Education Statistics. CCD contains fiscal and non-fiscal data about all public schools, public school districts and state education agencies in the United States. The database includes descriptive information, including school and district name, address, and phone number; information about students and staff, including demographics and enrollments; and fiscal data, including revenues and current expenditures. (Source E)

Confidence Interval – Used in inferential statistics, a range of values that a researcher can estimate, with a certain level of confidence, where the population parameter is located. (Source F)

Confidence Interval – When a reporting group does not make Adequate Yearly Progress by meeting the absolute bars or Annual Measurable Objectives, then a confidence interval approach is used. Georgia’s AYP plan uses a confidence interval approach as a second look at a group’s proficiency in order to strengthen the judgment of whether a group does or does not meet the criteria for AMO. (Source A Georgia-specific definition)

Constructed Response Item – An exercise for which examinees must create their own responses or products (performance assessment) rather than choose a response from an enumerated set (multiple choice). (Source C)

Content Standards – Broadly stated expectations of what students should *know and be able to do* in particular subjects and grade levels. Content standards define for teachers, schools, students, and the community not only the expected student skills and knowledge, but what schools should teach. An example of a language arts standard is: "Fourth-grade students will be able to gather information for a report using sources such as interviews, questionnaires, computers, and library centers." (Source B)

Continuous Improvement – Keeping the process of planning, implement, evaluating and improving alive over time. (Source F)

Continuous Improvement Plan – A plan for improvement, based on data that will help create and manage change. The process of answering the following questions: *Who are we? How do we do business? What are our strengths and areas for improvement? Why do we exist? Where do we want to be? What are the gaps? What are the root causes of the gaps? How can we get to where we want to be? How will we implement? How will we evaluate our efforts?* (Source F)

Correlation – A statistical analysis that helps one see the relationship of scores in one distribution to scores in another distribution. Correlation coefficients have a range of -1.0 to +1.0. A correlation of around zero indicates no relationship. Correlations of .8 and higher indicate strong relationships. (Source F)

Criteria – The dimensions or characteristics of standards used to judge learner work. When combined with a scale and performance descriptions, these elements valued in learner performance become rubrics or scoring guides to be used in assessment. Scoring guides embody and express criteria. For example, criteria for good reading might include characteristics such as fluency, flexibility, making connections with personal experience. (Source A)

Criterion-Referenced Assessment – An assessment where an individual's performance is compared to a specific learning objective or performance standard and not to the performance of other students. Criterion-referenced assessment tells us how well students are performing on specific goals or standards rather than just telling how their performance compares to a norm group of students nationally or locally. In criterion-referenced assessments, it is possible that none, or all, of the examinees will reach a particular goal or performance standard. (Source C)

Criterion Referenced Competency Test (CRCT) – State-required tests to measure student acquisition of the knowledge and skills set forth in the state curriculum. Georgia law requires that these tests be administered to students in grades 1 through 8 in the content areas of reading, English/language arts and mathematics, and in grades 3 through 8 in science and social studies. CRCTs are designed to measure how well students have mastered the content and skills that are unique to Georgia's Quality Core Curriculum at each grade level. (Source A – Georgia-specific definition)

Curriculum Alignment – A curriculum in which what is taught, how it is taught, and how it is assessed is intentionally based on, but not limited to, the state Core Curriculum Content

Standards and Assessment – the sequence of learning in an aligned curriculum is articulated and constantly discussed, monitored and revised. (Source F)

Cut Score – A specified point on a score scale, such that scores at or above that point are interpreted or acted upon differently from scores below that point. (See also Performance Standard.) (Source C)

Dashboard – A visual display of district or school data that is similar to a car’s instrumentation. Often color-coded as red, yellow or green using dial instrumentation, it provides a quick interpretation of where a school or district is on many different achievement or other results.

Data Cleansing – The process of ensuring that all values in a database are consistent and correctly recorded. (Source F)

Data Dictionary – In database management systems, a file that defines the basic organization of a database. A data dictionary contains a list of all files in the database, the number of records in each file, and the names and types of each field. A data dictionary includes information such as the precise definition of data elements, user names, roles and privileges, stored procedures, general database structure, and space allocations. (Source G)

Data-informed Decisionmaking – A process of making decisions about curriculum and instruction based on the analysis of classroom data and standardized test data. Data-informed decisionmaking uses data on function, quantity and quality of inputs, and how students learn to suggest educational solutions. It is based on the assumption that scientific methods used to solve complex problems in industry can effectively evaluate educational policy, programs and methods. (Source H)

Data Quality Campaign – A national, collaborative effort to encourage and support state policymakers to: improve the collection, availability, and use of high-quality education data, and implement state longitudinal data systems to *improve student achievement*. The campaign aims to provide tools and resources that will assist states in their development of quality longitudinal data systems, while also providing a national forum for reducing duplication of effort and promoting greater coordination and consensus among the organizations focusing on improving data quality, access and use. See www.dataqualitycampaign.org

Data Warehouse – Repository of an organization's data where the informational assets of the organization are stored and managed to support various activities such as reporting, analysis, decisionmaking, as well as other activities such as support for optimization of organizational operational processes. (Source G)

Database – A collection of information organized in such a way that a computer program can quickly selected desired pieces of data, similar to an electronic filing system. Traditional databases are organized by fields, records and files. A field is a single piece of information, a record is one complete set of fields, and a file is a collection of records.

Database Management System (DBMS) – A collection of programs that enables an individual to store, modify and extract information from a database. There are many different types of DBMSs, ranging from small systems that run on personal computers to huge systems that

operate on mainframes. The following are examples of database applications: computerized library systems, automated teller machines, flight reservation systems, or a student information system. (Source G)

Diagnostic – Assessment/evaluation carried out prior to instruction that is designed to determine a leader’s knowledge, skills, or misconceptions in order to identify specific learner needs. See also, *Formative and Summative*. (Source A)

Dimensions – Desired knowledge or skills measured in an assessment and usually represented in a scoring rubric. For example, a measurement of student teamwork skills on a performance assessment might include six dimensions: adaptability (recognizing problems and responding appropriately), coordination (organizing team activities to complete a task on time), decision-making (using available information to make decisions), interpersonal (interacting cooperatively with other team members), leadership (providing direction for the team), and communication (clearly and accurately exchanging information between team members). (Source B)

Demographic Data – Information describing the students and citizens of a school district; usually includes such characteristics as race, age, gender, income, educational level, birth rate, and profession. (Source E)

Disaggregated Data – Test results are sorted into groups of students by gender, those who are economically disadvantaged, by racial and ethnic minority groups, disabilities or by English fluency. This practice allows administrators and teachers to see more than just the average score for the school. Instead, administrators and teachers can see how each student group is performing. (Source F)

Economically Disadvantaged (ED) – In accordance with NCLB, AYP results must be disaggregated by students’ economic status. For a student to be classified as ED, the student must be coded in the Student Record as eligible for free or reduced-price lunch anytime during the school year. (Source A)

Electronic Grade Book – Usually a web-based resource where student information is entered by teachers who have access to these records. Electronic grade books often include the student’s attendance record, teacher grades and achievement scores on district and state assessment results. In some school districts, this information is available to parents using a password.

Elementary and Secondary Education Act (ESEA) – The federal education statute, originally passed by the U. S. Congress in 1965, that defines the role of the federal government in public education and authorizes many of the major federal education programs, including Title I. The act has been reauthorized by Congress several times, most recently in 2001 as the No Child Left Behind Act. (Source A)

Enhanced Georgia High School Graduation Tests (Enhanced GHSGT) – Georgia High School Graduation Tests (GHSGT) for English/language arts and mathematics were enhanced to comply with the No Child Left Behind Act of 2001 requiring more rigorous examinations. The Enhanced GHSGT is comprised of the base-GHSGT plus some additional items requiring more cognitive rigor in the areas of the areas of English/language arts and mathematics. For accountability purposes, Enhanced GHSGT results from first-time test-takers in the 11th grade

for English/language arts and mathematics are used in making AYP determinations at school, LEA, and state levels. (Source A- Georgia-specific definition)

Enrollment During Test Window (EDTW) – The student population enrolled for the duration of the state test window for the applicable test for the school, LEA, or the state. To derive EDTW, remove students with withdrawal date prior to or during the state test window from the population known as Students in AYP Grade Levels. The removal of these types of withdrawals results in Test Window Enrollees (TWE). The TWE count is then adjusted for any applicable federal flexibility such as medical emergencies. EDTW is the denominator used in calculating Test Participation Rate. (Source A – Georgia-specific definition)

Equity – The concern for fairness, i.e., that assessments are free from bias or favoritism. An assessment that is fair enables *all* children to show what they can do. At minimum, all assessments should be reviewed for (a) stereotypes, (b) situations that may favor one culture over another, (c) excessive language demands that prevent some students from showing their knowledge, and (d) the assessment's potential to include students with disabilities or limited English proficiency. (Source B)

Evaluation – Judgment regarding the collective quality, value or worth of responses, products or performances based on established criteria. Evaluations are usually based on multiple sources of information, a collection of evidence or snapshots gathered over a period of time. (Source A)

Examining Student Work – A process that engages teachers in a collective effort to learn about their practice by listening to a colleague(s) share samples of student work and provide focused feedback. A protocol or set of guiding questions is frequently used to structure the conversation so that the colleague describes the student work, receives and responds to questions about it, and explores implications for improving practice. (Source A)

Exemplars – Benchmark papers or performances. Exemplars provide a clear and stable reference point for providing feedback to learners, educators and the public in relation to standards. Exemplars give learners a clear picture of the targets they are aiming for, and they define levels of performance in concrete, meaningful and public ways. (Source A)

Flow Chart – A visual aid that illustrates the steps in a process. By visualizing the process, a flow chart can quickly help identify bottlenecks or inefficiencies where a process can be streamlined or improved. (Source F)

Formative Assessment – Ongoing assessment designed to provide direction for improvement and/or adjustment to a program for individual learners or for a whole class, for example: quizzes, initial drafts/attempts, homework, and questions during instruction. See also, *Diagnostic and Summative*. (Source A)

Full time equivalency files (FTE-1, FTE-3) – Counts of students reported to the Georgia Department of Education for purposes of funding. FTE-1 is reported in the fall and FTE-3 is reported in the spring. These files are used in conjunction with others to help determine enrollment during testing window and to identify full academic year students. See <http://techservices.doe.k12.ga.us/admin/datacollect/fte/fte2002.htm> for additional information on FTE. (Source A – Georgia-specific definition)

Full Academic Year (FAY) – Students who are continuously enrolled in the same school for school Adequate Yearly Progress (AYP) determination (or in same LEA for LEA AYP or in state for state AYP) from before the fall FTE count date through the end of the state’s spring testing window. (Source A – Georgia-specific definition)

Full Academic Year or FAY Participant – A student who is defined as continuously enrolled in a school or LEA before the October FTE count day through the end of the state’s test window for the student’s applicable grade level test. (Source A – Georgia-specific definition)

Gain Score – The difference between two administrations of the same test [or an equated test]. Gain scores are calculated by subtracting the previous score from the most recent score. One can have negative gains which are actually losses. (Source F)

Georgia Alternative Assessment (GAA) – An assessment administered to the small number of students with severe cognitive disabilities whose Individualized Education Program (IEP) teams determine they are unable to participate in the regular state-mandated assessments even with maximum accommodations. On the basis of ratings from five domains, the GaDOE and GOSA determine the proficiency level on the basis of the GAA. (Source A – Georgia-specific definition)

Georgia Department of Education (GaDOE) – The state agency charged with the fiscal and administrative management of certain aspects of K-12 public education, including the implementation of federal and state mandates. Such management is subject to supervision and oversight by the State Board of Education. (Source A – Georgia-specific definition)

Georgia High School Graduation Tests (GHS GT) – State-mandated curriculum-based assessments administered in grade 11 for graduation purposes. The tests are administered several times a year so that students have up to five opportunities to take each test within their 11th and 12th grade years. (Source A – Georgia-specific definition)

Goals – Achievements or end results. Goal statements describe the intended outcome of the vision and are stated in terms that are broad, general, abstract and non-measurable. Schools should have only two or three goals. (Source F)

Grade Level Equivalent – The school grade level for a given population for which a given score is the median score in that population. For example, if a test was administered during the month of October to a norming group of 6th-grade students and the median scale score obtained was 475, then the grade equivalent for a scale score of 475 on that test would be set at 6.1 – 6 representing Grade 6 and .1 representing the month of October (September is taken as the beginning of the school year and equals .0). (Source C)

Governor’s Office of Student Achievement (GOSA) – The state agency in Georgia mandated by state law to create a uniform performance-based accountability system for K-12 public schools that incorporates both state and federal mandates, including student and school performance standards. Additionally, GOSA is charged with the responsibility of publishing the State Report Card for schools and LEAs and to formulate a system of awards and consequences within the Single Statewide Accountability System. (Source A – Georgia-specific definition)

Grade – The number or letter reported at the end of a period of time as a summary statement of learner performance. (Source A)

Graduation Rate – Used as the second indicator for high schools. To comply with the No Child Left Behind Act, Georgia has defined a graduate as a student who leaves high school with a regular diploma (does not include Certificates of Attendance or Special Education diplomas) in the standard time (i.e., four years). High schools must be at or above the preset standard of 60% graduation rate or show an increase in graduation rate over the preceding year. (Source A – Georgia-specific definition) See also, *NGA graduation rate*.

Growth Models – Models of accountability that measure progress of tracking achievement scores of the same students from one year to the next with the intent of determining whether or not students have made progress. There are many different kinds of growth models including the commonly referenced “value-added” model, which takes into account not only growth, but other background characteristics in determining progress or the effect of a particular program or intervention. Growth models give schools credit for student improvement over time. These models assume that student performance is not simply a matter of where the school’s academic achievement is at any single point in time. Growth models assume a school’s ability to facilitate progress is a better indicator of its performance. (Source I)

Group – A combination of students based on a particular factor. In Georgia’s plan, Adequate Yearly Progress is not only determined for All Students (school, LEA or state LEA level) but also on groups based on race/ethnicity, disability, English language proficiency, and economic status. There are 10 groups: (1) All Students, (2) Asians/Pacific Islanders, (3) Black, (4) Hispanic, (5) Native American/Native Alaskan, (6) Multiracial, (7) White, (8) Students with Disability, (9) Limited English Proficient, and (10) Economically Disadvantaged. Student factors are coded in the Student Record. (Source A – Georgia-specific definition)

High Stakes Test – A test used to provide results that have important, direct consequences for examinees, programs, or institutions involved in the testing. For example, in some high stakes tests such as high school exit exams, students who do not pass the examination do not receive a high school diploma, regardless of their performance in other areas of their school education. (Source C)

Holistic Scoring – In assessment, assigning a single score based on an overall assessment of performance rather than scoring or analyzing dimensions individually. The product is considered to be more than the sum of its parts so the quality of a final product or performance is evaluated rather than the process or dimension of performance. A holistic scoring rubric might combine a number of elements on a single scale. Focused holistic scoring may be used to evaluate a limited portion of a learner's performance. (Source D)

Holistic Rubric – A rubric used to obtain an overall impression of the quality of a performance or product. Typically, a holistic evaluation yields a single score. See also, *Rubric, Contrast Analytic Rubric*. (Source A)

Horizontal Articulation or Coordination – Indicates that the curriculum is carefully planned within grade levels. In effect, this would mean that every primary grade throughout the

school/district will teach the same curriculum – also every grade 6 social studies class; every grade 10 health class; every grade 12 physics class, and so on. (Source F)

Improvement Initiative Statement – A SMART goal based on CRCT or HSGT data and submitted to Georgia Leadership Institute for School Improvement (GLISI) as the focus for results during the school and district’s three-year partnership with GLISI. See also, *SMART Goal*. (Source A)

Individualized Education Program (IEP) – A written statement of special education, related services, and, as appropriate, transition services, that meets the unique needs of the student with a disability. An IEP also includes any specific accommodations, needed modifications, and supports for the student with a disability. The IEP is developed, reviewed and revised in a meeting by an appropriately staffed IEP team, including the student’s parent(s). (Source A)

Inter-rater Reliability – A term used to describe the degree of agreement in scores among two or more judges (raters). (Source A)

Interim Federal Flexibility for Students with Disability – A one-year flexibility granted to states that met compliance with certain NCLB principles. For 2004-05 and 2005-06 AYP, Georgia qualified for this flexibility. The plan provides a mathematical adjustment to the proficiency rates for the SWD group in both reading/English language arts and mathematics. This applies to schools and LEAs not making AYP solely on SWD group scores. (Source A – Georgia-specific definition)

Item Analysis – Analyzing each item on a test to determine the proportions of students selecting each answer. Item analysis be used to evaluate student strengths and weaknesses, and may point to problems with the test's validity and to possible bias. (Source D)

Lagging Indicator – A measure that reflects results in areas of performance that are critical to the success of the organization. For example, scores on state exams are lagging indicators because they become available only a few times a year and possibly much later than the effects which they report actually occurred. Lagging indicators may be predicted by leading indicators, such as student attendance, teacher quality, stakeholder engagement, etc. that occur nearer in time to the effect reflected by the lagging indicator. See also, *Leading Indicator*. (Source A)

Large-Scale Assessment – Assessment for purposes and audiences outside the classroom. Large-scale assessments are usually standard to some extent and given to large numbers of learners at the same time. Use of large-scale assessment includes accountability, program/curriculum evaluation, certification of learner competency (e.g., high school graduation), and selection of learners for special programs. Design of large-scale assessments suggests their best use is for supplying trend data for groups as opposed to individuals. (Source A)

Leading Indicator – A measure of performance that affects critical lagging indicators. For example, measures of student attendance are leading indicators that affect lagging indicators such as scores on state exams. Monitoring and measuring leading indicators allows time to make changes and improvements which can impact lagging indicators. See also, *Lagging Indicator*. (Source A)

Learner Achievement – The demonstration of learner performance measured against established standards of performance.

Limited English Proficient (LEP) or English Language Learners (ELL) – A student is identified as LEP or ELL on the basis of having a language other than English as the primary home language and as being eligible for supplemental English language acquisition services. In future, LEP students will be known as English Language Learners (ELL). (Source A)

Local Educational Agency (LEA) – Local school system pursuant to local board of education control and management. (Source A)

Longitudinal Data System – Longitudinal data matches individual student records *over time*, from pre-kindergarten through 12th grade and into postsecondary education. According to the Data Quality Campaign, longitudinal data systems require these ten essential elements:

1. A unique statewide student identifier
2. Student-level enrollment, demographic and program participation information
3. The ability to match individual students' test records from year to year to measure academic growth
4. Information on untested students
5. A teacher identifier system with the ability to match teachers to students
6. Student-level transcript information, including information on courses completed and grades earned
7. Student-level college readiness test scores
8. Student-level graduation and dropout data
9. The ability to match student records between the preK-12 and higher education systems
10. A state data audit system assessing data quality, validity and reliability.

Matrix Sampling – An assessment method in which no student completes the entire assessment but each completes a portion of the assessment. Portions are allotted to different, representative samples of students. Group (rather than individual) scores are obtained for an analysis of school or district performance. (Source H)

Mean – One of several ways of representing a group with a single, typical score. Mean is figured by adding up all the individual scores in a group and dividing them by the number of people in the group. Can be affected by extremely low or high scores. (Source D)

Median – The point on a scale that divides a group into two equal subgroups. Median is another way to represent a group's scores with a single, typical score. The median is not affected by low or high scores as is the mean. (Source D)

Medical Emergency – US ED flexibility that allows a student who qualifies to be exempted from the Enrollment During Test Window (or test participation rate calculation). A medical emergency is a rare occurrence. Examples include (1) a student involved in a car wreck during the test window and sustains serious injury that prevents student from taking test, and (2) a student who may be out of LEA area for chemotherapy during the test window and cannot take test via LEA homebound teacher. Schools are required to keep documentation of medical

emergencies. Schools identify any student who qualifies as a medical emergency exemption during the Student Record and NCLB applications. (Source A)

Minimum N – The requisite group size in order for a group’s performance to be used in AYP determinations. Beginning in 2005, Georgia’s minimum group size for AYP groups is 40 members or 10% of the Students in AYP Grade Levels, whichever is greater, with a cap of 75 students. This applies to academic performance and second indicator only. The minimum number for participation is always 40. (Source A)

Mode – The score that occurs most frequently in a scoring distribution. (Source F)

Multiple Measures – Evaluations that provide more than one way for learners to demonstrate attainment of a standard. Learners need to be provided with multiple opportunities to perform in relation to standards. Multiple opportunities to perform can apply to the assessment approach (open-ended vs. close-ended), format (constructed response vs. selected response), or context (on demand vs. over time, the setting, the purpose of the assessment). (Source A)

Multi-Year Average as Second Look at AMO (MYA) – A second look at a group’s academic performance. This second look permits a further examination of the group’s performance to ensure that the judgment about whether the school, LEA, or state is did not make AYP is reliable. The MYA method allows the averaging of up to three years of proficient students. The MYA is then compared with the current year AMO to see if the group met the academic performance component of AYP. (Source A)

Multi-year Average Participation (MYAP) – If a school does not meet the 95% participation rate criterion, then Georgia’s plan allows for the school’s participation rates to be averaged up to three years to determine if the average meets the 95% criterion. This flexibility was allowed by US ED beginning in 2003-04, and Georgia used a two-year average. Starting with 2005 AYP, Georgia will use a three-year average. (Source A)

Needs Improvement – An identification for a school or LEA that has not made AYP for two or more consecutive years in the same subject for schools and in the same subject for both elementary and secondary school grade spans for LEAs. (Source A)

NGA Graduation Rate – A formula for calculating a uniform graduation rate endorsed by all 50 governors in 2005:

$$\text{Graduation rate} = \frac{\text{[on-time graduates in year x]}}{\text{[(first-time entering 9th graders in year x-4) + (transfers in) - (transfers out)]}}$$

Graduates are those earning a high school diploma or state approved modified diploma for students with individualized education plans. Passing General Educational Development (GED) tests or receiving a certificate of completion are *not* considered high school graduates by this definition created by the task force. This formula provides a more accurate graduation rate than can be compared across states.

No Child Left Behind Act of 2001 (NCLB) – A reauthorization of the Elementary and Secondary Education Act of 1965, the principal federal law affecting education from kindergarten through high school. NCLB is designed to improve student achievement and close

achievement gaps. States are required to develop challenging academic standards, to educate all students to 100% proficiency by 2014, and to create and implement a single, statewide accountability system. (Source A)

Norm-Referenced Assessment – An assessment designed to compare the performance of an individual learner or group to another learner or group by distributing performance across a normal curve and in which not all learners assessed can perform at the highest level. For example, “She typed better than 80% of her classmates.” This type of assessment is frequently contrasted with criterion-referenced assessment. (Source A)

Normal Curve Equivalent (NCE) – A score that ranges from 1-99, often used by testers to manipulate data arithmetically. Used to compare different tests for the same student or group of students and between different students on the same test. An NCE is a normalized test score with a mean of 50 and a standard deviation of 21.06. NCEs should be used instead of percentiles for comparative purposes. (Source D)

On-Demand Assessment – An assessment that takes place at a predetermined time and place, usually under uniform conditions for all students being assessed. The SAT, district and state tests, and most in-class unit tests and final exams are examples of on-demand assessments. (Source B)

Open-Ended – A term used to describe tasks and questions that do not lead to a single right answer. Rather, many different acceptable answers are possible. Such answers are justified, plausible, or well-defended as opposed to correct. Essay test questions, for example, are open-ended, whereas multiple-choice tests are not (by design). (Source A)

Opportunity to Learn (OTL) – Providing students with the teachers, materials, facilities, and instructional experiences that will enable them to achieve high standards. Opportunity to learn is what takes place in classrooms that enables students to acquire the knowledge and skills that are expected. OTL can include *what* is taught, *how* it is taught, by *whom*, and with *what* resources. (Source B)

Peer Assessment – A reflective practice in which learners make observations about their peers’ performances relative to expectations or specific criteria. (Source A)

Percentile – A ranking scale ranging from a low of 1 to a high of 99 with 50 as the median score. A percentile rank indicates the percentage of a reference or norm group obtaining scores equal to or less than the test-taker’s score. A percentile score does not refer to the percentage of questions answered correctly, it indicates the test-taker’s standing relative to the norm group standard. (Source D)

Performance Assessment – Direct observation and judgment of learner products or performances. Good-quality performance assessment has pre-established performance criteria. In standards-based assessment, the criteria are taken directly from the standards. (Source A)

Performance Standard – An established level of achievement, quality of performance, or degree of proficiency. A performance standard specifies how well learners are expected to achieve or perform. Georgia performance standards (GPS) include the standard description,

sample tasks, and samples of student work and teacher commentary on that work. (www.doe.k12.ga.us and see GPS). (Source A – Georgia-specific definition)

Performance Task – A task that uses one’s knowledge to effectively act or bring to fruition a complex product in which one’s knowledge and expertise are revealed (also called simply “performance”). A performance task is observable work with a beginning, end and one or more outputs. A task statement begins with one verb that states the main activity and may include a direct object of the verb, for example, “Analyze CRCT data.” (Source A)

Pie Chart or Pie Graph – A graph in a circular format used to display percentages, for example, the percentage of a school or district population by ethnicity. (Source F)

Pivot Table – A powerful data summarization tool in Microsoft Excel and other electronic spreadsheet programs. Among other functions, it automatically sorts, counts and totals data stored in a spreadsheet and creates a second table displaying the organized data. Its name describes the drag and drop "rotation" or pivoting of the summary table as selected variables or rows and columns are moved. New totals or other calculations are automatically updated. Pivot tables are useful to quickly create crosstabs. The name Pivot Table is a trademark of Microsoft Corporation. (Source G)

Processes – Measures that describe what is being done to get results, such as programs, strategies and practices. (Source F)

Portfolio – A purposeful, integrated collection of learner work that tells the story of the learner’s efforts, progress, and achievement of standards in a given area or in multiple areas. The portfolio typically has two purposes: (1) to provide documentation of a learner’s work and (2) to serve as the basis for evaluation of work in progress or work over time. (Source A)

Portfolio Assessment – A portfolio is 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. Portfolios can be designed to assess student progress, effort, and/or achievement, and encourage students to reflect on their learning. (Source B)

Prerequisite Knowledge / Skill – The knowledge and skill required to accomplish the goals of a unit in a culminating and complex performance task. Typically, prerequisites identify the more discrete knowledge and know-how required to put everything together in a meaningful final performance. For example, knowledge of the USDA food pyramid guidelines would be considered a prerequisite to the task of planning a healthy and balanced diet for a week. (Source A)

Primary Trait Method – A type of rubric scoring constructed to assess a specific trait, skill, behavior, or format, or the evaluation of the primary impact of a learning process on a designated audience. (Source D)

Proficiency – Having or demonstrating an expected degree of knowledge or skill in a particular area. (Source A)

Quality Program Indicator – A variable reflecting effective and efficient program performance; distinguished from a measure (data used to determine the level of performance) and a performance standard (the level of acceptable performance in terms of a specific numeric criterion). (Source C)

Quartile – The breakdown of an aggregate of percentile rankings into four categories: the 0-25th percentile, 26-50th percentile, etc. (Source D)

Query – A request made to a database in which the response is returned to the desktop. Understanding and knowing how to set up queries or to ask questions of the database is very important to the information discovery request. (Source F)

Quintile – Breakdown of an aggregate of percentile rankings into five categories: the 0-20th percentile, 21-40th percentile, etc. (Source D)

Race/Ethnicity – Classification of a student by his/her primary race/ethnic group with which he/she self identifies. This information is coded in the Student Record. There are six categories: (1) Asian/Pacific Islander, (2) Black, (3) Hispanic, (4) Multiracial, (5) Native American/Alaskan, and (6) White. (Source A)

Rater – A person who evaluates or judges student performance on an assessment against specific criteria. (Source B)

Rater Training – The process of educating raters to evaluate student work and produce dependable scores. Typically, this process uses anchors to acquaint raters with criteria and scoring rubrics. Open discussions between raters and the trainer help to clarify scoring criteria and performance standards, and provide opportunities for raters to practice applying the rubric to student work. Rater training often includes an assessment of rater reliability that raters must pass in order to score actual student work. (Source B)

Reliability – How accurately a score will be reproduced if an individual is measured again; the degree to which the results of an assessment are dependable and consistently measure particular student knowledge and/or skills. Reliability is an indication of the consistency of scores across raters, over time, or across different tasks or items that measure the same element. Thus, reliability may be expressed as (a) the relationship between test items intended to measure the same skill or knowledge (item reliability), (b) the relationship between two administrations of the same test to the same student or students (test/retest reliability), or (c) the degree of agreement between two or more raters (rater reliability). An unreliable assessment cannot be valid. (Source C)

Report Card – The formal mechanism used to communicate learner achievement of standards at designated points during the year. (Source A)

Root Cause – The primary underlying cause of a problem or situation. It is different from the problem's symptoms, which merely indicate that the problem exists. (Source A)

Rubric –An established set of parameters for scoring or rating learners’ performance on tasks. This scoring guide enables raters to make reliable judgments about learner work and enables learners to self-assess their work. Good rubrics consist of a fixed measurement scale (e.g., four-point), a set of clear criteria, and performance descriptions for each criterion at each point on the scale that signify the degree to which the criteria have been met. (Source A)

Safe Harbor – The last step in determining AYP status if the confidence interval approach and multi-year averaging do not enable a group of students to make AYP. To make safe harbor, a reporting group must decrease the percent not meeting proficiency/advanced levels by 10% and must meet the additional academic indicator requirement. (Source A)

Scale –Values given to student performance. Scales may be applied to individual items or performances. Examples include: *checklists*, i.e., yes or no; *numerical*, i.e., 1-6; or *descriptive*, i.e., the student presented multiple points of view to support his/her essay. Scaled scores occur when participants' responses to any number of items are combined and used to establish and place students on a single scale of performance. (Source B)

Scale Scores – A score to which raw scores are converted by numerical transformation (e.g., conversion of raw scores to percentile ranks or standard scores); units of a single, equal-interval scale that are applied across levels of a test. (Source C)

Scientifically Based Research – NCLB mandates that educators base school programs and teaching practices on scientifically based research, which includes everything from teaching approaches to drug abuse prevention. Scientifically based research means research that involves the application of rigorous, systematic and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs. (Source F)

Scoring – Assigning a number or letter to a particular piece of learner work or performance. (Source A)

Scoring Guide – A set of guidelines for rating student work. A scoring guide describes what is being assessed, provides a scoring scale, and helps the teacher or rater correctly place work on the scale. (Source H)

Second Indicator – One component of the Adequate Yearly Progress (AYP) process to produce annual determinations of adequacy. This component requires ALL students group and any group using Safe Harbor to demonstrate academic performance to meet requirements on an additional academic-related indicator. Georgia’s plan provides flexibility for the second indicator component by offering elementary and middle schools a menu of options. For high schools graduation rate is the federally required second indicator. (Source A)

Selected Response Items / Tasks – Items or tasks in which learners select from among response or answer choices presented to them (for example, true/false, matching, multiple choice). (Source A)

Self-Assessment – A reflective practice in which learners make observations about their own performance relative to criteria and standards. (Source A)

Schools Interoperability Framework (SIF) – A set of rules and definitions which enable software programs from different companies to share information, created by an association of developers and vendors of school technology and the federal, state and local educators who use that technology. This set of platform-independent, vendor-neutral rules and definitions is called the SIF Implementation Specification. The SIF Specification makes it possible for programs within a school or district to share data without any additional programming and without requiring each vendor to learn and support the intricacies of other vendors’ applications. (Source G)

Single Statewide Accountability System (SSAS) – The SSAS merges both federal and state education laws that relate to K-12 school accountability for student academic performance. The SSAS includes a school and system profile, which includes three components: (1) AYP, (2) Progress Index, and (3) Performance Highlights. School profiles include all three components and system profiles include AYP and Performance Highlights. The Progress Index provides school awards based on a school’s progress over the prior year for greatest gains or highest percent. The Performance Highlights capture a school’s and LEAs top academic-related indicators based primarily on State Report Card data. State Board Rules (SSAS 160-7-1-.01; Definitions 160-7-1-.02; Accountability Profile 160-7-1-.03; and Awards and Consequences 160-7-1-.04) provide more details of how the federal requirements of NCLB have been merged with state requirements. (Source A – Georgia-specific definition)

SMART Goals – Used for both organizational and personal goal setting, “SMART” is an acronym that helps leaders remember to develop goals that meet each these criteria:

- S = Specific – who, what, where, when, in what context
- M = Measurable – can be quantified; data is available
- A = Attainable – realistic; achievable by an average performer
- R = Relevant –aligned to mission, vision, objectives, etc.
- Results Focused – quantifies specific results expected
- T = Timely – start and end dates specified

Additionally, SMART goals are written concisely and includes the following elements related to measurement:

- Indicator – the specific measurable data point to be tracked
- Baseline – an established performance measurement
- Target – the desired measurable level sought over a specific time span. (Source A)

Standards – The broadest of a family of terms referring to statements of expectations for student learning, including content standards, performance standards, and benchmarks. (Source B)

Standards-Based System – A system in which the classroom curriculum is designed to help learners attain defined standards. There is congruence among a focus on standards, the learning-teaching activities and materials selected to engage learners, and the assessments used to document learner attainment of the standards. Published materials, units of study, skill sequences, instructional experiences, routines or strategies, and assessments are standards-based only to the extent they link learners with standards within a classroom and across classrooms and grades, in a consistent and purposeful way. (Source A)

Standards-Based Reform – A program of school improvement involving setting high standards for all students and a process for adapting instruction and assessment to make sure all students can achieve the standards. (Source B)

Standardization – A set of consistent procedures for constructing, administering and scoring an assessment. The goal of standardization is to ensure that all learners are assessed under uniform conditions so that interpretation of their performance is comparable and not influenced by differing conditions. Standardization is an important consideration if comparisons are to be made between scores of individuals or groups. (Source A)

Standard Deviation – A measure of variability in a set of scores. The standard deviation is the square root of the variance. Unlike the variance, the standard deviation is stated in the original units of the variable. Approximately 68% of the scores in a normal distribution lie between plus one and minus one standard deviation. The more scores cluster around the mean, the smaller the variance. (Source F)

Standard Error of Measurement – The difference between an observed score and the corresponding true score or proficiency; the standard deviation of an individual's observed scores from repeated administrations of a test or parallel forms of a test, under identical conditions. Because such data cannot generally be collected, the standard error of measurement is usually estimated from group data. (Source C)

Statistically Significant – The result of the probability that the observed relationship in a sample occurred by pure chance; indicates that there is at least a 95% probability of the result that did not happen by chance. (Source F)

Stoplight Method of Data Analysis – The GLISI Stoplight Method of Data Analysis is an adaptation of a North Central Regional Educational Laboratory (NCREL) technique used in a web-based tutorial titled, “The Toolbelt: A Collection on Data-driven Decision-Making Tools for Educators” (www.ncrel.org/toolbelt/tutor.htm.) This tool enables leaders to use colored highlighters to analyze the Georgia Office of Student Achievement Annual Report Card to identify strengths and weak areas by looking at year-to-year change. (Source A – Georgia-specific definition, although many sites now use this.)

Student Identifier – A unique statewide student identifier is a single, non-duplicated number that is assigned to and remains with a student throughout his or her PreK-12 career. Assignment of a unique statewide student identifier to every student in the PreK-12 system provides a way to follow students as they move from grade to grade and across campuses and/or districts within the state. (Source J)

Student Information System – A system by which student information is stored and available electronically, often in a data warehouse. Includes student records described below.

Student Record – An annual record that provides cumulative information about a student for the school year, such as education history and demographics. This information contains LEA, school and student-level data that can be used for both state and LEA reporting and analysis. (Source A)

State Board of Education (SBOE) – The constitutional authority that defines education policy for the public K-12 education agencies in Georgia. (Source A – Georgia-specific definition)

State Educational Agency (SEA or State) – The Georgia State Board of Education. The State Superintendent of Schools implements the administrative functions on behalf of the Georgia State Board of Education. (Source A – Georgia-specific definition)

Students With Disabilities (SWD) – A broadly defined group of students with physical and/or mental impairments such as blindness or learning disabilities that might make it more difficult for them to do well on assessments without accommodations or adaptations. (Source B)

Student with Disability (SWD) – A student has been identified as SWD on the basis of being coded in the Student Record as eligible for special education services and with a primary exceptionality code. (Source A - Georgia-specific definition)

Summative Assessment – A snapshot of learner performance at a given point in time, judged according to pre-established standards and criteria; often a culminating assessment for a unit grade level, or course. Summative assessment typically leads to a status report on success or degree of proficiency in relation to identified standards and is used to make judgments about a learner’s achievement. (Source A)

Symptom – An indicator that a problem exists, but different from the true “root cause” of a problem. For example, low test scores in schools are not the root cause of a problem but they indicate that a problem exists. See also, *Analysis and, Root Cause*. (Source A)

Task – An activity, exercise, or question requiring students to solve a specific problem or demonstrate knowledge of specific topics or processes. (Source B)

Test Participant (TP) – A student who has a complete test record or a PTNA (present test not attempted) status and therefore is considered as participating in the assessment. Limited English Proficient (LEP) students who are also first time in a U.S. school and who have a presence of a LAB or ACCESS score will also be counted as TP. (Source A – Georgia-specific definition)

Test Participation – One component of the Adequate Yearly Progress (AYP) process to produce annual determinations of adequacy. This component requires schools, LEAs and states to achieve at least 95% participation of ALL students and any group meeting the minimum N size. This component captures the spirit of NCLB that every student is entitled to an annual assessment of his/her academic progress. (Source A)

Test Window – The state administration window during which a test is to be administered. For the CRCT, the spring 2006 test window was from April 3 through May 3. For the Enhanced GHSGT, the spring 2006 test window was from March 20 through March 31. These dates are important for determining which students are Test Window Enrollees and for determining full-academic year status of students. (Source A – Georgia-specific definition)

Test Window Enrollee – A student whose student record shows that he/she was enrolled in the school or LEA during the state’s test window for the grade level test that is applicable. Further

adjustments may be made to this population for US ED flexibility such as medical emergencies. (Source A)

Triangulation – A process of combining methodologies to strengthen the reliability of a design approach. When applied to alternative assessment, triangulation refers to the collection and comparison of data or information from three different sources or perspectives. (Source C)

Validity – The extent to which an assessment measures what it is supposed to measure and the extent to which inferences and actions made on the basis of test scores are appropriate and accurate. For example, if a student performs well on a reading test, how confident are we that that student is a good reader? A valid standards-based assessment is aligned with the standards intended to be measured, provides an accurate and reliable estimate of students' performance relative to the standard, and is fair. An assessment cannot be valid if it is not reliable. (Source B)

Vertical Articulation or Alignment – Indicates that the curriculum is carefully planned and sequenced from beginning learning and skills to more advanced learning and skills. Vertical articulation speaks to what is taught from pre-school through upper grades and is sometimes noted simply as “K-12 Curriculum.” (Source F)

SOURCES

Source A: Glossary Of Data-driven Decisionmaking Terms
Created by the Governor's Office of Student Achievement and The Georgia Leadership Institute for School Improvement, August 2006.

Source B: CRESST/UCLA, Assessment Glossary.
www.cse.ucla.edu/CRESST/pages/glossary.htm

Source C: Glossary of Useful Terms by the System for Adult Basic Education in Massachusetts
<http://www.sabes.org/assessment/glossary.htm>

Source D: Assessment Terminology: A Glossary of Useful Terms by New Horizons for Learning, <http://www.newhorizons.org/strategies/assess/terminology.htm>

Source E: Glossary by the Center for Public Education
<http://www.centerforpubliceducation.org/site/c.kjJXJ5MPIwE/b.1461881/k.D5A2/Glossary.htm>

Source F: Glossary of Terms by Victoria Bernhardt, in *Using Data to Improve Student Learning for School Districts*. Larchmont, NY: Eye of Education. 2003.

Source G: Wikipedia, www.wikipedia.com

Source H: North Central Regional Education Laboratory, www.ncrel.org

Source I: Commission on No Child Left Behind, Commission Staff Research Report. *Growth Models: An examination within the context of NCLB*. The Aspen Institute, 2006.

Source J: The Data Quality Campaign, www.dataqualitycampaign.org

Copyrights on previous sources do not allow this glossary to be published or sold.

Wallace DID Interest Group Glossary Subcommittee:

Jane Armstrong, Armstrong & Associates
Ann Duffy, Georgia Leadership Institute for School Improvement
Nancy Haight, Governor's Office of Student Achievement
Jeff Nelson, Focus on Results
Beata Thorstensen, New Mexico Department of Finance and Administration
Cynthia Yoder, Ohio Department of Education

To add or refine glossary terms contact:

Jane Armstrong, janearmstrong11@comcast.net