Dr. John D. Barge, State School Superintendent

"Making Education Work for All Georgians"

Student Learning Objectives Operations Manual

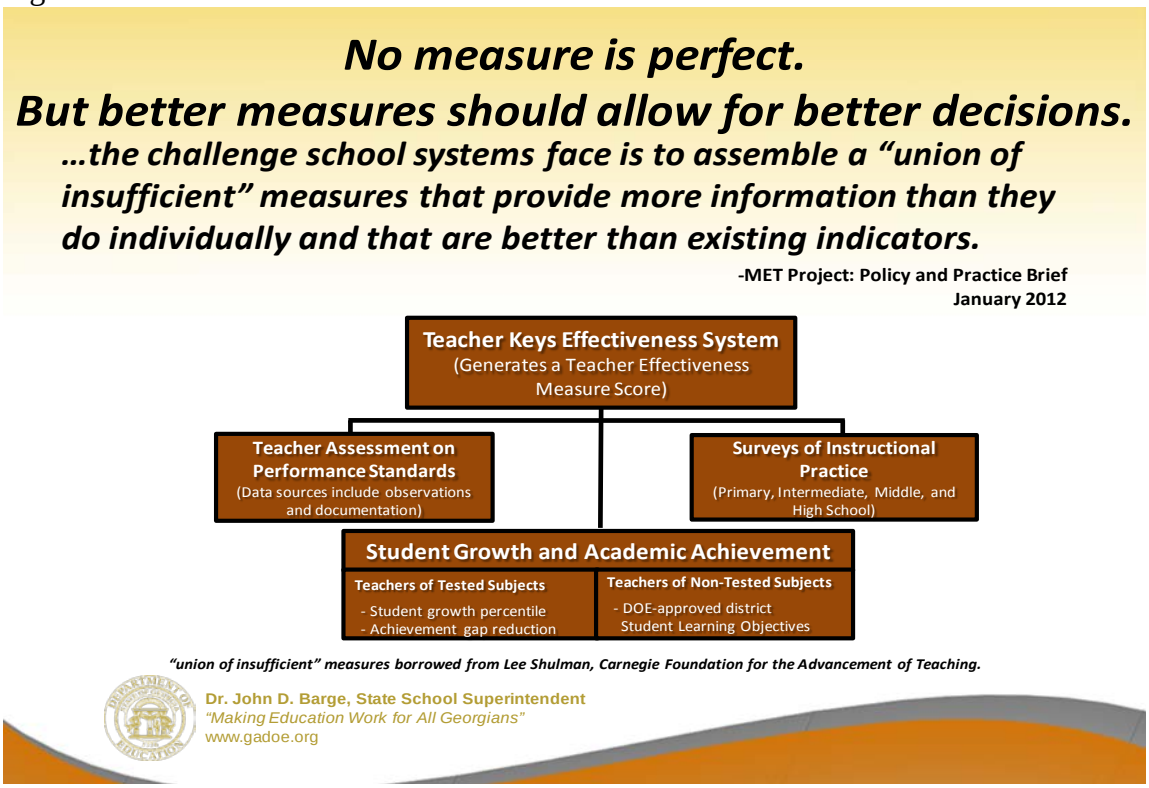
Contents

Introduction

In an effort to ensure that all schools and classrooms have great leaders and great teachers, Georgia, as part of the Race to the Top (RT3) plan has established the Teacher Keys Effectiveness System (TKES) and the Leader Keys Effectiveness System (LKES). As shown in Figure I, the TKES Evaluation System consists of three components which contribute to an overall Teacher Effectiveness Measure (TEM): Teacher Assessment on Performance Standards (TAPS), Surveys of Instructional Practice, and Student Growth and Academic Achievement. Georgia's focus on a multi-dimensional approach to teacher and leader evaluation will provide educators with high-quality and actionable feedback on their work with students, which will help them improve their effectiveness with students throughout their careers.

The focus of this manual is the Student Learning Objectives (SLOs) which comprise the Student Growth and Academic Achievement component of the teacher Keys Evaluations System for teachers of non-tested subjects.

Figure 1



Student Growth and Academic Achievement

Classrooms are complex places, and measuring student learning can be challenging due to unique grade level and subject characteristics. However, student learning is the ultimate measure of the success of a teacher and an instructional leader. The goal is to examine student growth and academic achievement by using components which will guide teachers as they design their instruction and determine student growth targets.

A vital component of the Teacher Keys Effectiveness System is Student Growth and Academic Achievement. For teachers of tested subjects, this component consists of a student growth percentile measure. Tested subjects include reading, English language arts, mathematics, science, and social studies for grades 4-8 and all high school courses for which there is an End-of-Course Test (EOCT).

Non-tested subjects include all courses not listed as tested subjects. Approximately 70-75% of all teachers teach non-tested subjects for at least some portion of the instructional day. For teachers of non-tested subjects, this component consists of the Georgia Department of Education (GaDOE)-approved Student Learning Objectives (SLOs) utilizing district-identified achievement growth measures. Districts have the option to develop SLOs for tested subjects, however this is not required.

The focus of this manual is the implementation of the SLOs development process. The professional practice of setting growth objectives to measure student growth is the cornerstone of the state's emphasis on using assessment results to guide instruction. Research has found that educators who set high quality objectives often realize greater improvement in student performance. Establishing this systematic approach will require unprecedented collaboration between state leaders, district leaders, and local school staffs. Under the leadership and participation of district leaders, content and assessment experts, along with effective classroom teachers, work together to determine one SLO for each course. Each district SLO is submitted to the GaDOE for review and approval.

Results from the 2011-2012 Pilot

Phase I: The training and preliminary work on SLOs began in October 2011. Over 37 training sessions, which included approximately 500 educators, were conducted. Each RT3 district chose 10 subjects for which to create an SLO. Phase I was designed for each district to learn the SLO development process and to write SLOs using current district assessments.

Two hundred five SLOs were submitted to the GaDOE for approval by December 2, 2011. A total of 306* SLOs were submitted during the month of December. Figure 2 describes the subjects for which the 234 reviewed SLOs were written:

Figure 2

Subject	ES	MS	HS
Reading/ELA	56		7
Math	37	3	8
Fine Arts	5	13	8
Physical Fitness	13	20	10
Social Studies		5	7
Science	8		12
Other	1	1	22
Total	**120**	**42**	**72**

- 120 or 51% were ES (78% of these SLOs were Reading/LA or Math)
- 42 or 18% were MS (48% were PE SLOs and the remaining SLOs were exploratory classes)
- 72 or 30% were HS (69% were non-EOCT academic class SLOs, 18% were CTAE-related, and 14% PE)

*As of 12.22.11, not all late submitted SLOs had been through the final review.

All Phase I SLOs are posted on the RT3 SharePoint site. SLOs are filed by state course number and by district. Well designed SLOs are noted and filed as exemplars on the SharePoint site. The RT3 SharePoint site is accessible only to districts who are piloting or implementing the Teacher Effectiveness System.

Phase II: The second phase of SLO submissions included 52 additional subjects which districts developed and submitted by August 1, 2012. The 52 additional subjects included pre-K and K-3 mathematics and reading/language arts and 32 high school courses. The most frequently taken courses for graduation were selected since those courses would impact the greatest number of teachers of non-tested subjects.

It was clear from Phase I that districts faced significant challenges in developing SLOs for courses for which the districts did not have valid and reliable measures. The RT3 SLO contacts met in January 2012 and again in February 2012 to consider a collaborative SLO development process which included collaborative development of valid and reliable measures. It was determined that 24 of the 26 RT3 districts would take the lead with two or three of the 52 Phase II SLOs and would develop the SLO as well as a pre and post-measure for the SLO. The GaDOE staff would train local educators on the qualities of effective measures during a three-day training and initiate the measurement development process. Several districts committed to

collaborating with other districts on certain subject SLOs and committed to involving higher education experts. The resulting SLOs and measures, called Georgia Public Domain SLOs, are currently posted on SharePoint for all districts to consider.

Twenty-five of these districts participated in the GaDOE three-day SLO training and completed SLO development locally. All RT3 districts will be required to create their own SLOs and measures for Phase II SLOs or adapt or adopt the public domain SLOs.

Next Steps in SLO Development: Current analysis of remaining undeveloped courses with SLOs, indicates that the remaining "most taught" courses, primarily include art, music, PE and Health, and foreign language courses. The GaDOE plans to support local districts as they develop their SLOs and accompanying SLO measures by sponsoring "content" weeks which include bringing together recommended teachers, state leaders, and other content experts to develop tasks which districts can utilize in their local SLO measures.

Student Learning Objective Overview

What is a Student Learning Objective (SLO)?

District determined SLOs are content-specific, grade level learning objectives that are measureable, focused on growth in student learning, and aligned to curriculum standards. As a measure of teachers' impact on student learning, SLOs give educators, school systems, and state leaders an additional means by which to understand, value, and recognize success in the classroom.

Purpose of SLOs

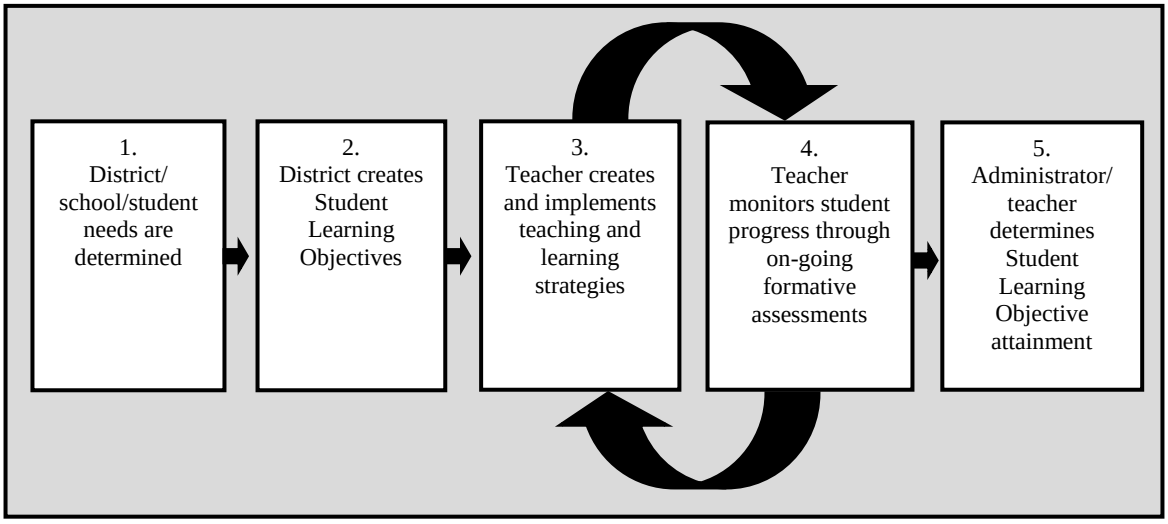
The primary purpose of SLOs is to improve student achievement at the classroom level. An equally important purpose of SLOs is to provide evidence of each teacher's instructional impact on student learning. The process of setting and using SLOs requires teachers to use common measures to determine student growth.

Teacher Effectiveness Measure (TEM) Requirements

1. Eventually SLOs will be written for all non-tested subject areas Pre-K through grade 12. This includes:
    a. All subjects in Pre-K through grade 2 (e.g. Language arts/reading, mathematics, science, social studies, fine arts, etc.) are non-tested subjects.
    b. All subjects in grade 3 are considered non-tested because there is no prior test score on which to determine Student Growth Percentile (SGP).
2. Teachers will be evaluated by one district SLO for each non-tested subject/course that they teach. SLOs are designed for the course, not individual teachers.
3. Teachers who teach both tested and non-tested subjects will be evaluated by SLOs for their non-tested subjects and by the SGP measure for their tested subjects.
4. If a teacher teaches the same course multiple periods/sections during the day, all students are included in the same SLO.
5. SLO results are reported at the student and class/group level. As teachers work with the district-designated SLOs, they should maintain a spreadsheet of each student's pre-assessment score and post-assessment score, as well as any other data needed to ascertain attainment of the SLO. Student data, classroom data, and school data will be submitted to the GaDOE, via an electronic method.
6. For Phase II SLOs, districts will submit SLOs on the District SLO Form for GaDOE approval no later than August 1, 2012. A separate form should be used for each SLO.
7. Prior to submission of district SLOs, appropriate district leaders should collect, review, and verify that each SLO is complete, aligned with content standards, and has rigor that is comparable to the standardized measures for tested subjects. Each superintendent or his/her

designee should approve and sign all SLOs prior to submission to the GaDOE. After Phase I, pre and post-assessments should also be submitted with all SLOs.

Overview of SLO Process



1. Districts, in collaboration with teachers and school leaders, examine current data and historical data to determine the focus of SLO for specified course.
2. Prior to the instructional period, districts develop an SLO based on the needs of students and/or school academic goals as they relates to the specified course, and the pre measure data. The District SLO form is completed and submitted to the GaDOE for review and approval by the specified date.
3. Using the approved district SLO for the specified course, teachers apply the SLO for their particular class(es) and complete the district-designated teacher SLO form which specifies how the teacher will implement the SLO with his/her class(es). Teachers and evaluators meet to discuss the teacher's SLO form/plan. (The GaDOE provides a Teacher SLO form. However, districts may choose to design and utilize their own form or method of reporting.)
4. Steps 3 and 4 are part of a recursive process, whereby the teacher continues to monitor student progress toward the given target.
5. Teachers and their evaluators will meet at the mid-point of the instructional period to review student progress. The purpose of this review is to determine if all students are on track to meet their growth targets or whether instructional interventions are warranted. This conference should identify the need and type of additional interventions necessary for student success.

6. At the end of the instructional period, the evaluator and teacher meet to review student data and progress. The evaluator scores the teachers' progress on the SLO Evaluation Rubric and submits the data to the GaDOE.

Essential SLO Components

Focus on student learning

By focusing on student learning, SLOs help teachers, principals, and districts pay close attention to the annual academic progress made by students (particularly those in non-tested subjects and grade levels). District-determined objectives are set using baseline data and are written with the expectation that student learning in each classroom will be measured against baseline data. Only those topics that clearly state expectations for student learning growth are to be included in objective setting. A teacher's professional growth objectives are not to be included.

Aligned with curriculum standards

SLOs must correlate with the Georgia Performance Standards (GPS), Common Core Georgia Performance Standards (CCGPS), or other national standards for the course being taught. District-selected standards should warrant the year-long or course-long focus of the students and teachers. They should be rigorous, measureable, and should deepen and extend knowledge for all students in the class/group/course. Each SLO must specify the exact course, subject, grade level, and set of standards for which it was designed.

Interval of instructional time

The interval of instruction is the length of time during which the SLO will be completed. Districts should determine the pre and post-measuremetn administration windows for each SLO. The majority of SLOs should be written for the entire length of the course being taught. However, the nature of specific courses may require that the pre-measure not be given at the very first of the instructional period but should be administered a short time into the instructional period. For example, in a beginning band class, students may need to learn to position and use their instruments before the progress on music standards can be pre assessed. For the majority of teachers, the instructional period is the full academic year. However, for teachers with courses that span only part of the academic, year, the instructional period will be the duration of that course, (e.g., a semester). The interval cannot change once approved.

Scope of SLOs

It is a district decision as to whether the SLO comprehensively addresses all standards taught in each course or if it addresses a prioritized set of standards. If a district chooses a set of prioritized standards, teachers are expected to address the entire curriculum and not exclude standards not assessed in the SLO.

Measureable objective

A measureable objective is one that quantifies growth in student learning, typically based upon the results of administration of pre- and post-measurements. Pre and post measurement scores are reported for each student in each teacher's class.

SLO measures

An SLO measure can be the instrument used to determine student learning of the objectives chosen. Each SLO must have a pre and post measure. Appropriate measures of student learning gains differ substantially based on the learners' grade level, content area, and ability level. Therefore the type and format of measures will vary based on the standards to be measured. Careful attention must be paid to how progress in relation to a given set of standards can most effectively be measured.

Commercially developed and validated measures that correlate with the standards selected for each subject SLO may be used. (Examples of externally-developed measures could include Advanced Placement tests, Lexile Framework for Reading, Dynamic Indicators of Basic Early Literacy Skills, DIBELS, etc.) Externally developed measures are selected, purchased, and used at each district's discretion. The GaDOE does not recommend any particular measures nor does the GaDOE endorse any particular product or assessment.

If aligned with the SLO's selected standards, the following measurement tools may be appropriate for determining student progress:

- Criterion-referenced tests, inventories, and screeners (e.g., state standardized tests, Advanced Placement tests, Scholastic Reading Inventory, Phonological Awareness Literacy Screening),
- School-adopted interim/common/benchmark assessments (e.g., county benchmark tests based on selected state standards, Career and Technical Education competency assessments, President's Physical Fitness Test),
- Authentic measures (e.g., learner portfolio, recitation, demonstration, performance) using district-developed performance scoring rubrics (e.g., writing rubrics) to document the performance,
- Regionally/locally developed common measures. Note: It is recommended that teacher-developed measures be considered as the last option only when other measures do not exist. If other measures do not exist, groups of teacher/district representatives with notable content expertise may develop common measures (test, rubrics, etc.). Beginning with Phase II SLOs, all locally/regionally developed common measures must be locally or regionally reviewed utilizing the SLO Table of Specifications and the SLO Assessment Criteria Table, as developed by the GaDOE.

The Foundation of Quality SLOs

The foundation of educational systems includes curriculum, instruction, and assessment. Curriculum, instruction, nor assessment can or should stand alone, and the alignment of these foundational systems is critical. Assessment is the process of using methods or tools to collect information about student learning. Careful and thoughtful attention to the selection and/or development of SLO measures is critical to the SLO process.

- Quality SLOs are built on quality measures.
- Quality measures inform teacher practice and student progress. A poor measure can negatively impact teacher instruction and student learning.
- Districts must develop measures that provide confidence and reassurance to teachers and administrators. In addition measures should directly inform instruction and offer true indications of attainment of the standards in the SLO. This collective confidence is built on content expertise, teacher input, increasingly valid and reliable measures, and immediate results which are meaningful for the teacher's work in the classroom.
- SLO measures should be selected and/or developed based on their appropriateness for the grade and content standards chosen for the SLO. Assessments may include traditional measures, performance or demonstrations, or work products, just to name a few.
- To the greatest extent possible, measures should be comparable between teachers, schools, and/or districts.
- SLOs provide unique opportunities for performance based measures which can give students models of high quality work.
- The capacity for developing increasingly valid and reliable measures should be built primarily at the local, regional and/or district level (and monitored at the state level).
- Improving locally developed measures will be a gradual, ongoing process. Measures and their results should be regularly examined by groups of trained educators using a common set of assessment criteria.
- It is not desirable that SLO measures become additional standardized tests.
- Well-designed SLOs should increase student achievement for students individually and collectively, therefore helping schools and districts attain their student achievement goals.

1. District/regional SLO Development Team(s)

It is critically important for districts or groups of districts to form trained teams for guiding the development of locally developed measures. Those team members might include the following:

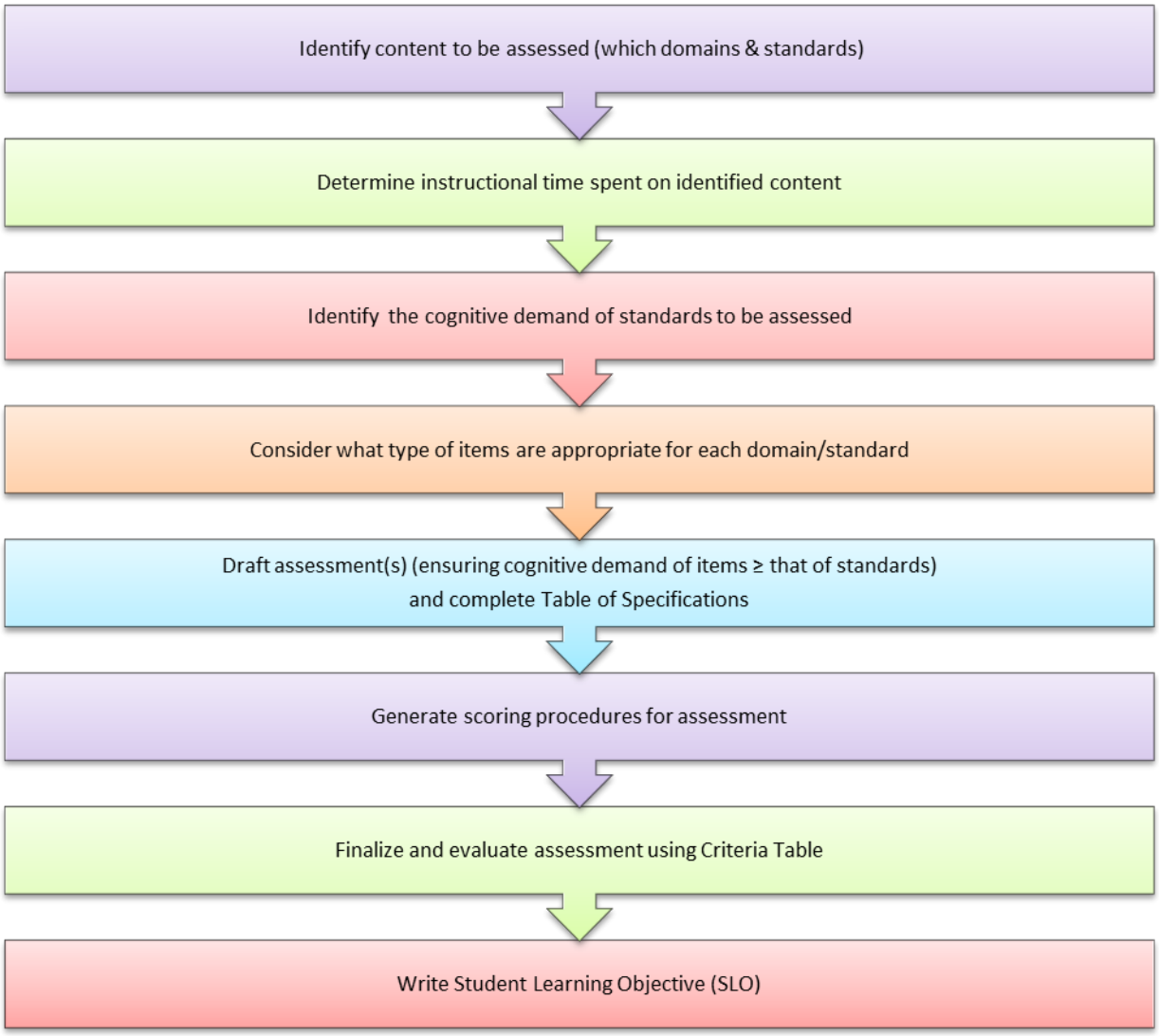
- Subject area experts
- Exemplary teachers
- Personnel with assessment design expertise
- Personnel with access to district and school data
- College/University content personnel
- RESA personnel or School Improvement personnel

2. Assessment Development

District/regional assessment teams need to have proficiency in:

- Aligning measures with course standards using the Aligning Curriculum and Assessment work tool.
- Completing or evaluating a measurement tool using the SLO Table of Specifications and the SLO Assessment Criteria Table (Described in Overview of the Measurement Cycle.)
- Assessing cognitive demand for each standard and measurement item
- Assessing the validity and reliability of the measurement items and measurement tool as a whole
- Assessing the measurment construction characteristics
- Understanding of the measurement cycle as described below

3. Overview of Measurement Cycle



The GaDOE Teacher and Leader Effectiveness Department has constructed a SLO Criteria Table for the development and evaluation of quality SLO measures. This tool is designed to help district/regional assessment teams evaluate the quality of their locally developed measurement tools which are used for SLOs. SLOs developed by inter-district collaboration are deemed Georgia Public Domain SLOs and SLO measures. These Public Domain SLOs are developed by the GaDOE trained teams and are available for other districts to use or adapt. The following steps provide a guide for the work of the GaDOE trained district/regional SLO development teams:

a. Identify the over-arching standards in all SLO courses.
b. Analyze the standards and identify the Language of the Standard utilizing the Aligning Curriculum and Assessment tool.
c. If there are no standards, identify the big ideas.
d. Use the SLO Table of Specifications to code the cognitive demand levels (Webb's DOK – all CCGPS formative assessments and PARCC assessments are aligned to DOK). Indicate the cognitive level with a check mark and include the corresponding verb (e.g. Comprehension- describe)
e. Use the SLO Criteria Table for the development and evaluation of quality assessments to construct or appropriately level SLO measurement items.
f. Create performance tasks, checklists, and rubrics that require subjective judgments.
g. Administer the SLO measures and conduct an item analysis.
h. Use the SLO Criteria Table for the development and evaluation of quality measuress to determine the reliability and validity of the measurement items.
i. Reach consensus regarding items requiring revision or removal from the item bank.

4. Components of Measurement Cycle

Selecting and understanding standards
District teams must identify or review (if this step has already been done by the district/regional SLO team) the standards to be assessed by the SLO for each course. The three-day SLO development training conducted by the GaDOE is not sufficient time to develop pre and post-measures which comprehensively assess all course standards. Initially choosing over- arching standards may help district teams focus on developing quality measures instead of developing extensive measurement tools.

The content of pre and post measures is driven by selected standards, the level(s) of cognitive demand required by the standards, and the emphasis/time devoted to the instruction of the selected standards. The first step in creating or evaluating valid and reliable measures is to analyze the standards which the measures are assessing. Analyzing the standard(s) means to really understand the standard(s) by determining the content and skills needed to achieve the standard(s) **and** the level of cognitive demand required for learning the standard(s). The GaDOE developed a tool to facilitate this process called the Aligning Curriculum and Assessment tool.

Content and skills: Each standard should be examined to determine the knowledge or content students are expected to acquire or demonstrate such as vocabulary, critical details, definitions, key facts, concepts, laws and formulas, sequence, and timelines. Likewise, the skills or behaviors which students are expected to apply or use to achieve the standard should be examined. Such skills may include listening, speaking, writing, thinking skills (e.g., compare, infer, analyze), research skills (e.g. inquire, investigate), and study skills (e.g. note-taking, outlining).

Cognitive demand: The level of cognitive demand is the expected level of thinking when engaged with the content. Determining the level of cognitive demand ensures that not only does the SLO focus on the subject matter/content, but it also provides parameters within which students can use the content in ways dictated by the standards.

There are various taxonomies that can be utilized by district/regional teams to determine cognitive demand. For the purposes of SLO assessments, the GaDOE is encouraging the use of Webb's Depth of Knowledge (DOK) since all CCGPS formative assessments and PARCC assessments are aligned to DOK. See Appendix E. However, district/regional teams may use any taxonomy which provides an adequate framework for complexity.

Creation of table of specifications

The SLO Table of Specifications (TOS) for measurement design and evaluation is used to align the standards, content, cognitive demand, and emphasis to the measures. A TOS is analogous to using a blueprint to build a house.

The Table of Specifications also includes types of measurement tools and measurement items which most effectively ascertain the students' knowledge and skills required by the selected standards.

Types of measurement items might include the following:

- Fill-in-the-blank
- Short answer
- Project(s)
- Essay
- Performance/demonstration
- Multiple choice

Validity of measurement tools and measurement items

Validity is the most important consideration in measurement design and evaluation of SLO measures. A valid measure assesses what it is intended to measure. Validity also refers to the level of confidence and trust in the judgments that educators can make about student learning as a result the measure. Validity is not an absolute characteristic; instead it is a matter of degree.

As district/regional assessment teams gain proficiency in measurement design and evaluation, they will be able to recommend ways to increase the degree of validity of regionally/locally developed measures. Conscientious use of the SLO Table of Specifications by a team of educators is one of the best methods for increasing and judging the validity of measurement tools.

Reliability of SLO measures

Reliability refers to the consistency of a measure. A test is considered reliable if the same results are obtained in a predictable manner over time and /or multiple administrations. For example, if a 150 pound person steps on a set of scales ten times, but gets results of 75, 100, 25, etc., then the scale does not produce reliable results. Rarely is an measurement tool perfectly reliable. As in the case of validity, reliability is a matter of degree. The goal is to design measures that are increasingly reliable.

An important concept which influences reliability is error in measurement. As stated in the book, *Teacher-made Assessments* (Grant and Gareis, p.41, 2008):

> Error in an assessment is when an assessment item does not adequately distinguish between the student who has truly mastered the intended learning outcome and the student who has not. For example, when a student gets a question correct, not because she knows it but because of something other than knowledge of or skill in the intended learning being assess, assessment error has occurred. Similarly, if a student misses a question for some reason other than a lack of knowledge or skill, then error has occurred.

Grant and Gareis suggest the following steps to improve a measure's reliability:

1. For non-performance tasks, include three or more test questions or items for each core objective/standard to reduce the unintended effects of error on the assessment results.
2. Review and proofread individual test questions, prompts, and directions for systematic error, including grammatical or mechanical mistakes, cultural bias, lack of clarity, etc.
3. Clarify and verify grading criteria for the test, including rubrics. Ensure intra-rater and inter-rater reliability for establishing scoring protocols and training (p. 46).

Creation of SLO measure(s)

Teachers and other team members should decide on the types and number of measurement items that will comprise the pre and post measures. A reasonable amount of class time should be allotted for pre and post-measures, typically one-two class periods. They may also create performance tasks, checklists, and rubrics that require subjective judgments. The SLO Criteria Table, as well as other sources can provide guidance with developing and evaluating measurement items.

Post item analysis

After measurements are administered, the SLO development team should reconvene to conduct an item analysis of the results. Items should be critiqued based on student results to determine if revision or removal of items is needed. Use the SLO Criteria Table for

the development and evaluation of quality measures to determine the reliability and validity of the measurement items. Consensus should be reached by SLO development teams regarding the revision items or their removal from the item bank.

Data analysis

This is the most important step of measurments in terms of instruction. Teachers of the same course should analyze the data to determine necessary instructional modifications to plans and reasonable next instructional steps based on student performance.

Integrity of SLO process and results

Opportunities to misrepresent student data or inappropriate interactions with students to affect pre and post-assessment results may be minimized by:

1- The use of signed assurances (See Appendix A)
2- On-going, systematic triangulation of formal and informal data by administrators/evaluators (observations, report card grades, tests, walk-throughs, documentation of teacher work). SLO data should be somewhat consistent with other student data.
3- Collaborative planning of groups of teachers around SLOs results/implementation
4- Utilization of Georgia Public Domain SLOs
5- Use of electronic item bank (under development)
6- Use of interchangeable passages, scenarios, numbers, etc. in assessment items
7- Increased use of performance tasks
8- Checking for inter-rater reliability of ratings; employ the use of sampling to ensure consistency of raters

SLO Approval

Once districts have completed their SLOs, the SLOs should be submitted to the GaDOE for approval. The SLO Approval Rubric criteria are used to determine whether each SLO will be sent back to the district for revision or whether it will be approved. Approved SLOs are posted on the RT3 SharePoint site for other districts to view.

How to Develop Student Learning Objectives

District SLO Leadership

In that SLO work has the potential to have a dramatic impact on student achievement and because that it impacts teacher and leader evaluations, strong instructional leadership of the SLO process is vital.

District SLO Team(s)

It is critically important for districts or groups of districts to form teams that have the needed expertise for designing district SLOs. Those team members might include the following:

- Subject area experts
- Exemplary teachers
- Personnel with assessment design expertise
- Personnel that has access to district and school data

The district team(s) assesses the needs of students by studying relevant data. While district student needs/data should be examined, individual school data are also important considerations in developing district SLOs. District/school trend data may also be examined.

Required Documentation for SLO Teams

Each team trained in the development of SLOs will be required to submit the District SLO Form. Below is a section-by-section description of the required District SLO form. (Directions for completing the form are in blue print.)

1. **The Standards**: Determine which standards are worthy of the students' and teachers' focus for the given instructional period (typically a school year or semester). List the standard reference number and a brief description of the standard in section 1.

   Based on the district/school data and needs assessment, district team(s) should determine appropriate state and national standards that will provide the basis for SLO development. Alignment of the SLO to standards is not only required merely for reference but to ensure validity. District-selected standards should warrant the year-long or course-long focus of the students and teachers and should be rigorous, measureable, and deepen and extend knowledge for all students in the class/group/course. It is up to the district to determine whether all standards are covered or if 5-15 over-arching standards are selected to determine teacher effectiveness.

Teams should consider the following questions when selecting standards:

- Do these standards focus on content and/or skills that capture the majority of the instructional period?
- Do these standards provide students with essential knowledge and skills that are necessary for success in the next level of instruction or next grade level?

Content and skills: Each standard should be examined to determine the knowledge or content students are expected to acquire or demonstrate such as vocabulary, critical details, definitions, key facts, concepts, laws and formulas, sequence and timelines. Likewise, the skills or behaviors which students are expected to apply or use to achieve the standard should be examined. Such skills may include listening, speaking, writing, thinking skills (e.g., compare, infer, analyze), research skills (e.g. inquire, investigate), and study skills (e.g. note-taking, outlining).

Cognitive demand: The level of cognitive demand is the expected level of thinking when engaged with the content. Determining the level of cognitive demand ensures that not only does the SLO focus on the subject matter/content, but it also provides parameters within which students can use the content in ways dictated by the standards.

2. Pre and Post Measures

A brief description of the pre and post SLO measures should be provided in section 2.

The quality of an SLO depends on the quality of the measures used to determine student growth. The validity of an measurement tool is, to a large degree, dependent on how well the chosen measurement tool assesses the students' learning of the determined standards. Teams should consider the following validity questions when selecting or designing SLO measures:

- Does the format and content of the pre measure allow students to demonstrate their current fundamental and/or background knowledge needed for this course?
- Do the results of the pre measure readily inform the teacher's instructional practice?
- Does the measure show what students should know and should be able to do at the end of this course? Does it measure the specified standards?
- Does the format and content of the post measure allow students to demonstrate their learning of the standards?

Districts should explore current formative and summative measures that they may already have to determine if those measures could be used or adapted as valid SLO measures. A list and description of commercial assessments which were used by RT3 pilot districts in Phase 1 may be obtained by contacting the GaDOE Teacher and Leader Department. Commercial assessments are selected, purchased, and used at the district's discretion. The GaDOE does not recommend any particular assessment or provide any such

endorsements. Districts must adhere to the guidance provided when using commercial measures for SLO purposes. These measures should be used according to the manufacturer's or designer's requirements for administration, fidelity of implementation, and limits of interpretation.

The information provided in The Foundation of Quality SLOs is designed to aid districts or groups of districts in evaluating the quality of current measures or in designing their own SLO measures.

Selected measures should determine student growth. Quality measues not only provide a pre and post score/result but should also be used to drive the teachers' instruction between the pre and post measurement results. Quality measures provide students and teachers with clear expectations and pictures/examples of quality work.

Measurement results, particularly in reference to standardized tests results, must be reported within the SLO cycle which ends on May 15.

If and when districts alter/adapt a Georgia Public Domain SLO or develop their own measure, a Table of Specifications and SLO Criteria Table should be submitted with the district SLO along with the locally developed measure(s).

3. Baseline or Trend Data

Baseline data, previous data, or data trends are the linchpin of the SLO since they provide the basis for the SLO growth targets and tiers. Before writing SLO growth targets, districts should analyze their pre measures data from the selected SLO subjects. Historical or trend data may also be examined. These data may include any of the following:

a. Formative Assessments based on the SLO's standards
b. Benchmark tests which focus on SLO's standards
c. Unit tests from course that assess SLO's standards
d. Grades from SLO course's performance based tasks
e. Student transiency rate for school system (High? Low?)
f. Pass/Fail Rate for SLO course for last two years
g. Percentage of students receiving As, Bs, Cs, Ds, and Fs in course
h. State-mandated standardized tests based on SLO's standards (EOCT, CRCT, GHSGT, etc.)

4. Design and Construction of SLO Statements

SLOs should describe observable behavior and/or measurable results that would occur when an objective is achieved. The acronym SMART (Figure 3) is a critical way to self-assess an objective's feasibility and worth.

Figure 3: SMART Acronym for Developing Student Learning Objectives

Specific:	The objective is focused, for example, by content standards; by learners' needs.
Measurable:	An appropriate instrument/measure is selected to assess the objective.
Appropriate:	The objective is within the teacher's control to effect change and is a worthwhile focus for the students' academic year.
Realistic:	The objective is feasible for the teacher.
Time limited:	The objective is contained within a single school year or instructional period.

- Specified components of the SLO include the following:

1. Course Name
2. Pre and post measurment administration dates or windows
3. Skill or content area to be measured
4. Name of measure
5. Level or scores or range of scores; proficiency levels; exceeding proficiency levels
6. Expected quantity of growth for each student taught by the teacher
7. Tiered targets, if applicable

Considerations when writing SLOs

- SLOs must be growth objectives not achievement objectives. SLOs should be designed and written so that individual student growth between the pre-measure and the post-measure can be determined.

  In contrast, achievement objectives would specify a percentage or number of students who would attain a specified level. Growth objectives specify the growth target for all students. Therefore, 100% of the students in the course will be included in the SLO and its growth targets.

  The SLO growth target(s) for students should be the growth needed to achieve the students' particular academic goals and/or the recommended achievement levels designed to meet the academic needs of the student population connected to the SLO. In order to ensure that all students perform well in terms of growth, SLOs may employ target tiers. Setting one growth goal and expecting all students in the district to meet the same level of

growth may not be realistic especially if there is high variability in student performance levels. Therefore, target tiers may be used to determine expected amounts of growth based on the variability of skills and knowledge students have upon beginning the course subject. As a reminder SLOs should also include the highest performers in the district population. This can be done by adding a "maintain" statement and including an additional task for advanced learners.

Growth targets should be set according to what a year's worth of growth would be for the course. When students attain targets, they should either make significant gains in "catching up" if they are below grade/course expectations or should demonstrate the knowledge and skills to be successful in their future course work.

It is also recommended that district staff speak with teachers who have historically produced high student achievement to determine acceptable and realistic student growth in the subject of the SLO. Before writing the SLO, the district should understand its student population and the needs of the individuals addressed in the SLO. However, it is predicted that districts will set more effective targets as they monitor the data collected on SLOs over time.

- SLOs should be written so that teachers implementing the SLOs are very clear on what to do and when to do it. It is advisable that prior to submitting SLOs to the GaDOE for approval, the SLO has been read by a teacher(s) who has not been a part of the SLO development. Such a "cold read" or novice read of the SLO can provide the district team with valuable feedback and may save the district team and the local school staff time with the implementation of the SLO.

- Well-designed and rigorous targets in SLOs will increase student achievement which will support the attainment of school and district goals.

- When asking teachers for a "cold read' on a proposed SLO, it will be helpful to provide them with examples of other SLOs which have been through the review process. Figures 5 and 9 provide appropriate examples.

- The language of the assessment(s) should be reflected in the SLO. For example, if the SLO measure uses performance levels, a score on a 100-point test, etc. congruent terminology should be used in the SLO.

- SLOs should be written so that local school evaluators can successfully use the SLO Evaluation Rubric (See Figure 4) to determine if the teacher's students met the SLO.

Figure 4: Student Learning Objective Evaluation Rubric

Exemplary (3 pts)	Proficient (2 pts)	Developing/Needs Improvement (1 pt)	Ineffective (0 pts)
Greater than or equal to 90% of students met or exceeded their Student Learning Objective target and 50% or more of these students exceeded their Student Learning Objective target.	Greater than or equal to 80% of students met or exceeded their Student Learning Objective target.	Greater than or equal to 50% and less than 80% of students met or exceeded their Student Learning Objective target.	Less than 50% of students met or exceeded their Student Learning Objective target.
☐	☐	☐	☐
Comments:			

Sample SLOs

While each SLO must have the components specified above, the SLO itself may vary. The ranges used, the reporting score, the level of proficiency expected, and the use of multiple assessment measures are all factors that affect the design of the SLO. The samples provided here are meant to demonstrate the required components for the SLO. All SLOs may not match these samples, but these models demonstrate the basic structure that the SLO will take. Each example demonstrates a different approach to measuring growth (Figure 5 uses tiered targets while Figure 7 uses a uniform approach, but each addresses the progress of all students.) Targets for meeting and exceeding are needed in each SLO.

Figure 5: Sample SLO with Tiered Targets

Sample SLO for Grade 9-12 Environmental Science (26.06110)

From the fall assessment window (September 1-15, 2012) to the spring assessment window (April 1 - April 30, 2013), 100% of ____ County's Environmental Science students will improve their pre to post measurment scores as measured by the _____ County Environmental Science Benchmark Assessment. Students will increase from their pre-measuement score ranges to the post-measurement score ranges as follows:

- Students scoring form 0-25 on the pre measure will increase to the post measure by 50 to 60 points;
- Students scoring from 26-40 will increase by 40 to 50 points;
- Students scoring from 41-65 will increase by 30 to 40 points;
- Students scoring from 66-100 will increase by 25-35 points or score >93.
- Students scoring above the target range by 10 points will exceed their targets. Students scoring above 91 may exceed their target by completing an applicable project/demonstration that has been approved by the teacher's evaluator.

Figure 6: Guide for Constructing SLO with Targeted Tiers

From _____(date)_____ to _____(date)_____ , 100% of _____(student group)_____will improve their _____(skill/content area)_____ as measured by the _____(assessment measure)_____.

Students will increase from their pre-assessment scores to the post-assessment scores as follows:

- **Tier 1**: Students scoring on the _____(*pre-assessment)_____ at ___(level/score or range of scores)____ will improve/progress from ____(level/score or range of scores)____ to _____.
- **Tier 2:** Students scoring on the _____(*pre-assessment)_____ at ___(level/score or range of scores)____ will improve/progress from ____(level/score or range of scores)____ to _____.
- **Tier 3:** Students scoring on the _____(*pre-assessment)_____ at ___(level/score or range of scores)____ will improve/progress from ____(level/score or range of scores)____ to _____.
- Students score above the target range by ___ points will exceed their targets. Students scoring above ___ may exceed their target by completing an applicable project/demonstration that has been approved by the teacher's evaluator.

The following sample and guide are used when a uniform growth target is incorporated into the SLO. A uniform growth target simply means that the expectation is that all students will demonstrate equal growth as determined by the assessment. In the SLO sample in Figure 7, all students are expected to increase by one or more levels from the pre assessment to the post assessment.

Figure 7: Sample SLO with Uniform Growth Target

Sample SLO for Grade 6 Intermediate Chorus (54.0131)

From September 1-15, 2012 to March 15- April 1, 2013, 100% of grade 6 chorus students will demonstrate an increase of 1 or more levels from the pre to the post assessment as measured by the regionally developed four-level rubric for sight-singing composition and sight-singing performance. Students who demonstrate an increase of 2 or more levels have exceeded their targets.

The common performance based four-level rubric assessment titled "Sight Singing Assessment" was developed by representatives from the local RESA and its districts.

Figure 8: Guide for Constructing SLO with Uniform Growth Target

From _____(date)_____ to _____(date)_____, 100% of _____(student group)_____ will improve their _____(skill/content area)_____ as measured by the _____(assessment measure)_____. Students will demonstrate progress by increasing their pre-assessment score/level on the _____(name of post-assessment)_____ by a minimum of _____(quantity of increase of numerical points, percentage increase, or rubric level)_____. Students who demonstrate an increase of ___ or more levels have exceeded their targets.

The following sample and guide are used when growth targets are unique for each student depending on their pre assessment score. In the sample in Figure 9, growth is ½ the difference between the pre assessment score and 100.

Figure 9: Sample SLO with Individualized Growth Target

Sample SLO for Grade 9-12 World History (45.0830038)

From September 1-15, 2012 to April 1-30, 2013, all students enrolled in World History will demonstrate measureable growth from the pre measure score to their post measure scores as measured by X District's pre measure and post measure as follows:

The minimum expectation for individual student growth is based on the formula which requires students to grow by 70% of their potential growth.

- Pre measure score + [(100- pre score) * .7] = target score
- Students who score 10 points above their target score have exceeded their target.

*Example using 40 on a pre measure: 40 + (100-40) *.7*
*40 + (60 * .7)*
40 + 42
82 is the target for post-measure
A score of 92 would indicate exceeding target.

Figure 10: SLO Guide for Individualized Growth Target

From _____(date)_____ to _____(date)_____ , all _____(student group)_____ enrolled in _____(class/subject)_____ will demonstrate measureable growth from their pretest score to their posttest score as measured by the _____(assessment measure)_____ and the following criteria:

- Pre measure score + [(100- pre score) *__ = target score
- Students who score __ points above their target score have exceeded their target.

5. Powerful Strategies to Attain SLO Targets

Section 5 of the District SLO Form is an optional portion of the form. Districts may decide to include suggested or required strategies for SLO attainment. This information may be very helpful for novice teachers or may be used by school leaders as they work to reinforce effective instructional strategies. The frequency of these strategies may also be included.

6. Mid-year or Mid-course Review

A description of the mid-year or mid-term review should be added to section 6.

A mid-year or mid-term review of student progress toward growth targets is required. The purpose of this review is for teachers to examine and share student progress with their evaluator. It is important to determine if students are on track to achieve growth targets and whether instructional adjustments or interventions are needed. The district may determine the format of the mid-year or mid-term review, may recommend/suggest specific mid-year or mid-term actions, or may leave this decision up to the school evaluator and/or teacher.

Teacher's role with SLOs

As stated earlier in this document, it is critical to include expert teachers and content specialists in the development of district SLOs. After districts have developed SLOs and received GaDOE approval of SLOs, the SLOs should be given to teachers who will administer the pre-measures. Pre and post-measures should be administered during the district determined administration windows. The GaDOE provides a suggested but not required Teacher SLO Form (See Appendix F). The purpose of this form or a similarly developed district form is to outline the meetings required to fulfill the guidelines for the SLO process. Evaluators should meet a minimum of three times each year with their teachers regarding progress toward SLO targets. This form should include an opportunity to address each of the following components of the SLO process:

- Purpose of SLOs
- Effective teaching practices
- When pre and post measurement data are collected
- How pre and post measurement data are recorded
- Format for mid-year or mid-course review

Submission Process

1. Phase II: Districts must submit each SLO for GaDOE approval before local teachers begin implementation of their SLO plans. A separate District SLO Form should be completed for each district SLO and should be submitted to the GaDOE on or before August 2, 2012. SLOs should be submitted to the GaDOE by a person designated at the district level to review and sign-off on all submissions. Individual teachers or staff members should not submit SLOs for approval to the GaDOE. SLOs completed and reviewed at the district level should be submitted to the Teacher and Leader Effectiveness Program Manager at SLO@doe.k12.ga.us.

Approval Rubric

	Stage 3	Stage 2	Stage 1
	Exemplary SLO **Integrity of SLO Process is Increased** **(Stage 3 also include criteria for Stage 2)**	**Proficient SLO** **All Requirements Met**	**Developing SLO** **Needs Revisions**
Identifying Information		☐ State Course number and name provided ☐ District name and grade level/s provided	☐ No state course number ☐ No state course title ☐ No grade level/s provided ☐ No district name provided
Standards	☐ Selected standard are appropriate for teacher/student focus for the instructional period ☐ Selected standards are an important and overarching concept and approved by GaDOE-trained SLO development team	☐ Focused on content standards ☐ Standards are selected by collaborative team ☐ Brief description of standard(s) provided	☐ Too few standards are selected to adequately assess student knowledge in course ☐ In order for the large number of standards chosen to be assessed, the pre and post measures would be too lengthy. ☐ No brief description of the standard/s provided ☐ No collaboration is evident

Baseline or Trend Data	☐ Is based on specific, related district baseline or trend data and supports growth targets	☐ General baseline and/or trend data are provided Or ☐ Convincing rationale is provided	☐ No baseline data or rationale are provided to support the standard/s chosen
Measures for pre and post assessment	☐ Alignment between standards and assessment has been approved by GaDOE-trained SLO development team using the SLO Table of Specifications ☐ Utilizes externally developed, reliable and valid purchased measures Or ☐ Locally developed measures have been approved by GaDOE-trained SLO development team using the SLO Assessment Table of Specifications ☐ Paper/pencil or performance based measures are used as appropriate for the characteristics or standards of the non-tested subject	☐ Measurement tool is aligned with the standards ☐ It appears that an appropriate instrument/measure is selected to assess SLO ☐ Measurement is adequately described (for purchased assessments) Or ☐ Locally developed measures are submitted; SLO Assessment Table of Specifications has been accurately completed for locally/regionally developed measures. ☐ Pre and post measures are utilized by multiple teachers/schools	☐ Measurement tool is not aligned with standards ☐ Locally developed measures are not submitted ☐ Purchased measures are not described ☐ SLO Table of Specifications was not utilized in designing or evaluating locally/regionally developed measure(s) ☐ Table of Specifications does not accurately reflect measurement items/tasks

SLO Statement	☐ SLO is clear and coherent on first read ☐ Results of pre-measures can be used to drive instruction and not for the sole purpose of SLO data. ☐ Attainment of SLOs reinforces school and district student achievement goals ☐ Expected growth is rigorous, yet attainable during instructional period. Rigorous DOK items/tasks are noted in SLO Assessment Table of Specifications ☐ SLO was developed by content experts and practitioners ☐ Course proficiency is stated. ☐ Meets and exceeds targets are specified.	☐ SLO is clear and coherent ☐ Uses SMART criteria ☐ SLO appears to be feasible for teacher ☐ Teachers are able to align work directly to the district SLO ☐ Growth targets appear realistic and meet the needs of all students ☐ SLO is within teachers' control to effect change and appears to be a worthwhile focus for the instructional period ☐ Growth targets appear to be rigorous	☐ SLO is not clear and coherent ☐ Does not completely follow SMART criteria ☐ Attainment of SLO is outside teachers' influence ☐ Growth targets do not appear to be realistic ☐ Growth targets do not address the needs of all students ☐ Growth targets do not appear to be rigorous ☐ Growth targets not supported by baseline data
Time Bound	☐ Instrument(s) is used to measure student growth from beginning of instructional period to end of instructional period. ☐ District standardized time frames for administration of pre and post-measures have been determined and will be observed.	☐ SLO specifically states the instructional period.	☐ No instructional time period is listed ☐ Time period listed is outside the SLO target calendar
District approved	☐ District establishes and provides procedures/guidance/requirements for usage of district SLOs and accompanying measures. ☐ Rigor of SLO is comparable to the rigor of "tested" subjects as determined by analysis of district data	☐ District approves/recommends this SLO for teachers at the designated grade level(s) and in these subject area(s) ☐ Signature of Superintendent or designee	☐ SLO not submitted via the proper procedures ☐ Required signature of Superintendent or designee is not provided ☐ Locally/regionally developed measures have

	☐ Time for post-analysis of student data and SLO revision is predetermined and scheduled ☐ All locally/regionally developed measures have been evaluated using the SLO Criteria Table.	is provided with SLO ☐ Locally/regionally developed assessments have been evaluated using the SLO Criteria Table.	not been evaluated using the SLO Criteria Table.

Timeline

Figure 11 : Student Learning Objectives Timeline

February - June 2012	• Districts considers needs of students, demands of grade level standards, and baseline data and creates SLOs for Phase II, including pre- and post measures s • Collaborating districts participate in three-day training on SLO development provided by the GaDOE
May 2012	• District reviews the end of the year data on Phase I SLOs and revises as necessary. • Train-the-Trainer SLO Assessment Training for RT3 and new districts
June 2012	• Districts begin work on SLOs for Phase III.
July 2, 2012	• The district submits Phase II SLOs to the GaDOE for review and approval
August - September 2012	• District shares revised SLOs for Phase I and SLOs for Phase II with teachers.
September - October 2012	• Teachers use district SLOs and administer pre measures. • Teachers complete a spreadsheet with student pre scores; analyze the class/group data. • Teachers complete Teacher SLO Form, and implement teaching strategies. • Teachers meet with their evaluators to finalize their SLO plan.
August 2012 – March 2013	• Teachers implement teaching strategies and monitor student progress toward attainment of SLO(s).
December - January	• Teachers complete mid-year or mid-courses review for SLO(s) • Teachers meet with their evaluator to discuss mid-year progress on SLO(s) • Evaluators determine if Professional Growth Plan is needed for SLO attainment
April 1, 2013	• Teachers administer post-measure.
April 15, 2013	• Teachers submit class/group data to building level evaluator. Evaluator completes SLO Evaluation Rubric and submits SLO information (TBD) to the GaDOE.
May 2013	• GaDOE calculates TEM using all components of the TKES.

They will assign an end-of-year rating using an evaluation rubric with the following levels:

Exemplary, Proficient, Developing/Needs Improvement, and *Ineffective* as shown in Figure 12.

Figure 12: Student Learning Objective Evaluation Rubric

Exemplary (3 pts)	**Proficient (2 pts)**	**Developing/Needs Improvement (1 pt)**	**Ineffective (0 pts)**
Greater than or equal to 90% of students met or exceeded their Student Learning Objective target and 50% or more of these students exceeded their Student Learning Objective target.	Greater than or equal to 80% of students met or exceeded their Student Learning Objective target.	Greater than or equal to 50% and less than 80% of students met or exceeded their Student Learning Objective target.	Less than 50% of students met or exceeded their Student Learning Objective target.
☐	☐	☐	☐
Comments:			

The Student Learning Objective score then will be scaled so that it counts for a specific amount of the overall Teacher Effectiveness Measure.

Appendix A: Teacher Assurances

As related to Teacher and Leader Keys Evaluation Systems

Any action that compromises test/assessment security, leads to the invalidation of an individual student's or a group of students' assessment scores, or interferes with the components of the Teacher and Leader Keys Effectiveness System will be viewed by the Georgia Department of Education as inappropriate. In order to maintain the fidelity of TKES and LKES all teachers and administrators must adhere to following assurances. The list is not exhaustive. Any concerns about test/assessment security or proper implementation of the TKES and LKES components must be reported to the GaDOE immediately.

	Initials	Assurances
1.		Teachers have been trained in the appropriate use of all components of the Teacher and Leader Keys Evaluation System.
2.		Students are prepared for the post assessment by the alignment of curriculum and instruction to the district content standards.
3.		Students are appropriately informed about the assessment prior to its administration, including its purposes, uses, consequences, and how the assessment information will be judged or scored. However, communication as to how the results will be used for an individual teacher's evaluation is prohibited.
4.		Students are encouraged to put forth optimal effort based on the purpose of the assessment. Results of pre assessments will be appropriately shared with students.
5.		An appropriate testing environment is provided.
6.		All eligible students are assessed.
7.		All reasonable and allowable accommodations for the administration of the assessment are provided to persons with disabilities or special needs.
8.		Appropriate security precautions are taken before, during, and after the administration of the assessment and the survey.
9.		Reasonable quality control procedures are maintained before, during, and after administration and scoring of the assessment and the survey.
10.		No part of the assessment or survey is revealed to students prior to the administration.
11.		Distribution of assessment and survey materials occurs immediately prior to administration.
12.		The assessment/survey occurs during the specified schedule of administration.
13.		The specified schedule of administration provides for make-up opportunities for students absent during the administration of the assessment or the survey.

14.		Teacher actions before, during, or after assessments should not give any particular student or class of students an unfair advantage over others.
15.		All standards within the course are taught with the appropriate level of time/emphasis. No course standards are taught to the exclusion of other standards for the sole purpose of SLO attainment.
16.		Pre and post assessment will administered within the district-designated pre and post assessment windows. Assessments be scored and recorded (as determined by evaluator) in a timely manner.
17.		A whistle blower procedure has been established and communicated to all staff. Staffs are responsible for reporting any breaches of assurances.
18.		Student assessments and all scoring documents are maintained according to the district's records retention schedule.

It is a breach of proper assessment and survey administration if anyone performs any of the following:

19.		Coaches examinees during testing, performance assessments, or surveys or alters or interferes with examinees' responses in any way;
20.		Gives examinees access to assessment or survey questions or prompts prior to administration;
21.		Copies, reproduces, or uses in any manner inconsistent with test security regulations including all or any portion of test booklets, assessments, or surveys;
22.		Reads or reviews test questions before, during or after testing (unless specified in the IEP, IAP or ELL/TPP);
23.		Fails to follow security regulations for distribution and return of secure test materials as directed, or fails to account for all secure test materials before, during and after testing (NOTE: Lost test booklets constitute a breach of test security and will result in a referral to the PSC.);
24.		Uses or handles secure assessments, prompts, survey questions, and/or answer documents for any purpose other than examination;
25.		Fails to follow administration directions for the assessment or survey;
26.		Erases, marks answers, or alters responses on an answer document or interferes with student as they respond to computerized questions, etc.
27.		Participates in, directs, assists, counsels, encourages or fails to report any of the above listed acts.

Failure to safeguard assessment and survey materials or to comply with proper administration procedures could adversely affect an individual's certification status.

_____
Teacher Name

_____ _____
Teacher signature Date

Appendix C: Aligning Curriculum and SLO Measurement Tool

Session I – Aligning Curriculum & Measures
Ensuring Content & Construct Validity

Contact the Department of Teacher and Leader Effectiveness at the GaDOE for the latest version of this document.

Appendix B: Table of Specifications

Contact the Department of Teacher and Leader Effectiveness at the GaDOE for the latest version of this document.

Appendix E: Depth of Knowledge

Depth-of-Knowledge Definitions
Norman L. Webb
University of Wisconsin-Madison
Wisconsin Center for Education Research
January 16, 2008

Reading DOK Levels

In language arts, four DOK levels were used to judge both reading and writing objectives and assessment tasks. The reading levels are based on Valencia and Wixson (2000, pp. 909-935). The writing levels were developed by Marshá Horton, Sharon O'Neal, and Phoebe Winter.

Reading Level 1. Level 1 requires students to receive or recite facts or to use simple skills or abilities. Oral reading that does not include analysis of the text, as well as basic comprehension of a text, is included. Items require only a shallow understanding of the text presented and often consist of verbatim recall from text, slight paraphrasing of specific details from the text, or simple understanding of a single word or phrase. Some examples that represent, but do not constitute all of, Level 1 performance are:

Support ideas by reference to verbatim or only slightly paraphrased details from the text.
Use a dictionary to find the meanings of words.
Recognize figurative language in a reading passage.

Reading Level 2. Level 2 includes the engagement of some mental processing beyond recalling or reproducing a response; it requires both comprehension and subsequent processing of text or portions of text. Inter-sentence analysis of inference is required. Some important concepts are covered, but not in a complex way. Standards and items at this level may include words such as summarize, interpret, infer, classify, organize, collect, display, compare, and determine whether fact or opinion. Literal main ideas are stressed. A Level 2 assessment item may require students to apply skills and concepts that are covered in Level 1. However, items require closer understanding of text, possibly through the item's paraphrasing of both the question and the answer. Some examples that represent, but do not constitute all of, Level 2 performance are:

Use context cues to identify the meaning of unfamiliar words, phrases, and expressions that could otherwise have multiple meanings.
Predict a logical outcome based on information in a reading selection.
Identify and summarize the major events in a narrative.

Reading Level 3. Deep knowledge becomes a greater focus at Level 3. Students are encouraged to go beyond the text; however, they are still required to show understanding of the ideas in the text. Students may be encouraged to explain, generalize, or connect ideas. Standards and items at Level 3 involve reasoning and planning. Students must be able to support their thinking. Items may involve abstract theme identification, inference across an entire passage, or students' application of prior knowledge. Items may also involve more superficial connections between texts. Some examples that represent, but do not constitute all of, Level 3 performance are:

Explain or recognize how the author's purpose affects the interpretation of a reading selection.
Summarize information from multiple sources to address a specific topic.
Analyze and describe the characteristics of various types of literature.

Reading Level 4. Higher-order thinking is central and knowledge is deep at Level 4. The standard or assessment item at this level will probably be an extended activity, with extended time provided for completing it. The extended time period is not a distinguishing factor if the required work is only repetitive and does not require the application of significant conceptual understanding and higher-order thinking. Students take information from at least one passage of a text and are asked to apply this information to a new task. They may also be asked to develop hypotheses and perform complex analyses of the connections among texts. Some examples that represent, but do not constitute all of, Level 4 performance are:

Analyze and synthesize information from multiple sources.
Examine and explain alternative perspectives across a variety of sources.
Describe and illustrate how common themes are found across texts from different cultures.

Writing DOK Levels

Writing Level 1. Level 1 requires the student to write or recite simple facts. The focus of this writing or recitation is not on complex synthesis or analysis, but on basic ideas. The students are asked to list ideas or words, as in a brainstorming activity, prior to written composition; are engaged in a simple spelling or vocabulary assessment; or are asked to write simple sentences. Students are expected to write, speak, and edit using the conventions of Standard English. This includes using appropriate grammar, punctuation, capitalization, and spelling. Students demonstrate a basic understanding and appropriate use of such reference materials as a dictionary, thesaurus, or Web site. Some examples that represent, but do not constitute all of, Level 1 performance are:

Use punctuation marks correctly.
Identify Standard English grammatical structures, including the correct use of verb tenses.

Writing Level 2. Level 2 requires some mental processing. At this level, students are engaged in first-draft writing or brief extemporaneous speaking for a limited number of purposes and audiences. Students are expected to begin connecting ideas, using a simple organizational structure. For example, students may be engaged in note-taking, outlining, or simple summaries. Text may be limited to one paragraph. Some examples that represent, but do not constitute all of, Level 2 performance are:

Construct or edit compound or complex sentences, with attention to correct use of phrases and clauses.
Use simple organizational strategies to structure written work.
Write summaries that contain the main idea of the reading selection and pertinent details.

Writing Level 3. Level 3 requires some higher-level mental processing. Students are engaged in developing compositions that include multiple paragraphs. These compositions may include complex sentence structure and may demonstrate some synthesis and analysis. Students show awareness of their audience and purpose through focus, organization, and the use of appropriate compositional elements. The use of appropriate compositional elements includes such things as addressing chronological order in a narrative, or including supporting facts and details in an informational report. At this stage, students are engaged in editing and revising to improve the quality of the composition. Some examples that represent, but do not constitute all of, Level 3 performance are:

Support ideas with details and examples.
Use voice appropriate to the purpose and audience.
Edit writing to produce a logical progression of ideas.

Writing Level 4. Higher-level thinking is central to Level 4. The standard at this level is a multi-paragraph composition that demonstrates the ability to synthesize and analyze complex ideas or themes. There is evidence of a deep awareness of purpose and audience. For example, informational papers include hypotheses and supporting evidence. Students are expected to create compositions that demonstrate a distinct voice and that stimulate the reader or listener to consider new perspectives on the addressed ideas and themes. An example that represents, but does not constitute all of, Level 4 performance is:

Write an analysis of two selections, identifying the common theme and generating a purpose that is appropriate for both.

Georgia Department of Education
Student Learning Objectives Manual
Depth-of-Knowledge Definitions
University of Wisconsin-Madison
(Permission to replicate by author, Dr. Norman L. Webb)

Mathematics DOK Levels

Level 1 (Recall) includes the recall of information such as a fact, definition, term, or a simple procedure, as well as performing a simple algorithm or applying a formula. That is, in mathematics, a one-step, well-defined, and straight algorithmic procedure should be included at this lowest level. Other key words that signify Level 1 include "identify," "recall," "recognize," "use," and "measure." Verbs such as "describe" and "explain" could be classified at different levels, depending on what is to be described and explained.

Level 2 (Skill/Concept) includes the engagement of some mental processing beyond an habitual response. A Level 2 assessment item requires students to make some decisions as to how to approach the problem or activity, whereas Level 1 requires students to demonstrate a rote response, perform a well-known algorithm, follow a set procedure (like a recipe), or perform a clearly defined series of steps. Keywords that generally distinguish a Level 2 item include "classify," "organize," "estimate," "make observations," "collect and display data," and "compare data." These actions imply more than one step. For example, to compare data requires first identifying characteristics of objects or phenomena and then grouping or ordering the objects. Some action verbs, such as "explain," "describe," or "interpret," could be classified at different levels depending on the object of the action. For example, interpreting information from a simple graph, or reading information from the graph, also are at Level 2. Interpreting information from a complex graph that requires some decisions on what features of the graph need to be considered and how information from the graph can be aggregated is at Level 3. Level 2 activities are not limited only to number skills, but may involve visualization skills and probability skills. Other Level 2 activities include noticing or describing non-trivial patterns, explaining the purpose and use of experimental procedures; carrying out experimental procedures; making observations and collecting data; classifying, organizing, and comparing data; and organizing and displaying data in tables, graphs, and charts.

Level 3 (Strategic Thinking) requires reasoning, planning, using evidence, and a higher level of thinking than the previous two levels. In most instances, requiring students to explain their thinking is at Level 3. Activities that require students to make conjectures are also at this level. The cognitive demands at Level 3 are complex and abstract. The complexity does not result from the fact that there are multiple answers, a possibility for both Levels 1 and 2, but because the task requires more demanding reasoning. An activity, however, that has more than one possible answer and requires students to justify the response they give would most likely be at Level 3. Other Level 3 activities include drawing conclusions from observations; citing evidence and developing a logical argument for concepts; explaining phenomena in terms of concepts; and deciding which concepts to apply in order to solve a complex problem.

Level 4 (Extended Thinking) requires complex reasoning, planning, developing, and thinking, most likely over an extended period of time. The extended time period is not a distinguishing factor if the required work is only repetitive and does not require applying significant conceptual understanding and higher-order thinking. For example, if a student has to take the water temperature from a river each day for a month and then construct a graph, this would be classified as a Level 2. However, if the student is to conduct a river study that requires taking into consideration a number of variables, this would be a Level 4. At Level 4, the cognitive demands of the task should be high and the work should be very complex. Students should be required to make several connections—relate ideas *within* the content area or *among* content areas—and have to select one approach among many alternatives on how the situation should be solved, in order to be at this highest level. Level 4 activities include designing *and* conducting experiments and projects; developing and proving conjectures, making connections between a finding and related concepts and phenomena; combining and synthesizing ideas into new concepts; and critiquing experimental designs.

Levels of Depth-of-Knowledge for Science

Interpreting and assigning depth-of-knowledge levels to objectives both within standards and assessment items is an essential requirement of alignment analysis. Four levels of depth of knowledge are used for this analysis. Because the highest (fourth) DOK level is rare or even absent in most standardized assessments, reviewers usually will be making distinctions among DOK levels 1, 2 and 3. Please note that, in science, "knowledge" can refer both to content knowledge and knowledge of science processes. This meaning of knowledge is consistent with the National Science Education Standards (NSES), which terms "Science as Inquiry" as its first Content Standard. The science levels were developed with the help of Edward Britton and Gwen Pollock.

Level 1. Recall and Reproduction

Level 1 is the recall of information such as a fact, definition, term, or a simple procedure, as well as performing a simple science process or procedure. Level 1 only requires students to demonstrate a rote response, use a well-known formula, follow a set procedure (like a recipe), or perform a clearly defined series of steps. A "simple" procedure is well-defined and typically involves only one-step. Verbs such as "identify," "recall," "recognize," "use," "calculate," and "measure" generally represent cognitive work at the recall and reproduction level. Simple word problems that can be directly translated into and solved by a formula are considered Level 1. Verbs such as "describe" and "explain" could be classified at different DOK levels, depending on the complexity of what is to be described and explained.

A student answering a Level 1 item either knows the answer or does not: that is, the answer does not need to be "figured out" or "solved." In other words, if the knowledge necessary to answer an item automatically provides the answer to the item, then the item is at Level 1. If the knowledge necessary to answer the item does not automatically provide the answer, the item is at least at Level 2. Some examples that represent but do not constitute all of Level 1 performance are:

Recall or recognize a fact, term, or property.
Represent in words or diagrams a scientific concept or relationship.
Provide or recognize a standard scientific representation for simple phenomenon.
Perform a routine procedure such as measuring length.

Level 2. Skills and Concepts

Level 2 includes the engagement of some mental processing beyond recalling or reproducing a response. The content knowledge or process involved is more complex than in level 1. Items require students to make some decisions as to how to approach the question or problem.

Keywords that generally distinguish a Level 2 item include "classify," "organize," "estimate," "make observations," "collect and display data," and "compare data." These actions imply more than one step. For example, to compare data requires first identifying characteristics of the objects or phenomenon and then grouping or ordering the objects. Level 2 activities include making observations and collecting data; classifying, organizing, and comparing data; and organizing and displaying data in tables, graphs, and charts.

Some action verbs, such as "explain," "describe," or "interpret," could be classified at different DOK levels, depending on the complexity of the action. For example, interpreting information from a simple graph, requiring reading information from the graph, is a Level 2. An item that requires interpretation from a complex graph, such as making decisions regarding features of the graph that need to be considered and how information from the graph can be aggregated, is at Level 3. Some examples that represent, but do not constitute all of Level 2 performance, are:

Specify and explain the relationship between facts, terms, properties, or variables.
Describe and explain examples and non-examples of science concepts.
Select a procedure according to specified criteria and perform it.
Formulate a routine problem given data and conditions.
Organize, represent and interpret data.

Level 3. Strategic Thinking

Level 3 requires reasoning, planning, using evidence, and a higher level of thinking than the previous two levels. The cognitive demands at Level 3 are complex and abstract. The complexity does not result only from the fact that there could be multiple answers, a possibility for both Levels 1 and 2, but because the multi-step task requires more demanding reasoning. In most instances, requiring students to explain their thinking is at Level 3; requiring a very simple explanation or a word or two should be at Level 2. An activity that has more than one possible answer and requires students to justify the response they give would most likely be a Level 3. Experimental designs in Level 3 typically involve more than one dependent variable. Other Level 3 activities include drawing conclusions from observations; citing evidence and developing a logical argument for concepts; explaining phenomena in terms of concepts; and using concepts to solve non-routine problems. Some examples that represent, but do not constitute all of Level 3 performance, are:

Identify research questions and design investigations for a scientific problem.
Solve non-routine problems.
Develop a scientific model for a complex situation.
Form conclusions from experimental data.

Level 4. Extended Thinking

Tasks at Level 4 have high cognitive demands and are very complex. Students are required to make several connections—relate ideas *within* the content area or *among* content areas—and have to select or devise one approach among many alternatives on how the situation can be solved. Many on-demand assessment instruments will not include any assessment activities that could be classified as Level 4. However, standards, goals, and objectives can be stated in such a way as to expect students to perform extended thinking. "Develop generalizations of the results obtained and the strategies used and apply them to new problem situations," is an example of a Grade 8 objective that is a Level 4. Many, but not all, performance assessments and open-ended assessment activities requiring significant thought will be Level 4.

Level 4 requires complex reasoning, experimental design and planning, and probably will require an extended period of time either for the science investigation required by an objective, or for carrying out the multiple steps of an assessment item. However, the extended time period is not a distinguishing factor if the required work is only repetitive and does not require applying significant conceptual understanding and higher-order thinking. For example, if a student has to take the water temperature from a river each day for a month and then construct a graph, this would be classified as a Level 2 activity. However, if the student conducts a river study that requires taking into consideration a number of variables, this would be a Level 4. Some examples that represent but do not constitute all of a Level 4 performance are:

Based on provided data from a complex experiment that is novel to the student, deduct the fundamental relationship between several controlled variables.
Conduct an investigation, from specifying a problem to designing and carrying out an experiment, to analyzing its data and forming conclusions.

Social Studies Depth-of-Knowledge Definitions
Four levels of depth of knowledge were used for this analysis. Because the highest (fourth) level is rare or even absent in most standardized assessments, reviewers usually made distinctions among DOK levels 1, 2 and 3. The social studies levels were developed by Ann Prewitt and Fred Czarra.

Level 1 Recall of Information
Level 1 asks students to recall facts, terms, concepts, trends, generalizations and theories or to recognize or identify specific information contained in graphics. This level generally requires students to identify, list, or define. The items at this level usually ask the student to recall who, what, when and where. Items that require students to "describe" and "explain" could be classified at Level 1 or 2 depending on what is to be described and explained. A Level 1 "describe or explain" would recall, recite or reproduce information. Items that require students to recognize or identify specific information contained in maps, charts, tables, graphs or drawings are generally level 1.

Level 2 Basic Reasoning

Level 2 includes the engagement of some mental processing beyond recalling or reproducing a response. This level generally requires students to contrast or compare people, places, events and concepts; convert information from one form to another; classify or sort items into meaningful categories; describe or explain issues and problems, patterns, cause and effect, significance or impact, relationships, points of view or processes. A Level 2 "describe or explain" would require students to go beyond a description or explanation of recalled information to describe or explain a result or "how" or "why."

Level 3 Application

Level 3 requires reasoning, using evidence, and a higher level of thinking than the previous two levels. Students would go beyond knowing "how and why" to justifying the "how and why" through application and evidence. The cognitive demands at Level 3 are more complex and more abstract than Levels 1 or 2. Items at Level 3 include drawing conclusions; citing evidence; using concepts to explain "how and why;" using concepts to solve problems; analyzing similarities and differences in issues and problems; proposing and evaluating solutions to problems; recognizing and explaining misconceptions or making connections across time and place to explain a concept or big idea.

Level 4 Extended Reasoning

Level 4 requires even more complex reasoning and the addition of planning, investigating, or developing that will most likely require an extended period of time. The extended time period is not a distinguishing factor if the required work is only repetitive and does not require applying

significant conceptual understanding and higher-order thinking. At this level the cognitive demands should be high and the work should be very complex. Students should be required to connect and relate ideas and concepts *within* the content area or *among* content areas in order to be at this highest level. The distinguishing factor for Level 4 would be evidence through a task or product that the cognitive demands have been met. A Level 4 performance will require students to analyze and synthesize information from multiple sources, examine and explain alternative perspectives across a variety of sources and/or describe and illustrate how common themes and concepts are found across time and place. In some Level 4 performance students will make predictions with evidence as support, develop a logical argument, or plan and develop solutions to problems.

Many on-demand assessment instruments will not include assessment activities that could be classified as Level 4. However, standards, goals, and objectives can be stated so as to expect students to perform thinking at this level. On-demand assessments that do include tasks, products, or extended responses would be classified as Level 4 when the task or response requires evidence that the cognitive requirements have been met.

References

Valencia, S. W., & Wixson, K. K. (2000). Policy-oriented research on literary standards and assessment. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Vol. III.* Mahwah, NJ: Lawrence Erlbaum.

Appendix F: SLO Criteria Table

SLO CRITERIA TABLE
for the Development & Evaluation of Quality Assessments

Subject: _____ **Course:** _____ **Grade(s):** _____

Test Title: _____ **District(s):** _____ **Date:** _____

I. Test Item Construction					
Select-Response Items (Multiple Choice) 1. Question stem is clear. 2. Item is stated in the positive. (For example, refrain from using items, such as "Which of the following is NOT a purpose for the passage?") 3. Item does not give away correct answer. 4. Emphasize qualifiers (e.g., most likely, best) and avoid using "all" or "none of the above." 5. Answer choices are plausible. 6. Answer choices are parallel in length (e.g., words, phrases, sentences). 7. Answer choices are parallel in grammar, semantics, and syntax. 8. Answer choices are in a logical order (e.g., numerical, alphabetical, sensible). 9. Avoid clues in the answer choices. 10. Ensure correct response is the only correct response. 11. Arrange items for easy to more difficult.	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	Test meets 10 or 11 of the 11 select-response item criteria.	Test meets 8 or 9 of the 11 select-response item criteria.	Test meets 6 or 7 of the 11 select-response item criteria.	Test meets 5 or less of the 11 select-response item criteria.	
Supply-Response Items (Short Answer, Essay, etc.) 1. Question stem is clear. 2. Scoring rubric is included. 3. Adequate space for response is provided.	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	Question or prompt is written to utilize higher-order thinking at DOK Levels 3-4 and elicit a unique response. The rubric clearly delineates the expectation of the response.	Question or prompt is written to elicit the appropriate response. The rubric provides general expectations of the response.	Question or prompt is too broad or too narrow to elicit the intended response. The rubric minimally provides expectations of the response.	Question or prompt is unclear and invites a wide range of responses. The rubric does not provide clear expectations of the response.	
Performance Tasks **(Student-created answers or products)**	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐

1. The task is clear and is detailed enough to provide students with an understanding of the expectations (e.g., purpose, product, process, required/suggested resources, time, due dates, presentation format, etc.). 2. Student product and/or performance will sufficiently illustrate student attainment of outcomes. 3. A scoring rubric or evaluative criteria document is included and plainly outlines observable and measurable indicators of the task. 4. A student checklist that is aligned to the evaluative criteria is provided to support student performance and self-monitoring of progress.	The task clearly relates to the intended outcome(s) and indicator(s) to be assessed. The task is written to promote higher-order thinking at DOK Levels 3-4. The rubric clearly delineates the expectation of the response.	The task relates to the intended outcomes or indicators. The task is written to elicit the appropriate student response or product. However, the rubric provides general expectations of the response.	The task relates to the intended outcomes or indicators. However, the task is too broad or too narrow to elicit the intended student response or product. The rubric minimally provides expectations of the response.	The task does not relate to the intended outcomes and indicators. The task is unclear and invites a wide range of student responses or products. The rubric does not provide clear expectations of the response.	

Rubric Development (*If applicable)* 1. The rubric type is appropriately matched to the assessment. 2. The levels of performance are clearly identified. 3. Each level of performance is appropriately and adequately described taking into account the critical elements of the task. 4. Point values are assigned to each performance level.	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	Rubric meets 4 of the 4 criteria.	Rubric meets 3 of the 4 criteria.	Rubric meets 2 of the 4 criteria.	Rubric meets less than 2 of the 4 criteria.	
Formatting 1. All parts of a test question are presented on a single page. 2. The number of items on the test is appropriate for the developmental level of the students. 3. The testing period is the appropriate length for the developmental level of the students. 4. The test is visually easy to read. The graphics, charts, and pictures support test content appropriately. 5. There is consistency in the presentation of item types. 6. All directions are stated clearly and explicitly. 7. If applicable, state the point value of each item type or task.	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	Test meets 6 of the 6 formatting criteria. *If point values are applicable, test meets 7 of the 7 formatting criteria*	Test meets 5 of the 6 formatting criteria. *If point values are applicable, test meets 6 of the 7 formatting criteria*	Test meets 4 of the 6 formatting criteria. *If point values are applicable, test meets 5 of the 7 formatting criteria*	Test meets 3 or less of the 6 formatting criteria. *If point values are applicable, test meets 4 of the 7 formatting criteria*	
Test and Item Bias	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐

The wide range of student experiences and exposure is honored and acknowledged. The test is not biased or offensive with regard to race, gender, native language, ethnicity, geographic region or other factors.	The test and/or items do not contain any words or phrases that would put any student at a disadvantage.	The test may contain words or phrases that could be considered biased or offensive with regard to race, gender, native language, ethnicity, geographic region or other factors.	The test is biased or offensive with regard to race, gender, native language, ethnicity, geographic region or other factors.	The test is biased and offensive with regard to race, gender, native language, ethnicity, geographic region or other factors.	
Section I Test Construction Comments:					

II. Test Validity and Reliability					
Content Validity The test questions sufficiently represent the skills in the specified subject area and adequately assess the skills in the specified standard. (Research indicates that 6 or more items per domain/strand or standard, where appropriate, should be included depending upon the instructional emphasis or weight of the standard in the course.) *Consider the following question(s) in determining content validity.* • Does this test measure what it is intended to measure? • Does the assessment adequately sample the intended learning outcomes? • Are there items on the assessment with no intended learning outcomes?	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	The test adequately samples the intended standards or objectives, and it *does not* assess any learning outcomes that were not intended nor taught.	The test samples the majority of the intended standards or objectives, and it *does not* assess any learning outcomes that were not intended nor taught.	The test inadequately samples the intended standards or objectives.	The test does not sample the intended standards or objectives.	
Construct Validity Use of the SLO Table of Specifications is evident and reflects a clear alignment between Categorical Concurrence (extent to which the items or performance tasks cover the standards, six items or one performance task per domain/strand or standard, where appropriate), Depth of Knowledge (cognitive processing required by each item or performance task compared to the requirements implied by the content objectives), and Range of Knowledge (alignment of items to the multiple objectives within a standard; at least 50% of the standard's objectives must be matched to one or more items or tasks).	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	There is a balanced representation of the content objectives/ cognitive levels, with at least 50% of the test items at or above the standards'	There is a balanced representation of the cognitive objectives/ cognitive levels, with at least 40% of the test items at or above	There is an unbalanced representation of the content objectives/ cognitive levels, with 30% or less of the test items at or above the standards' respective DOK	There is not a balanced representation of the content objectives/ cognitive levels. Most of the test items fall below the standards' respective DOK levels and	

Consider the following question(s) in determining construct validity: • Does the assessment have a sufficient number of items or performance tasks to target each standard to be assessed? • Is the assessment comprised of items that correspond to at least 50% of each standard's objectives (or elements)? Items should assess multiple objectives where possible. • Can logical inferences be made about students' knowledge and/or skills in the course from the assessment?	respective DOK levels and objectives.	the standards' respective DOK levels and objectives.	levels and objectives.	objectives.	
Reliability Six or more test questions or items are included for each domain/strand or standard, where appropriate, (depending upon the instructional time spent or the weight) to reduce the unintended effects of error on the assessment results. *Consider the following question(s) in determining reliability.* • Are there enough questions for each domain or strand assessed? • Is the test length appropriate? Does the test length reduce measurement error and support reliability? • Does the assessment provide for student-specific factors (e.g., fatigue, guessing, marking errors), test-specific factors (e.g., ambiguous items, poor directions), scoring-specific factors (e.g., non-uniform scoring guidelines, computation errors)? • Are the questions, directions, and formatting on the assessment free from systematic error? • Are the grading criteria objective?	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	An adequate number of items are included, the test is free from systematic error, **and** the grading criteria are objective.	An adequate number of items are included, and the test is free from systematic error **or** the grading criteria are objective.	An adequate number of items are included, **but** the test is subject to systematic error and/or the grading criteria are not objective.	There are an inadequate number of items, the test is subject to systematic error, **and** the grading criteria are not objective.	

Section II Test Validity & Reliability Comments:

III. Test Administration Procedures

Test Administration Plan The plan provides detailed and clear instructions that outline appropriate test administration procedures to include the following: 1. Specifications for proper identification and training of testing coordinators and proctors 2. Clearly communicated test administration procedures 3. Clearly outlined time length and testing accommodations 4. Provisions for a script (where appropriate) 5. Adequate access to the appropriate test materials and testing tools for all test participants 6. Clearly communicated test scoring procedures 7. Provisions for inter-rater reliability training (where appropriate)	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	Clear guidelines for test security are provided. Test administration guidelines meet 7 out of the 7 test administration criteria.	This category is not applicable as the assessment must meet 7 out of 7 for test administration criteria. **The assessment cannot move forward.**	This category is not applicable as the assessment must meet 7 out of 7 for test administration criteria. **The assessment cannot move forward.**	This category is not applicable as the assessment must meet 7 out of 7 for test administration criteria. **The assessment cannot move forward**.	

Section III Test Administration Comments:

IV. Test Reporting

Detailed and clear test reporting procedures are provided. 1. The proficiency criteria for the SLO are clearly communicated. 2. The time between test administration, scoring, and reporting of results is timely. 3. The district's data reporting method is clear and consistent with classroom data reports. 4. The data reporting format provides for aggregate data (district, school, class) and individual student data. 5. A protocol is established to provide feedback to students, teachers, administrators, and parents.	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	Test reporting guidelines meet 5 of the 5 test reporting criteria.	Test reporting guidelines meet 4 of the 5 test reporting criteria.	Test reporting guidelines meet 3 of the 5 test reporting criteria.	Test reporting guidelines meet two or less of the test reporting criteria.	

Section IV Test Reporting Comments:

V. Post-Administration (Test Reliability)					
Item Analysis Item analysis was conducted to improve the effectiveness of test items and the validity of test scores. Items were critiqued to determine revision or removal from item bank.	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	Item analysis was conducted and items were critiqued resulting in the revision or removal of test items, if appropriate.	Item analysis was conducted and items were critiqued for future assessment construction.	Item analysis was conducted.	Item analysis was not conducted.	
Reliability of Results The results of the assessment are consistent and dependable. *Consider the following question(s) in determining reliability.* • Do the items discriminate between students with different degrees of mastery: Did the "higher performing" students tend to answer the item correctly while the "lower performing" students responded incorrectly? • Did each item distinguish between those who have learned the standard/or objective and those who have not? • Are test scores free of errors of measurement due to things like student fatigue, item sampling, student guessing? • Do the results reflect the intended learning outcomes?	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐
	The assessment contained 6 or more items or 1 or more tasks to assess each domain or standard. The items/tasks were free from bias. The items were free from ambiguity. The items were free from grammatical or mechanical mistakes.	The assessment contained 6 or more items or 1 or more tasks for most domains or standards. The items/tasks were free from bias. The items were free from ambiguity. The items were free from grammatical or mechanical mistakes.	The assessment contained 6 or more items or 1 or more tasks for some domains or standards. The items/tasks were biased, ambiguous, or included grammatical or mechanical mistakes.	The assessment contained less than 6 items or no task for each domain or standard. There was evidence of bias and ambiguity. The test contained several grammatical or mechanical mistakes.	
Data Use Items are diagnostic and/or conclusive in nature,	**3**☐	**2**☐	**1**☐	**0**☐	**Not Applicable** ☐

providing information regarding misunderstanding and misconceptions in learning and/or demonstration of intended learning outcomes based on student responses. The information can be used to determine student performance of the standard and to prescribe appropriate remediation and inform future test construction. (modified statement)	Item analysis and/or standard analysis data were used to determine student learning trends, inform instruction, **and** assessment development.	Item analysis and/or standard analysis data were used to determine student learning trends and inform instruction **but** were not used to inform assessment development.	Item analysis and/or standard analysis data were used to determine student learning trends.	Item analysis and/or standard analysis were not conducted.	
Section V Post Administration Comments:					

Appendix G: District SLO Form

	SLO GENERAL INFORMATION	
A. District Name	GaDOE Public Domain SLO	
B. State Funded Course Number	Enter the state course number, not the local course number	
C. State Funded Course Title	Enter state course title	
D. Grade(s)	May be a single grade or grade range	
E. Pre-Assessment	☐ Commercially Developed ☒ Locally/Regionally Developed	
F. Pre-Assessment Window	To be determined by the district during a pre-selected assessment window made available by GaDOE, preferably during the first 30 calendar days of the year.	
G. Post-Assessment	☐ Commercially Developed ☒ Locally/Regionally Developed	
H. Post-Test Window	To be determined by the district during a pre-selected assessment window made available by GaDOE, preferably during the last 30 calendar days of the year.	
I. Collaboratively Developed	List assessment/SLO team members and position:	
J. Developed by GADOE Trained Assessment Team	☒ Yes	☐ No

Georgia Department of Education
Phase II District Student Learning Objective (SLO) Form

	SLO CONTEXT AND STATEMENT
1. Selected Standards	Enter the standard number and short description. Elements aren't necessary.
2. Pre and Post Assessment Indicate level of proficiency.	The (course name) Public Domain Post Assessment is a (name type of assessment) and is comprised of (number of items and/or tasks). It assesses students (content, skills, etc.). The proficiency criteria are as follows: Exceeds Proficiency Meets Proficiency Does Not Meet Proficiency The (course name) Public Domain Pre Assessment is a (name type of assessment) and is comprised of (number of items and/or tasks). It assesses students (content, skills, etc.). The proficiency criteria are as follows: Exceeds Proficiency Meets Proficiency Does Not Meet Proficiency **Note: proficiency criteria are usually set by the test creators.**
3. Baseline Data or Historical Data/Trends	
4. SLO Statement	
5. Strategies for Attaining Objective	☐ Required

	☐ Recommended
6. Mid-year Review	The mid-year review is a district and/or school-based decision. It is recommended that teachers review *formative and benchmark classroom and grade-level or content-area data to monitor student progress.* Consider collaborative teacher data review within the content area and across grade-levels where appropriate.

Appendix H: Teacher SLO Form

Teacher Student Learning Objective (SLO) Form

Directions: This suggested form is a tool to assist teachers in meeting the student learning objective set by their district.

Teacher _____ **Course Title** _____ **Grade** ___ **Date(s) of pre assessment**_____ **Date(s) of post assessment** _____	
TI. Setting (Describe the population and special learning circumstances)	
T2. Content/Subject/Field Area (The area/topic addressed based on learner achievement, data analysis, or observational data)	
T3. Classroom Baseline Data (Results of pre assessment)	☐ *Data attached*

T4. Means for Attaining Objective (Strategies used to accomplish the objective)

Strategy	Evidence	Target Date

T5. Mid-year or Mid-course Results	
T6. End-of-year Results	

☐ ***Appropriate Data Received***

Strategies used and data provided demonstrate appropriate Student Growth ☐ ***Yes*** ☐ ***No***

Student Learning Objective Evaluation Rubric

Exemplary (3 pts)	Proficient (2 pts)	Developing/Needs Improvement (1 pt)	Ineffective (0 pts)
Greater than or equal to 90% of students met or exceeded their Student Learning Objective target and 50% or more of these students exceeded their Student Learning Objective target.	Greater than or equal to 80% of students met or exceeded their Student Learning Objective target.	Greater than or equal to 50% and less than 80% of students met or exceeded their Student Learning Objective target.	Less than 50% of students met or exceeded their Student Learning Objective target.
☐	☐	☐	☐
Comments:			

Final Student Learning Objective Score

Teacher's Signature _____ *Date* _____

Evaluator's Signature _____ *Date* _____

Appendix I: Glossary

Appendix J: Resources

Bergan, J.R., Bergan, J. Robert, Bergan, S., Burnham, C., Cunningham, S.A., … Smith, K. (2008). *Building reliable and valid benchmark assessments. A resource for district leaders and staff responsible for benchmark planning, construction, and review.* Tucson, AZ: Assessment Technology, Inc.

Burke, Kay. (2010). *Repacking the Standards.* Solution Tree Press.

Denver Public Schools. (2005). *Student Growth Objective and Monitoring Process Manual.* Denver, CO: Denver Public Schools.

Educational Testing Service. "*Linking Classroom Assessments with Student Learning."*

Flowers, C., Wakeman, S., Browder, D. & Karvonen, M. (2007). *Links for academic learning: An alignment protocol for alternate assessments based on alternate achievement standards*. Charlotte, North Carolina: University of North Carolina at Charlotte.

Gareis, Christopher R., & Grant, Leslie W. (2008). *Teacher-Made Assessments: How to Connect to Curriculum, Instruction, and Student Learning.* New York: Eye On Education.

Herman, Joan L., Heritage, Margaret, & Goldsmith, Pete. (2011). "*Developing and Selecting Assessments and Student Growth for Use in Teacher Evaluation Systems."* California: National Center for Research on Evaluation, Standards, & Student Testing.

Hess, Karin K. (2006). *Exploring Cognitive Demand in Instruction and Assessment.* Dover, NH: National Center for Assessment.

Hindman, Jennifer L., Grant, Leslie W. & Stronge, James H. (2010). *The Supportive Learning Environment: Effective Teaching Practices*. Larchmont, NY: Eye On Education.

Indiana Department of Education. (2011-2012). *RISE Evaluation and Development System: Student Learning Objective Handbook.* Indiana.

Joint Committee on Standards for Educational Evaluation. (2003). *The student evaluation standards.* Thousand Oaks, CA: Corwin
Rhode Island Board of Regents.(2011-2012). *Guide to Measures of Student Learning for Administrators and Teachers.* Rhode Island.

Sato, Edynn, Lagunoff, Rachel, & Worth, Peter. (2011). *SMARTER Balanced Assessment Consortium Common Core State Standard Analysis: Eligible Content for Summative Assessment Final Report.* West Ed.

Stiggins, Rick, Arter, Judith, Chappuis, Jan, & Chappius, Steve. (2006). *Classroom Assessment for Student Learning: Doing It Right – Using It Well*. Portland, OR: Educational Testing Service.

Stronge, James H., & Tucker, Pamela D. (2003). *Handbook on Teacher Evaluation: Assessing and Improving Performance*. Columbus, OH: Eye On Education.

Stronge, James H., Tucker, Pamela D., & Hindman, Jennifer L. (2004). *Handbook for Qualities of Effective Teachers*. Alexandria, VA: Association for Supervision and Curriculum Development.

Stronge, James H. (2007). *Qualities of Effective Teachers*. Alexandria, VA: Association of Supervision and Curriculum Development.

Stronge, James H., Richard, Holly B., & Catano, Nancy (2008). *Qualities of Effective Principals*. Alexandria, VA: Association for Supervision and Curriculum Development.

Stronge, James H., & Grant, Leslie W. (2009). *Student Achievement Goal Setting: Using Data to Improve Teaching and Learning*. Larchmont, NY: Eye On Education.

Stronge, James H. (2010). *Effective Teachers = Student Achievement: What the Research Says*. (2010). Larchmont, NY: Eye On Education.

Stronge, James H. (2010). *Evaluating What Good Teachers Do: Eight Research-Based Standards for Assessing Teacher Excellence*. Larchmont, NY: Eye On Education.

Taylor, Joann, Stautinger, Paul, Spencer, Tessa, & Preston, Jon Mark. *Student Learning Objective Manual.*Austin, TX: Austin Independent School District.

Taylor, L.,Wise, L., Thacker, A., Moody, R., Koger, L., Dickinson, E., Matthew, T.D. (2007). *Independent evaluation of the alignment of the california standards tests (CSTs) and the california alternate performance assessment (CAPA).* Sacremento, CA: Human Resources Research Organization (HumPRO).

Webb, Norman L. (2009). *Depth of Knowledge Guide: Career and Technical Education Definitions.* http://www.mde.k12.ms.us

Webb, Norman, L. & Others. (2005). *"Web Alignment Tool"*. Madison, Wisconsin*:* Wisconsin Center of Education Research.