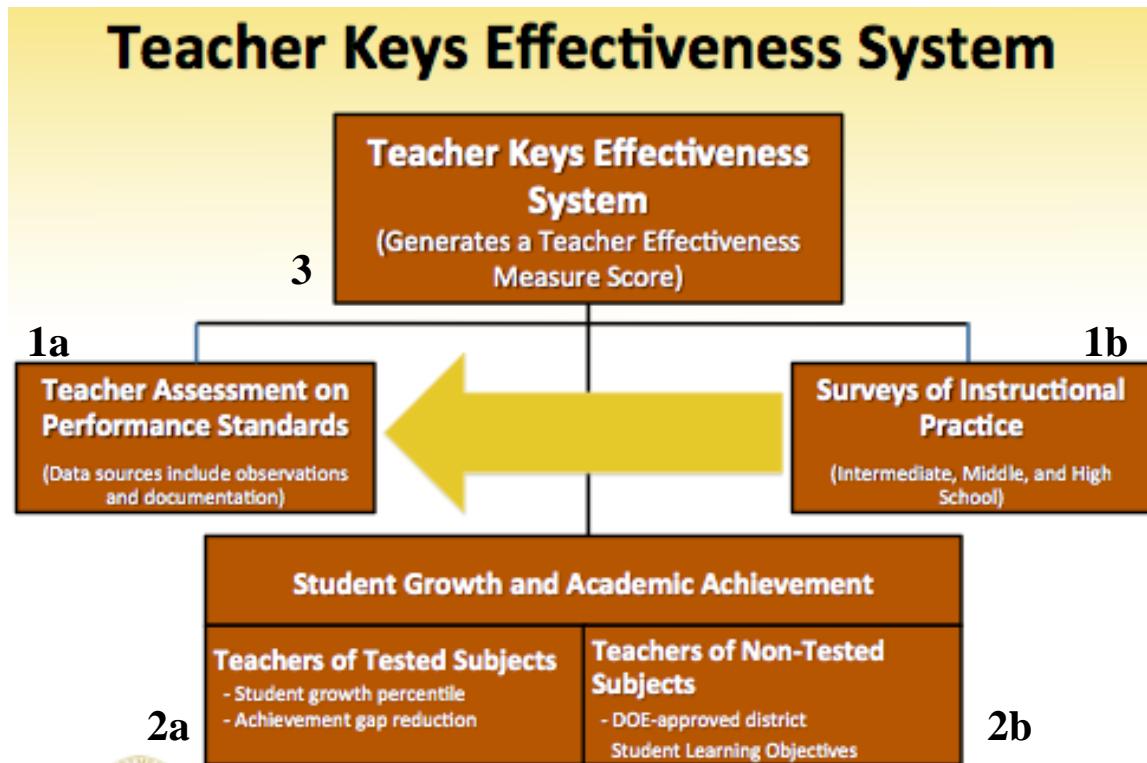




Dr. John D. Barge, State School Superintendent
"Making Education Work for All Georgians"

Overview to the 2012 TKES/LKES Pilot Evaluation Report

The contents of this report were developed under a grant from the U.S. Department of Education. However, those contents do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the Federal Government.



From January to May, 2012, schools piloted the Teacher Assessment on Performance Standards, which included classroom observations and documentation of teacher practice to determine proficiency on the ten TAPS standards (1a). The student surveys of instructional practice were also piloted during this time frame, and that information was used by evaluators to inform ratings of teacher performance on the standards (1b).

The data from this component of the pilot resulted in the distributions illustrated in the following charts from the pilot evaluation report published in December.

Pilot LAPS Evaluation Data	
Standard	% Proficient and Exemplary
1. Instructional Leadership	90.7
2. School Climate	90.4
3. Planning and Assessment	87.2
4. Organizational Management	93.4
5. Human Resources Management	85.9
6. Teacher / Staff Evaluation	92.5
7. Professionalism	96.0
8. Communication and Community Relations	89.5
TOTAL	89.8

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

State TAPS Ratings by Domain and Level of Evaluation				
Standard # and Name	% Ineffective	% Developing /Needs Improvement	% Proficient	% Exemplary
Planning	0.24	4.42	75.13	20.22
Instructional Delivery	0.43	8.93	72.53	18.12
Assessment of and for Learning	0.55	7.67	82.77	9.01
Learning Environment	0.28	5.82	68.54	25.36
Professionalism and Communication	0.12	2.92	73.24	23.86
TOTAL	0.32	5.95	74.44	19.41

This summarizes the performance rating data by standard and domain for the pilot, but it does not reflect an overall performance rating for the ten TAPS standards. Those ratings have not been calculated yet, but they will be calculated in the coming weeks and published in an addendum to the report.

The second component of the Teacher Keys Effectiveness System (TKES), measures of student growth and achievement, was also a part of the 2012 pilot. These measures include Student Growth Percentiles (SGPs) for teachers of courses that have state assessments (2a) and Student Learning Objectives (SLOs) for teachers of courses that do not have state assessments (2b). These measures are designed to assess a teacher's effectiveness in increasing student learning in a course. Quantitative data from the SGP and SLO measures are currently being analyzed and will be published in an addendum to the report.

The overall Teacher Effectiveness Measure (3) is determined using a combination of performance on the TAPS ratings and on the growth measures. Once the overall TAPS measures have been calculated (1) and the quantitative SGP and SLO data has been analyzed (2a and 2b), the two component measures will be combined to determine the TEM (3). The TEM performance (3) data will also be reported in the addendum to the report.

The lag in the student growth analysis and overall calculations is due to two factors:

- the lack of a comprehensive data collection system for the pilot that required collecting, cleaning, and analyzing data from multiple sources and in different formats, and
- the need to put the GaDOE evaluation team into place prior to completing the analysis.

These two factors that delayed the analysis have been appropriately addressed, and the work is now moving forward.

As stated in the pilot report, referring to the data reported on individual TAPS standards, not overall TAPS ratings:

TAPS data appears to be strongly skewed to the positive, with 93.85% of the ratings proficient or exemplary, compared to 6.27% ineffective and developing/needs improvement. Having this large of percentage of overwhelming positive scores indicates a need for further training or increased monitoring by the GaDOE. However, consideration should be given to factors

Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report

that potentially influenced the strong positive ratings. The Teacher Assessment on Performance Standards (TAPS) evaluation was part of a parallel evaluation system occurring concurrently with the district's standard evaluation process. Because the principals had to do TAPS classroom assessments in addition to their "regular" evaluation system, extra demands were placed on the principals. They knew this was a pilot and results would not count so, in some cases, it was not taken as seriously or implemented with fidelity. The short time frame for implementation and the newness of the materials, standards, and process should also be taken into account.

During the pilot, we think it less appropriate to conclude that there was lack of principal buy-in than that there was a double burden on the principals for evaluation, and that the time frame for the pilot was shorter than would allow for ideal implementation. It is also appropriate to conclude that the challenge of moving from a predominately binary rating system (satisfactory/unsatisfactory), where more than 99% of teachers have historically been rated satisfactory, to a four-level rating system where a number of teachers' performance will be rated as other than proficient, presents a major culture shift not likely to be adequately addressed in a five month pilot or even in the first year of full implementation. The magnitude of this cultural shift cannot be overstated in considering the work. Feedback we have received from the RT3 districts that participated in the 2012 pilot has clearly indicated significant principal buy-in and that teachers especially value the specific feedback the system is designed to provide.

We are not prepared at this time to speculate on what the percentage of ineffective, needs development, proficient, and exemplary ratings on the TEM (3) will be once the final calculations have been completed with the pilot data. The pilot data analyzed and reported thus far reflects performance on individual components only, and not on composite measures (TEMs). The student growth measures have not been included in the calculations yet, and it is important that we not draw unwarranted conclusions based on this partial report. However, during the 2012-2013 and 2013-2014 implementation years, as opposed to the 2012 pilot, we do expect that the ineffective/needs development TEMs will account for up to 10% or so of the performance, that the exemplary TEMs will account for 5-10% of the performance, and that most teachers will have TEMs in the proficient range.

At this time there are no plans to limit districts in the future to a certain percentage of top scorers. However, over time, as district and school personnel gain proficiency with the system, we expect to observe several changes in culture, in practice, and in the resulting data.

- Lower performing teachers will be more consistently identified and then provided with professional development to support improving performance in areas of concern.
- More teachers will become more proficient as a result of this type of individually targeted, effective professional development.
- Fewer lower performing teachers will be hired and/or retained in districts.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Conclusions from the pilot evaluation report, initial release are:

As GaDOE makes adjustments to TKES and LKES, the following recommendations need to be considered:

- **Strengthen the fidelity of implementation for TAPS and LAPS rubrics.** If the TAPS and LAPS standards are to accomplish the instructional changes in classrooms and schools across the state, it is imperative the rubrics are applied with consistency and greater balance. District and school leaders need a stronger understanding of the standards, rubrics and exemplars. Training for TAPS and LAPS needs to include a deeper focus on application of the rubric, and participants must reach a specific level of competence at the end of the training. Increasing inter-rater reliability among evaluators will be critical to consistency of implementation statewide.
- **Create consistency in TKES orientation training.** Those being evaluated, teachers, require a strong understanding of the TAP standards and process. To apply the standards without effective orientation to them is an improbable expectation. Orientation videos developed by GaDOE trainers may be one option to help standardize training.
- **Continue to focus on improving the development, application and use of Student Learning Objectives (SLOs).** Because the assessments are ultimately district developed or identified, there are concerns about quality, standardization, consistency, and confidentiality. The limited time period for development, and the inconsistency in the level of expertise across developers, influence the quality of the assessments. It is suggested that GaDOE use the current and new SLOs to develop an item bank of exemplars that can be used by districts. Districts and non-tested subject teachers need training and assistance on how to set appropriate growth targets and use pre-assessments to inform instruction.
- **Assist teachers and leaders with analysis and use of instructional survey data.** Professional learning may be beneficial for teachers and leaders on using the instructional surveys for improvement at the classroom, grade, school, and district level. Over time, teachers will likely become less anxious about students completing the instructional surveys. However, if teachers and leaders do not use the survey data effectively, the potential value is lost.
- **Provide data to districts and schools in a timely manner.** It is understandable that the data related to the pilot would take time to process and return to districts. As previously mentioned, pilot data was received later than anticipated because of the relaxation of initial deadlines and challenges experienced with data collection. Deciding to move the deadlines back was the correct decision for the pilot, schools, and districts; however, the delay in returning TKES and LKES data prevented its use for improvement planning. As processes and procedures are finalized, and as the SGP process is completed, data needs to be returned to the participants as soon as possible.

With slight revisions and continued analysis of progress, TKES and LKES have the potential to positively influence education in Georgia for years to come.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**



*Dr. John D. Barge, State School Superintendent
"Making Education Work for All Georgians"*

Teacher Keys and Leader Keys Effectiveness Systems

2012 Pilot Evaluation Report

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

TABLE OF CONTENTS

Introduction.....	4
The Pilot.....	5
Evaluation Plan	6
Pilot Data	6
Observation Data	6
Focus Group Data	7
Survey Data.....	8
Document Analysis.....	8
Reported Data	9
Training	10
Research Question 1	10
Research Question 2	15
Implementation	19
Research Question 3	19
Research Question 4	44
Pilot Verification and Validation.....	46
Research Questions 5 and 6	46
Research Question 7	46
Research Questions 8 and 9	47
Research Question 10	47
Research Question 11	50
Conclusion	51
References.....	56
Appendices	
Appendix A: Areas for Focus Groups.....	57
Appendix B: District Responses to End of Pilot Surveys	58

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

The contents of this report were developed under a grant from the U. S. Department of Education. However, those contents do not necessarily represent the policy of the U. S. Department of Education, and you should not assume endorsement by the Federal Government.

Acknowledgments

The Georgia Department of Education's (GaDOE) Teacher and Leader Keys Effectiveness Systems Pilot Evaluation Report was developed with the thoughtful contributions of the Georgia Race to the Top districts. We wish to express our appreciation for their conscientious and insightful efforts for a successful pilot and evaluation of that pilot.

Project Consultant

Dr. Georgia G. Evans, Quality Educational Services

Teacher Keys and Leader Keys Effectiveness Systems Pilot Evaluation Report Usage Statement

The Teacher Keys and Leader Keys Effectiveness Systems Pilot Evaluation Report was developed on behalf of the Georgia Department of Education to assist with implementation of Georgia's Race to the Top (RT3) plan.

The materials in this report are copyrighted by the Georgia Department of Education. The materials may not be revised or modified without the express written permission of the copyright holder. Georgia public schools may use these materials without alteration to meet applicable requirements or for educational purposes as long as the materials continue to reflect: "All Rights Reserved."

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

**Data Report on the
Georgia Teacher Keys and Leader Keys Effectiveness Systems
2012 Pilot**

Introduction

As part of the Race to the Top Initiative (RT3), Georgia piloted the Teacher Keys Effectiveness System (TKES) and Leader Keys Evaluation System (LKES) from January through May 2012. The TKES and LKES create a common definition of teacher and leader effectiveness that allows for consistency and comparability across districts.

The Teacher Keys Effectiveness System (TKES) consists of the following three components:

1. Teacher Assessment on Performance Standards (TAPS) is a qualitative rubric-based method by which evaluators use quality performance standards to measure teacher performance.
2. Student Growth and Academic Achievement uses Student Growth Percentiles (SGPs) for teachers of tested subjects and Georgia Department of Education (GaDOE) approved Student Learning Objectives (SLOs) for teachers of non-tested subjects. Student Learning Objectives (SLOs) will use district achievement growth measures.
3. Surveys of Instructional Practice are student surveys of both tested and non-tested K-12 teachers.

The three components above contribute to an overall Teacher Effectiveness Measure (TEM). The amount contributed to the TEM by each component depends on whether teachers are in a tested or non-tested course.

The Leader Keys Effectiveness System (LKES) also consists of three components, which contribute to an overall Leader Effectiveness Measure (LEM).

1. Leader Assessment on Performance Standards (LAPS) is a qualitative rubric-based method by which evaluators can measure leader performance related to quality performance standards.
2. Student Growth and Academic Achievement includes Student Growth Percentiles (SGPs) and GaDOE approved Student Learning Objectives (SLOs) utilizing district achievement growth measures. This area also includes a measure of Achievement Gap Reduction at the school level.
3. Governance and Leadership measures include climate survey data, student attendance data, and will include data on the retention of effective teachers when it becomes available.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

The Pilot

Twenty-six RT3 districts participated in the spring 2012 pilot of the Teacher Keys Effectiveness System (TKES) and Leader Keys Effectiveness System (LKES). The purpose of the pilot was to (1) improve administration procedures, training, and educator familiarity with the system, and (2) collect preliminary data that can inform instrument revision and validation plans prior to full operational implementation.

The participating districts provided critical feedback and data used to assess the effectiveness of training, implementation, and evaluation efficacy (verification and validation measures). Prior to the pilot, an evaluation plan composed of eleven key research questions was developed to guide the data collection process during the pilot. Due to the time limitations associated with the pilot, several implementation decisions were made prior to the development of this evaluation plan.

First, districts were allowed to select one of two random sampling options for the pilot:

Option 1: An across-school model in which approximately 10% of teachers and 25% of the leaders across all schools in the district were randomly selected for participation by the Georgia Department of Education (GaDOE).

Option 2: A whole-school model in which all teachers and leaders in selected schools within the district participated. These schools represented approximately 10% of the total number of teachers and leaders in the district.

This process resulted in approximately 5,800 of the more than 48,000 teachers being randomly selected to participate in the 26 RT3 districts.

Option 1: Across School Model Districts	Option 2: Whole School Model Districts
1. Atlanta City Schools 2. Burke County Schools 3. Carrollton City Schools 4. Chatham County Schools 5. Cherokee County Schools 6. Clayton County Schools 7. Dade County Schools 8. DeKalb County Schools 9. Dougherty County Schools 10. Gainesville City Schools 11. Hall County Schools 12. Meriwether County Schools 13. Muscogee County Schools 14. Peach County Schools 15. Pulaski County Schools 16. Rabun County Schools 17. Rockdale County Schools 18. Spalding County Schools 19. Treutlen County Schools	1. Ben Hill County Schools 2. Bibb County Schools 3. Gwinnett County Schools 4. Henry County Schools 5. Richmond County Schools 6. Valdosta City Schools 7. White County Schools

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Second, the GaDOE required each district to select 10 courses for participation in the Student Learning Objectives (SLO) component during the pilot period.

As a result, not all of the non-tested subjects and teachers were included during the pilot. Approximately 3,360 teachers of non-tested subjects (70% of the sample) have been included.

Evaluation Plan

The Georgia Department of Education (GaDOE) Teacher and Leader Effectiveness Division undertook the formidable task of implementing this new, large-scale evaluation system during the 2011-2012 school year. Piloting the Teacher Keys Effectiveness System (TKES) and Leader Keys Effectiveness System (LKES) involved significant training on a variety of new instruments for evaluators, districts, principals, and teachers from the 26 participating districts. The evaluation plan was written prior to the rollout of the pilot and it identified research questions for which the Teacher and Leader Division of the GaDOE planned to collect data. Two of the questions in the plan (3b, 3e) referred to pilot components being completed on schedule. Given the magnitude of the work, the newness of the instruments, human resource limitations at the district level, and the other demands on participating districts, the GaDOE relaxed deadlines for work completion given to districts. This decision was prompted by the GaDOE's desire for quality data instead of simply compliance. The GaDOE anticipated districts would provide high quality data and better meet fidelity expectations if they had adequate time to collect and record the various TKES and LKES components. Therefore, the deadlines for submission were relaxed.

Pilot Data

Pilot data was collected from observations, focus groups, surveys, document analyses and both qualitative and quantitative reported data. The data collected for research questions was identified and aligned to the respective question. Multiple sources of data were available for each question to allow for triangulation of information.

(1) Observation data

Consultants from Stronge and Associates Educational Consulting, LLC visited and reviewed 10 site-training sessions provided by nine trainers to collect data on the GaDOE trainers' delivery of the professional learning modules and to determine the congruence between their behavior and training specifications.

The two reviewers rated the trainers' performance through the use of an observation protocol. Eight presenter skills related to the training are included in this protocol: 1)

Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report

accuracy of information presented; 2) clarity of information; 3) clarity of directions; 4) logical flow; 5) pacing; 6) rapport with audience; 7) engagement of audience; and 8) response to audience questions. Trainers' performance on each of these eight domains was rated based on a rubric, which consists of a four-point rating scale: 1-Weak, 2-Inconsistent, 3-Proficient, and 4-Advanced.

In addition, the observation protocol examined trainers' performance on specific professional learning content. A total of 10 topics and activities were included: (1) Qualities of Effective Teachers PowerPoint presentation; (2) Design a Teacher activity; (3) TKES orientation PowerPoint presentation; (4) Sample Documentation activity; (5) Scavenger Hunt activity; (6) Look-fors and Red Flags activity; (7) Math video vignette; (8) AP Biology vignette; (9) Survey presentation; and (10) Student Learning Objectives presentation. Trainers' performance on these ten topics/activities was rated by the same rubric mentioned above on a continuum from "Weak" to "Advanced." The chart below provides information on the results of the trainers' ratings by the reviewers, and validates that TKES and LKES trainers provided accuracy and clarity of information, clear directions, a logical flow and pacing, and positive interactions with the training participants. On a four-point scale, the average mean for each domain ranged from 3.2 to 3.6.

TKES Site Visit Summary of GaDOE Trainers									
Domain	Advanced		Proficient		Inconsistent		Weak		Average of Ratings
	#	%	#	%	#	%	#	%	
Accuracy of information presented	5	50%	4	40%	1	10%	0	0%	3.4
Clarity of information	4	40%	6	60%	0	0%	0	0%	3.4
Clarity of directions	4	40%	6	60%	0	0%	0	0%	3.4
Logical flow	6	60%	4	40%	0	0%	0	0%	3.6
Pacing	3	30%	6	60%	1	10%	0	0%	3.2
Rapport with audience	6	60%	4	40%	0	0%	0	0%	3.6
Engagement of audience	4	40%	5	50%	1	10%	0	0%	3.3
Response to audience questions	4	40%	6	60%	0	0%	0	0%	3.4

(2) Focus group data

The Georgia Department of Education's Teacher and Leader Effectiveness division contracted with the Center for Community Studies, Inc. to assist with the focus group interviews, as outlined in the TKES and LKES evaluation plan. The Center for Community Studies, Inc. (1) selected a random sample of the systems for focus groups; (2) developed a focus group moderator's guide for teachers, principals, and district representatives; (3) provided staff training for the moderator's guide; (4) pilot tested the moderator's guide; (5) developed an exit survey for the teachers' focus group; (6) selected a random sample of teachers for the focus groups; and (7) consolidated the focus group findings into an executive summary. The topic areas for moderator questions for teachers, principals, and districts are found in Appendix A.

Each focus group had a facilitator and at least one recorder from the GaDOE team, who were assigned to groups in districts where they did not provide support and assistance

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

during the pilot process. The evaluator requested transcriptions of the flip chart paper pages, written notes of the recorder, and a general summary prepared by the facilitators/recorders. These documents were analyzed and consolidated into the Focus Groups' Executive Summary.

(3) Survey data

There were 2,584 respondents to the End of Pilot Survey for Teacher Keys and Leader Keys Effectiveness Systems from 26 different school systems in the state of Georgia. The greatest numbers of respondents were from the largest participating school districts: DeKalb County Schools (21%), Gwinnett County Schools (15%), and Chatham County, Clayton County, and Muscogee County Schools (9% each). A breakdown of district responses is included in Appendix B.

The majority of the respondents were classroom teachers (77%), followed by school principals (20%), and district level personnel (3%). Broken down by grade level groups, the majority of survey participants (47%) were from elementary schools (K-5). Twenty percent were from middle schools, high schools (9-12) had 26% participation, and 7% were from other categories. Respondents in the "other" category included those from alternative schools, district office, magnet schools, and special education programs. Eighty-eight percent of the respondents participated in TKES only, 10% participated in both TKES and LKES, and 2% participated in only LKES. Less than one percent served as coordinators, facilitators or supervised the program.

Students in grade K-12 participated in Surveys of Instructional Practice of their teachers. Over 164,000 students took part in the surveys during spring 2012. The data from these have been disaggregated by grade level assignments: K-2, 3-5, 6-8, and 9-12. Surveys at the elementary level (K-5) had 3 selection options: no, sometimes, and yes. Middle school and high school surveys contained the selection options: strongly disagree, disagree, agree, and strongly agree.

Teachers and staff members of the participating schools completed a 25-question climate survey related to the Leader Keys Effectiveness System (LKES). While the number of respondents to each question varied, more than 13,300 surveys were completed across the state. The data from these surveys were disaggregated at the state, district, and individual school level.

4) Document Analysis

Document analysis was conducted on Student Learning Objectives (SLOs), which comprise the Student Growth and Academic Achievement component of the TKES for teachers of non-tested subjects. Ten SLOs were developed by each of the 26 Race to the Top districts for non-tested courses. SLOs measure student academic growth between two points in time. Each SLO was submitted to the GaDOE for review and approval. A rubric was used to assess rigor of SLOs, and that information will be used in the evaluation for research question 11.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

5) Reported Data

Various types of data were collected before, during, and after the pilot (see chart below). It was analyzed at the state and district level using SPSS and Excel pivot tables to identify patterns, trends, and data outliers. Data has been compared and triangulated to answer the research questions identified in the evaluation plan. As additional data is analyzed and comparisons completed, an addendum to this report will be published.

Reported Data
<ul style="list-style-type: none">• Pilot Survey data from District leaders, Principals, and Teachers• Student Learning Objectives (SLOs) data by State and District• LKES State and District data on School Leaders' Climate Surveys• LAPS data by State and District• TKES Student Survey data by State and District• TAPS data by State and District• Site Visit Reports from 26 RT3 Districts• Student Growth Percentiles (SGP) data

The focus of this report was on the three key areas identified in the evaluations plan: training, implementation, and validation. Data was analyzed primarily at the state level to allow for a comprehensive look at the pilot across all districts. While district level data is important to the individual district, it does not allow for the broad analysis needed for a state pilot.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Training

With the pilot and preparation for implementation of a new state evaluation system, significant training had to occur for various groups. Evaluators, districts, principals, and teachers were all trained on the various components of the Teacher Keys Effectiveness System (TKES) and the Leader Keys Effectiveness System (LKES). The first set of research questions relates to the overall understanding of the TKES and LKES evaluation systems, and the training provided for them. Data from (1) observations, (2) focus group interviews, and (3) surveys were collected and analyzed in order to answer the following questions.

Research Question #1: Do educators understand the TKES and LKES systems?

Teacher Keys Effectiveness System (TKES):

In fall 2011 and winter 2012, Georgia Department of Education trainers from the Teacher and Leader Effectiveness Division led a series of 56 workshops for more than 1,650 participants following which data were collected. The trainings were designed to orient and prepare district and school leaders to implement the spring 2012 pilot for the Teacher Keys Effectiveness System (TKES). Fourteen GaDOE Teacher and Leader Effectiveness trainers led most of the sessions. Dr. James Stronge, who developed several TKES components, led a session with 40 participants. District leaders from Cherokee County, Clayton County, and Gwinnett County Schools assisted with several sessions.

At the conclusion of each workshop, participants were asked to complete an evaluation form on the effectiveness of the training. Ten Likert scale questions were to be rated on a scale of one to five, with one being “strongly disagree” and five being “strongly agree”. When combined, the agree/strongly agree percentages from the 56 workshops reflect an overwhelmingly positive response to the training sessions. All but one question scored above 90% agree/strongly agree. Question eight, which had the lowest score, with 83.6% of the evaluations being rated agree/strongly agree, stated, “As a result of this training, I have a clear picture of my next steps with the TKES pilot.” The lower positive response to this question is typical of a new program implementation and is indicative of participant uncertainty about any new process. Results of the evaluations are in the chart below.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Evaluation Responses to GaDOE TKES Training Fall 2011 – Winter 2012			
Questions	% Agree/ Strongly Agree	% Strongly Agree	% Agree
1. These sessions effectively engaged me as a learner.	96.4	62.9	33.5
2. The presenters were very knowledgeable and effectively presented the information.	96.8	75.1	21.7
3. The training materials were designed so I can use them with my district staff/schools.	93.9	62.9	31.0
4. This training strengthened my understanding of the TKES pilot.	95.8	69.5	26.3
5. This training will help my district/school to increase the effectiveness of our teachers.	90.1	54.1	36.0
6. I had ample opportunities to ask questions and share information during the training sessions.	98.2	80.8	17.4
7. There were opportunities for my active participation, for reflection, and for processing and applying the information.	97.8	77.9	19.9
8. As a result of this training, I have a clear picture of my next steps with the TKES pilot.	83.6	42.4	41.2
9. This training will help me perform my role in the RT3 pilot more efficiently and effectively.	90.5	50.1	40.4
10. I clearly understand where I can get support from the Georgia Department of Education if I have questions about our TKES pilot.	93.7	65.0	28.7

In addition to the 10 Likert scale questions on the evaluation, there were four open-ended questions (11-14) associated with the value of the workshop. Question 11 asked the participants to identify the “most valuable thing” they learned. There were a total of 1,130 responses to this question from the 56 trainings. The majority made reference to some variation of “having gained a greater understanding of the TKES documents, process, and components.”

Question 12 of the evaluation asked participants to identify the “least valuable component” of the training. There were 470 responses to this question, including comments such as, “All were valuable;” “Can’t think of anything;” “Everything was important.” While the question indicates the likelihood of the identification of a negative part of the training, the responses were still predominantly positive. The only theme detected in the comments to this question points not to the least valuable component, but to a few of the participants being somewhat overwhelmed with the amount of material presented.

The 621 responses to Question 13 about the need for more information on some component of the training acknowledged that participants wanted additional training and answers to the unknowns. Because the pilot was still in the development phase and answers for some components had not been determined, participant questions could not always be answered fully. Many acknowledged an understanding of this being part of the

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

development process, and the majority looked forward to receiving more information and additional training when it was available, particularly closer to implementation.

Numerous participants included comments regarding SLOs, though the nature of those varied. Some wanted additional information, such as how to develop SLOs, how pre/post assessments would be used, and how SLO data will be used for teacher evaluation. The amount of training time, either too much or not enough, seemed to be another theme for question 13. Some believed there was not enough time for practice and they would have liked more time for conversations. Others said the two-day format of the workshop was too long, and being out of their school for two days was a challenge. Many indicated that they wanted more training, but it should be divided up and not provided all at one time. There were also comments mentioning that some of the training materials could be improved. In particular, participants felt that the videos were of poor quality (hard to see and hear) and too long. They would also have preferred to see videos of Georgia teachers using Georgia standards-based teaching strategies, and multiple examples of good teaching or a parallel of good and bad teaching on the same standard.

A small number of the responses to this question demonstrated misunderstandings about the TKES system. For example, one participant requested information on writing SLOs for media specialists, when media specialists are not subject to TKES evaluation. Another participant wanted information on how to turn in the documentation before an unannounced observation. Someone else wanted to know how teachers without CRCT/EOCT tests would be evaluated. All of these comments indicate that several participants had missed key components of training, and that they did not have a full understanding of TKES.

Question 14 was an opportunity for participants to provide “open-ended” comments. There were 565 responses noted during the course of the various workshops. The majority of these involved praise for the trainers. Approximately 70% of the comments were positive and included statements about trainers’ knowledge, disposition, and presentation abilities. A large number of the 565 comments included a “thank you” for the training.

An additional 14 workshop sessions were held for training on the TKES summative process following which data were collected. These sessions were attended by 463 participants from mid-December through mid-March. There were six trainers who conducted these sessions, including five GaDOE trainers and Dr. James Stronge. The same 14 questions used for the initial TKES training were asked on the evaluation forms completed at the end of these sessions. Responses to the 10 Likert scale questions are in the chart below.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Evaluation Forms for TKES Summative Training			
Question	% Agree/ Strongly Agree	% Strongly Agree	% Agree
1. These sessions effectively engaged me as a learner.	94.9	54.9	40.0
2. The presenters were very knowledgeable and effectively presented the information.	98.1	67.4	30.7
3. The training materials were designed so I can use them with my district staff/schools.	95.9	61.7	34.2
4. This training strengthened my understanding of the TKES pilot.	95.2	62.2	33.0
5. This training will help my district/school to increase the effectiveness of our teachers.	90.3	53.5	36.8
6. I had ample opportunities to ask questions and share information during the training sessions.	98.0	72.0	26.0
7. There were opportunities for my active participation, for reflection, and for processing and applying the information.	97.9	71.9	26.0
8. As a result of this training, I have a clear picture of my next steps with the TKES pilot.	88.5	45	43.5
9. This training will help me perform my role in the RT3 pilot more efficiently and effectively.	93.9	52.6	41.3
10. I clearly understand where I can get support from the Georgia Department of Education if I have questions about our TKES pilot.	96.1	65.9	30.2

Responses to the open-ended questions 11-14 from these follow-up sessions reflected a stronger understanding of how to apply and score the summative assessment for TKES. The 250 responses to question 11 clearly demonstrate that participants gained knowledge and increased their understanding of TKES by attending this summative training.

Question 12 had 67 responses, and many of these were again positive statements about all components of the training being valuable. Several noted there were too many simulations, and videos were very long.

Inter-rater reliability, SLOs, and the electronic platform appeared to be the key themes identified by participants as topics for additional information in response to Question 13. Forty-three of the 73 comments on Question 14 were positive and complimented the trainers' presentation style, expertise, and interaction with participants.

Based on the preponderance of evidence from the training evaluation data, district leaders and principals had a sufficient understanding of TKES. This was substantiated during focus group interviews where principals spoke positively about the training, expertise, and support provided by state team members. Most principals reported orientation was thorough and the manuals were useful for reference.

While data analysis indicated district leaders and principals had a strong understanding of TKES, teachers did not have consistent knowledge and training. Teacher focus groups and pilot surveys revealed a vast difference in the level of understanding among teachers.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Teacher Keys Effectiveness System (TKES) orientation training was conducted by school leaders. While some teachers expressed that orientation was comprehensive and effective, others noted that their school leaders were ill prepared and could not answer questions about the TKES program. Many teachers believe that the orientation training implementation was inconsistent in quality. Orientation training times ranged from 15 minutes to a full day, and in some cases, teachers were not trained, but had to attempt to figure everything out on their own. Descriptions of the training sessions ranged from “confusing and chaotic” to “necessary and helpful.”

Leader Keys Effectiveness System (LKES):

There were 15 training sessions conducted by seven trainers for Leader Keys Effectiveness System (LKES) following which data were collected. Sessions occurred in January and early February 2012. The evaluation form completed at the end of each training session mirrored the evaluation form used for the TKES sessions, with the same 10 Likert scale and four open-ended questions. The evaluations, completed by 172 participants, were tremendously positive with all questions receiving over 90% agree/strongly agree responses. The highest response was to question two, which referenced the knowledge of presenters and the effectiveness of the presentation. The agree/strongly agree percentages to the 10 Likert scale questions are found in the chart below:

Evaluation Forms for LKES Training			
Question	% Agree/ Strongly Agree	% Strongly Agree	% Agree
1. These sessions effectively engaged me as a learner.	97.0	63.7	33.3
2. The presenters were very knowledgeable and effectively presented the information.	98.8	71.5	27.3
3. The training materials were designed so I can use them with my district staff/schools.	95.9	66.9	29.0
4. This training strengthened my understanding of the LKES pilot.	96.6	64.0	32.6
5. This training will help my district/school to increase the effectiveness of our teachers.	91.9	59.3	32.6
6. I had ample opportunities to ask questions and share information during the training sessions.	98.2	75.3	22.9
7. There were opportunities for my active participation, for reflection, and for processing and applying the information.	98.3	72.7	25.6
8. As a result of this training, I have a clear picture of my next steps with the LKES pilot.	93.6	53.5	40.1
9. This training will help me perform my role in the RT3 pilot more efficiently and effectively.	94.2	55.6	38.6
10. I clearly understand where I can get support from the Georgia Department of Education if I have questions about our LKES pilot.	97.6	67.8	29.8

The LKES open-ended questions 11-14 provided similar responses to the TKES comments. There were 85 responses to question 11, where participants were asked about the “most valuable” aspect of the training. The responses were overwhelmingly positive and the vast

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

majority indicated a strong understanding of the LKES components, process, and expectations.

Question 12 asked about the “least valuable” component of the training. There were 25 responses to this question, 10 of which were positive statements about the value of all components of the training, or the trainer’s knowledge and expertise. A few LKES participants noted their frustration with questions that could not be answered due to the simultaneous rollout of training and the pilot. These comments were similar to those expressed during the TKES training.

Twenty-four of 29 participants responded to question 13. The predominant theme was a request for more information on LKES expectations, documentation, and pilot feedback. Question 14 was an open-ended opportunity for comments. Thirty-nine of the 48 statements were positive in nature and related to the trainers or training in general.

Research Question #2: Were the training sessions successful in preparing evaluators to implement the TKES and LKES evaluation systems?

TKES Implementation: District Leaders / Evaluators and School Leaders / Evaluators

Based on the TKES training evaluation ratings of the 10 Likert scale questions on pages 13, the training was highly successful in preparing evaluators for implementation of the Teacher Keys Effectiveness System. The percentage of agree/strongly agree responses clearly reflect that participants had opportunities to ask questions, be actively engaged, apply the information, and strengthen their understanding of TKES.

The 3,147 responses to questions 11-14 continued to support the positive responses noted on questions one through 10. In general, participants thought the workshops provided a good overview and orientation to TKES. Most had little or no knowledge of the new system prior to the training, and appreciated the opportunity to practice and learn about the various components. The majority of comments were extremely positive about the trainers. Many mentioned the trainers’ knowledge, effectiveness, and how well the trainer performed during the presentation. Data compiled separately for each trainer supported their effectiveness, as 13 of the 14 trainers had mean scores on their evaluations above 4.4 on the five-point scale. Participant comments on the training supported this with remarks such as, “I have been to much training over the years. This was undoubtedly one of the best – hats off and thanks!” while another responded, “In 18 years in education in Georgia, and the last 10 as an administrator, [trainer’s name] has done the best job of any state training I have ever attended.” When critical feedback was provided, most noted that their questions or concerns were not with the trainer, but with the unknown aspects of the system due to its development and pilot nature.

Training evaluations indicated participants appreciated the opportunity to watch videos and practice rating the videos on the TKES rubric. They also saw value in collaborating and discussing ratings with their peers. Practice ratings of the videos gave participants real-

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

world experience with the new TAPS rubric, and allowed them to begin building inter-rater reliability among their peers.

While the evaluators appeared to be prepared to effectively implement the TKES process, additional data from the teachers needs to be collected and analyzed to fully determine the success of the pilot. Without this analysis, the key participants in the pilot, the teachers, do not have an opportunity to convey their perception of the evaluators' preparedness.

TKES Implementation: Teachers

As previously mentioned, school leaders conducted Teacher Keys Effectiveness System (TKES) orientation trainings, which led to a vast difference in preparation at the teacher level for implementation of TKES. Some teachers were comfortable with the expectations, standards, and process, while others were anxious, uneasy, and had limited knowledge of TKES. During a focus group interview, one group of six teachers from the same school relayed their leader used an interactive approach to the TKES orientation. As a result, teachers were engaged and this principal turned the orientation into a type of a game that was both enjoyable and informative. They left the training feeling like they had a good understanding of the process and content. Others, upon hearing this, voiced their concerns about the inconsistencies of the training from building to building. While many teachers initially felt overwhelmed by the training, they also expressed this sentiment, "At first the orientation was a bit overwhelming, but at the end, it sort of relieved my apprehension when he said do exactly what you do in your classroom." Another said, "I agree. It was overwhelming and I was unsure of what to expect. But then we realized it was what we already did."

The consensus among teachers was there was too much information to remember and they recommended more sessions with smaller amounts of information conveyed. Most of the focus group participants found the TKES handbook to be helpful, well organized, and thorough. Some of the teachers, who were not core academic teachers (band, ROTC, ESOL, foreign language, and some special education), stated they did not see how the standards applied to them.

Data from a survey question about what "support was needed for TKES" indicated one of the themes emerging from the 1,514 surveys was related to effective training. Respondents requested a more hands-on, organized orientation that included demonstrations and videos of teaching, as well as a training portfolio and time to practice, while getting feedback to integrate into future implementation. Survey respondents requested training on various components related to TKES including: 1) administrator training to properly evaluate and provide feedback to teachers; 2) teacher training on how to complete assessments and guidelines for completing the documentation; and 3) evaluator training to look for the same indicators across various settings.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

LKES Implementation: District Leaders/Evaluators

Data from the LKES training sessions, page 14, provides evidence that evaluators were comfortable with the LKES components and pleased with the training they received. Comments on the open-ended questions were predominately positive and reflected an understanding of the LKES training.

District leaders/evaluators stated at the end of the pilot that while they support and like the eight standards for the principals, especially the emphasis on principals' being instructional leader, the documentation requirements of LKES made it difficult to implement. They also identified a need for refresher training. There were 1,082 responses to the LKES pilot survey. Training was again identified as a theme that emerged from the surveys, as respondents requested initial and ongoing training for evaluators and communication of clear expectations of the LKES process.

As with the teachers' perspective on their evaluator's preparedness for successful implementation of TKES, the data from the LKES participants, the principals, needs to be included in order to gather a well-rounded perspective on the evaluator's preparation for successful implementation of LKES. Teacher and principal surveys related to both TKES and LKES evaluator preparation would provide worthwhile information that could be used to increase the effectiveness of the training, as well as identify schools and districts needing additional support for implementation.

LKES Implementation: Principals

There appeared to be some correlation between the teachers' understanding of TKES and the principals' understanding of LKES. In both cases, a "train the trainer" process was used to provide information on the orientation and process to the evaluators. Principals received the TKES training directly from GaDOE trainers and had a positive experience. They went back to their schools as the trainers, and vast differences in teachers' orientation trainings occurred. The same process was replicated with LKES training. GaDOE effectively trained district leaders, and then they became LKES trainers to the principals in their districts.

During focus group sessions, principals confirmed they had received orientation training that was well organized, and noted there was a great deal of information to assimilate in one session. Principals agreed that several shorter orientation sessions, rather than one long, comprehensive one, would be more beneficial. Additionally, several leaders who were among the focus group participants suggested a more interactive presentation style would benefit participants and aid in learning about the evaluation system.

Principals that received the two-day TKES training prior to LKES found it provided a foundation upon which to build understanding of LKES. As with TKES, there were questions about LKES that could not be answered by the district leaders/evaluators. Some principals stated they were unsure of the documentation that was needed. Additionally, many focus group participants expressed the belief that follow-up training and coaching from the Georgia Department of Education evaluation system specialist assigned to the school district would be most helpful. LKES survey respondents requested additional

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

training and discussion of the process, including better training for evaluators, coaching sessions, and opportunities to take online refresher courses.

As with TKES, the quality of the LKES district training for principals by the evaluators directly impacts the understanding of the LKES components and process. Research identifies the “train the trainer” model of professional learning as being limited in its effectiveness (Reeves, 2010). While it is difficult, if not almost impossible, to provide professional learning to large numbers of educators, following the pilot the GaDOE needs to strengthen the TKES and LKES training for those being evaluated.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Implementation

The second group of research questions pertains to implementation. The questions seek to determine the opinions of district leaders, principals, and teachers at the end of the pilot and obtain information about the implementation of each component of TKES and LKES. Data from focus groups, surveys, and various reported data were analyzed and used in response to the research questions.

Research Question #3: Did district leaders and teachers implement TKES and LKES systems as intended? How does that vary by sampling method (across school model vs. whole school model)?

Using multiple data points (focus group interviews, surveys, and data analysis of TAPS and LAPS evaluations), there appears to be consistency in TKES and LKES implementation across the districts during the pilot. While implementation is consistent across the participating districts, implementation as intended by GaDOE Teacher and Leader Effectiveness Division does not fully exist. The statistical data from the TAPS and LAPS pilot reveal a much higher percentage of Proficient and Exemplary ratings than would be anticipated with implementation fidelity, as expected by GaDOE. Comments from principals about TKES during focus group interviews confirm some rationale behind the skewed data:

- All of my teachers are exemplary, or they are gone.
- Subjective, unreliable, contentious.
- Didn't use it, it makes my job harder, too much work.
- I follow the law, and TKES came last.
- Gave everyone exemplary to save time.

TAPS data was available for 26 districts. Standards 4, 6, 7, and 10 each had one "not complete," which indicates that no score was given for one teacher on each of those standards. This resulted with a 0.02% difference or a total percentage of 99.98%. The biggest variation was at the application level of the four rating levels of TAPS evaluations. State TAPS data analyzed by the 10 standards indicates that 74.44% of teachers were rated proficient and 19.31% were identified as exemplary for a total of almost 94% of the teachers meeting expectation on the standards.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

State TAPS Ratings by Standard and Level of Evaluation					
Standard # and Name	% Ineffective	% Developing / Needs Improvement	% Proficient	% Exemplary	TOTAL
PS 1: Professional Knowledge	0.22	3.01	74.40	22.38	100
PS 2: Instructional Planning	0.26	5.82	75.86	18.06	100
PS 3: Instructional Strategies	0.15	5.97	72.46	21.42	100
PS 4: Differentiated Instruction	0.70	11.88	72.59	14.82	99.98
PS 5: Assessment Strategies	0.39	6.84	84.07	8.69	100
PS 6: Assessment Uses	0.70	8.50	81.46	9.33	99.98
PS 7: Positive Learning Environment	0.17	4.07	62.65	33.08	99.98
PS 8: Academically Challenging Environment	0.39	7.56	74.42	17.63	100
PS 9: Professionalism	0.11	2.27	70.58	27.04	100
PS 10: Communication	0.13	3.57	75.90	20.68	99.98
TOTAL	0.32	5.95	74.44	19.31	99.99

TAPS Standard 4, Differentiated Instruction, and Standard 6, Assessment Uses, had the highest ineffective percentage of 0.7% for both. Standards 5 and 6, which dealt with Assessment, each had less than 10% rated as exemplary. These lower ratings on differentiation and assessment appear to have a logical connection since assessment results are most often criteria for determining flexible groups for differentiated instruction. Even with these lower ratings on Standards 5 and 6 in the exemplary category, the combined proficient and exemplary ratings for these three standards does not indicate a significantly lower percentage than the other standards. Differentiation, Standard 4, does have a slightly lower overall rating than other performance standards measured by TAPS. Standard 7, Positive Learning Environment, has a noticeably higher exemplary rating (33.08%) than the other nine performance standards.

When TAPS data was analyzed by the two sampling options, *Across School Model* and *Whole School Model*, the results show no significant variation in data from the analysis for all 26 districts combined. The *Whole School Model* districts had higher exemplary ratings (23.17%) than the *Across School Model* sample districts (16.43%). However, the combined proficient and exemplary ratings for the two sampling methods provide identical results of 93.72% as compared to the combined proficient and exemplary ratings of 93.85% for all districts. The ineffective and developing/needs improvement percentages have little difference between the two sampling options.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

TAPS Ratings for Whole School Samplings Districts					
Standard # and Name	% Ineffective	% Developing / Needs Improvement	% Proficient	% Exemplary	TOTAL
PS 1: Professional Knowledge	0.36	3.21	69.03	27.41	100
PS 2: Instructional Planning	0.36	4.99	71.37	23.28	100
PS 3: Instructional Strategies	0.25	5.76	66.23	27.76	100
PS 4: Differentiated Instruction	1.12	10.59	68.47	19.82	100
PS 5: Assessment Strategies	0.61	5.20	83.09	11.11	100
PS 6: Assessment Uses	0.92	8.30	79.32	11.46	100
PS 7: Positive Learning Environment	0.31	5.09	56.34	38.21	100
PS 8: Academically Challenging Environment	0.71	7.28	70.25	21.75	100
PS 9: Professionalism	0.20	2.80	68.21	28.78	100
PS 10: Communication	0.15	4.48	73.20	22.16	100
TOTAL	0.50	5.77	70.55	23.17	100

TAPS Ratings for Across School Samplings Districts					
Standard # and Name	% Ineffective	% Developing / Needs Improvement	% Proficient	% Exemplary	TOTAL
PS 1: Professional Knowledge	0.11	2.86	78.41	18.62	100
PS 2: Instructional Planning	0.19	6.44	79.21	14.17	100
PS 3: Instructional Strategies	0.08	6.13	77.11	16.68	100
PS 4: Differentiated Instruction	0.38	12.87	75.67	11.08	100
PS 5: Assessment Strategies	0.23	8.07	84.81	6.89	100
PS 6: Assessment Uses	0.53	8.64	83.05	7.73	100
PS 7: Positive Learning Environment	0.08	3.31	67.36	29.25	100
PS 8: Academically Challenging Environment	0.15	7.77	77.53	14.55	100
PS 9: Professionalism	0.04	1.87	72.35	25.74	100
PS 10: Communication	0.11	2.89	77.38	19.57	100
TOTAL	0.19	6.09	77.29	16.43	100

Over 220 principals statewide participated in the LKES pilot during the 2011-2012 school year. State level LAPS data was available for 24 districts. The data was comparable to the TAPS performance level ratings for the combined proficient and exemplary levels, with 89.8% of the summative evaluations statewide rated in these two highest levels.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Pilot LAPS Evaluation Data	
Standard	% Proficient and Exemplary
1. Instructional Leadership	90.7
2. School Climate	90.4
3. Planning and Assessment	87.2
4. Organizational Management	93.4
5. Human Resources Management	85.9
6. Teacher / Staff Evaluation	92.5
7. Professionalism	96.0
8. Communication and Community Relations	89.5
TOTAL	89.8

When comparing implementation by the two sampling options, *Across School Model* and *Whole School Model*, the data was analyzed by percentages rather than frequencies. Since there were 19 districts using the *Across School Model*, and only 7 districts using the *Whole School Model*, a numeric comparison would have been invalid. Data indicates a substantial difference in ratings between the two sampling methods. The *Whole School Model* districts' exemplary ratings were noticeably higher than those using the *Across School Model*. Overall, this difference in ratings between the two sampling methods is less meaningful than any corresponding differences might have been in the TAPS pilot due to the fact that 25% of the principals were randomly sampled to participate in the pilot regardless of whether his or her school was participating in the TKES pilot.

Standard	State LAPS Pilot Data - Ratings by Sampling Method							
	% Ineffective		% Developing /Needs Improvement		% Proficient		% Exemplary	
	Across School	Whole School	Across School	Whole School	Across School	Whole School	Across School	Whole School
1. Instructional Leadership	0.61	1.96	7.88	9.80	67.88	49.02	23.64	39.22
2. School Climate	0.60		9.64	7.84	69.28	54.90	20.48	37.25
3. Planning and Assessment		1.96	13.94	9.80	66.06	56.86	20.00	31.37
4. Organizational Management			8.48	1.96	69.70	60.78	21.82	37.25
5. Human Resources Management			12.12	15.69	72.73	56.86	15.15	27.45
6. Teacher / Staff Evaluation		1.96	7.88	5.88	76.97	64.71	15.15	27.45
7. Professionalism			3.64	5.88	82.42	60.78	13.94	33.33
8. Communication and Community Relations	0.60	3.92	11.45	3.92	66.27	49.02	21.69	43.14

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

3a. What are participants' opinions about implementing the LKES and TKES systems as a whole?

Critical data used to answer this research question came from site visits to each of the 26 districts during the pilot. Georgia Department of Education personnel from Teacher and Leader Effectiveness, School Improvement, and Race to the Top Divisions made site visits to all RT3 districts during spring 2012. During each visit, there was a discussion of four reform focus areas: CCGPS Rollout, TKES/LKES Implementation, School Turnaround, and Data Systems/Technology. Site visit minutes from each district were collected and prepared for document analysis.

Based on evidence from site visit reports, as well as information from surveys and focus groups, the majority of district personnel, principals, and teachers support and like TKES and LKES.

Districts

Districts had a strong support base for both TKES and LKES. They like the standardization of the TKES performance standards across the state and the uniformity and consistency of them. Districts believe the TKES standards validate and strengthen good teachers, and will assist in raising the bar for all teachers. District comments about TKES included:

- Standards are good; like that we can see the overall picture of the schools.
- TKES will be rewarding down the road and the rubric is specific.
- Wonderful tool for new teachers to use for growth.

The LKES process was also viewed positively by most at the district level. They appreciate the emphasis on principals being an instructional leader and believe the standards capture the work of the building level principals. Comments included:

- Establishes clear criteria for leaders.
- Aligns with improvement plans.
- Focuses everyone on the research.

Principals

Principals having previously implemented CLASS Keys found the transition to TKES to be easy and they like that TKES has fewer standards. Most principals believe the process, standards, and implementation of TKES will have a positive influence on student achievement. Sample comments from principals included:

- No concerns around implementation, looking forward to implementing 100% next year.
- Process encourages principals to practice together.
- TAPS has allowed for a lot of collaboration amongst administrators. "Great tool."
- There are benefits to TKES. It is more detailed and there are more sources of information. With the GTOI a teacher could appear to be doing fine, but with TKES there is nowhere to hide.

As mentioned earlier in the report, not all principals were positive about TKES. They were candid during the focus group interviews and voiced frustration with the time needed to

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

evaluate the standards, the stress and anxiety for teachers, and the disillusionment with the amount of paperwork. Two principals stated that a teacher in each of their schools quit because of TKES and the stress it caused. One of the principals was adamant that he lost an outstanding teacher because of this process.

Overall, school leaders were also supportive of LKES. They believe the eight standards are sufficient and capture everything that is needed. They also noted some of the things identified by the districts, as well as other comments that provide insight into their perceptions:

- Snapshot, good tool.
- Helps me identify where I am weak and need to improve.
- Fits with the school improvement plan.
- Raised my awareness of effective, research-based practices.
- Validates what I already do.

Again, some principals were negative about LKES. Statements from principals during the focus group interviews included:

- Overwhelming.
- Won't change anything.
- We just do what the superintendent wants.

Teachers

The majority of the teachers saw value in the TKES process and clearly understood the value on classroom practice. Generally, teachers had positive comments about the impact of standards and indicators for improving student growth and achievement. Many teachers envision the alignment of standards and indicators generating opportunities to support school visions, operations, and consistency across the state. Teachers believe it will raise awareness and provide a "road map" for instruction in a quality classroom. By and large, teachers are positive about the performance standard indicators and they like having examples of tangible behaviors of a proficient teacher. Many also used the indicators to self-assess and make changes to their teaching practices. Positive comments about TKES included:

- Good and fair – eliminates bias.
- Collaborative process with open conversations between us (teacher and administrator).
- Enjoyed showing my administrator my documentation. It proved I do my job well.
- Encompassed more awareness of expectations of me as a teacher.
- Good tool, easy to understand.
- Like the evaluation system better than in the past. Teachers want administration to be in the classrooms more often to see instruction as well as to see the students.
- Like the ability to use it (TKES) to focus my practice and use it (TKES) as a road map.

While the majority of the comments and opinions about the TKES and LKES were positive, there were some consistent themes mentioned as concerns or areas needing improvement.

Georgia Department of Education TKES and LKES 2012 Pilot Evaluation Report

District leaders, principals, and teachers identified these common concerns about TKES: differentiated instruction, SLOs, documentation, surveys, and time.

Differentiated Instruction

District leaders, principals, and teachers discussed differentiated instruction in the focus group interviews. Their comments did not have a consistent focus but were all related to Standard 4. Comments included these statements:

- Differentiation is spread out over several standards, rather than being specific to one place.
- Teachers need more clarity and training about differentiated instruction and assessment strategies.
- The indicators of Differentiated Instruction should be more inclusive of all research-based strategies.

Student Learning Objectives (SLOs)

Student Learning Objectives (SLOs) were another area of concern identified by all focus groups, and SLOs were most frequently acknowledged in the open-ended training questions as a topic needing more information. Comments and concerns varied from the time needed to build each SLO to the quality, fairness, and legality of SLOs. Teachers seemed to have the most questions, comments, and the greatest amount of angst about SLOs, but concerns from all groups were noted. Examples from district leaders, principals, and teachers were:

- Inconsistent targets. Some targets were too high, but others were extremely low. Some teachers felt SLO targets are unreachable and unrealistic.
- Worry about consistency and standardization of the SLOs.
- Building SLOs is time-consuming work, which takes time away from instruction.
- Apprehension about the number of SLOs and making sure that they are of good quality.
- Anxieties about SLOs being equitable and legally defensible. Teachers have a say in SLOs, yet teachers have NO say regarding CRCT questions. Also, teachers of non-tested subjects enter their own scores.
- Lack of communication by the district.
- Unclear how SLOs will be used with high achieving students who have little room for growth.

Documentation

Documentation was another common theme among all three levels of participants. While district leaders mentioned documentation less than principals and teachers, they referenced the amount of paperwork and that some principals and teachers were overwhelmed with the expectations. Other district leaders believe that TKES and LKES were difficult to implement because of the required documentation.

Principals were very forthcoming with their opinions about documentation, especially regarding LKES. Many of the participants indicated the documentation is available for the majority of the standards. However, some believe it was an “exercise in collecting paper.” Principals like the standards and indicators as a simple way of knowing the district’s

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

expectations. During the pilot, many commented on the increased visible presence of their supervisors.

Teachers were often confused about what, and how much, documentation must be provided, as well as whether it had to be from the current school year. Application of the documentation was occasionally inconsistent. Some teachers were told they had to have documented evidence from two indicators per standard, while at other schools no specifications were given. There is also confusion among some teachers about how and when standards are evidenced. Several teachers incorrectly believe that all ten standards must be viewed in a single observation.

Surveys

Instructional and climate surveys were also of concern for all groups. District leaders, principals, and teachers expressed concerns about the wording of the questions, the purpose or intent of the results, and the potential for misuse.

Teachers were especially anxious about the instructional survey process. Their greatest concerns were about “vindictive or angry” students using the surveys for “payback”. At one school, teachers who administered the surveys were very worried that students did not carefully consider the questions. They described how some students chose a single answer for each question, with no thought to the question itself. Several teachers mentioned the wording of the survey questions, saying that they had students who did not understand them. Other comments about the instructional surveys included:

- Not developmentally appropriate.
- Unreliable responses.
- Emotional responses by students.

Most teachers saw value in the feedback provided by the surveys, and acknowledged that if the surveys were used to improve instruction, they would be beneficial. One teacher said, “I want immediate feedback so I can incorporate change into the classroom for the current school year.” Others noted that students are their customers and they want to consider their feedback thoughtfully. While many teachers see worth in the surveys for feedback if it is used for evaluative purposes, they are opposed to surveying students regarding implementation.

Recent research associated with the MET (Measures of Effective Practice) project has confirmed that student surveys provide vital information. MET research, funded by the Bill and Melinda Gates Foundation, began in 2009 with the goal of developing fair, useful measures of teacher effectiveness. The preliminary findings from this project have confirmed student perceptions of the “7 C’s (Raudonis, 2012) and a strong correlation with classroom improvement. The “7 C’s” align perfectly with the instructional surveys developed by GaDOE for the pilot and include the following seven practices: (1) teachers caring for students, (2) class control, (3) ability to clarify things, (4) challenge students, (5) captivate students with interesting lessons, (6) confer with students and respect them, and (7) consolidate by ensuring students are understanding (Ferguson, 2010).

District leaders and principals were not as apprehensive about the instructional surveys but they did identify some of the same trepidations as the teachers. In addition, some

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

concerns expressed by teachers about the instructional survey were consistent with principal responses about the climate survey. Generally, the principals like the feedback since most already use some survey instrument with their staff. As with the teachers, many of the principals stated it is a good tool for reflection, and they plan to use the survey data for growth and improvement.

District representatives are in support of obtaining feedback in general. However, they expressed caution about the intent and use of this information. Comments about student (instructional) surveys were:

- Believe student feedback is very important.
- Students need to be instructed on how to give accurate feedback on the surveys.
- Improve the verbiage of the survey questions.
- Students will use them to “get back” at teachers.
- Can be dangerous if used punitively.
- Too much loss of instructional time. In some cases, it took two weeks to administer the surveys.
- Logistics and technology capacity were issues.
- Public perceptions of results could be problematic and cause misunderstandings.
- Concerned over security of surveys.

Observations by principals and district leaders about the climate surveys included:

- Climate surveys were geared toward teachers; not appropriate for support staff.
- Can be useful data for improvement.
- Some staff will use the surveys to punish principals.
- It may be hard to apply the survey in schools with new principals.
- Difficult for teachers to focus on just the principal, when in many cases the administrative team works cohesively as one unit or [members] bring different strengths.
- Hard for teachers to make a judgment about principal if staff is split for evaluations amongst various administrators.

Time

Whenever something new or different is being implemented, more time is needed to become familiar and comfortable with the expectations and processes. The same was true with the piloting of TKES and LKES. Training took time away from the schools and districts. More time was needed to do teacher observations than had previously been required. Collecting documentation, building SLOs, giving feedback, and the short time frame of the pilot were all identified as “time issues.” Principals were especially concerned about time and found having to do two different evaluation programs at the same time stressful. During the focus group interviews, principals consistently mentioned the “overwhelming” and “impossible task” of using TKES with all of the teachers in their school during the upcoming school year.

District leaders and principals also commented on time in their discussion of LKES. The efficacy of the observation was questioned because of the limited amount of time district

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

level observers/evaluators spent in the building. One principal stated, "They (evaluator) need more time in my building to see everything I do every day."

3b. Was TAPS administrated as intended?

At the State level, TAPS data was available from 26 districts. There were 4,589 participants from 542 schools in the pilot of the Teacher Assessment on Performance Standards (TAPS). Almost 107,000 formative and summative assessments were completed. Of these, 70.1% were formative, and 29.9% were summative. TAPS data by the four rating levels of Ineffective, Developing/Needs Improvement, Proficient, and Exemplary was also analyzed and displayed in the chart on page 20.

When TAPS data was analyzed by the 10 standards and four ratings, the areas of Differentiation and Assessment, Standards 4, 5, and 6, had the lowest overall ratings. Standard 7, Positive Learning Environment, and Standard 9, Professionalism, had the highest overall ratings. Standard 1, Professional Knowledge, and Standard 10, Communication, were also rated highly, with over 95% identified as proficient or exemplary.

At the district level, 26 school systems in the state of Georgia participated in the TAPS evaluations. The frequency of evaluations directly correlated to the size of the district. The number of evaluations ranged from 420 in a small district to over 35,000 in larger districts. By looking at the various ratings by level (ineffective, developing/needs improvement, proficient, and exemplary) the following observations were made:

- Eighteen of the 26 districts had at least one ineffective rating, with three of the 18 having only one ineffective rating each.
- Standard 4, Differentiation, and Standard 6, Assessment Uses, had the highest number of ineffective ratings.
- The highest percentage of ineffective ratings given by a district was 11.48%.
- Twenty-five of the 26 school districts had evaluation ratings in the developing/needs improvement level.
- Developing/needs improvement percentages in districts ranged from 0% to 55.56%.
- 5.95% of the evaluations across the state were rated developing/needs improvement.
- Twelve districts had less than 5% of their evaluations in the ineffective and developing/needs improvement categories for all ten standards, while one district had only proficient and exemplary ratings.
- Thirteen districts had over 90% of their ratings in the proficient and exemplary levels for all standards.
- Proficient ratings were the highest percentage for all districts, and ranged from 61.3% to 91.1%.
- Across the 26 districts, 74.4% of the evaluations were proficient.
- Exemplary ratings fluctuated from 0% to 77.78% with the state at 19.41%.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Along with examining the data by performance standards, TAPS observations can also be analyzed by the five domains: (1) Planning, (2) Instructional Delivery, (3) Assessment of and for Learning, (4) Learning Environment, and (5) Professionalism and Communication.

State TAPS Ratings by Domain and Level of Evaluation				
Standard # and Name	% Ineffective	% Developing /Needs Improvement	% Proficient	% Exemplary
Planning	0.24	4.42	75.13	20.22
Instructional Delivery	0.43	8.93	72.53	18.12
Assessment of and for Learning	0.55	7.67	82.77	9.01
Learning Environment	0.28	5.82	68.54	25.36
Professionalism and Communication	0.12	2.92	73.24	23.86
TOTAL	0.32	5.95	74.44	19.41

TAPS data appears to be strongly skewed to the positive, with 93.85% of the ratings proficient or exemplary, compared to 6.27% ineffective and developing/needs improvement. Having this large of percentage of overwhelming positive scores indicates a need for further training or increased monitoring by the GaDOE. However, consideration should be given to factors that potentially influenced the strong positive ratings. The Teacher Assessment on Performance Standards (TAPS) evaluation was part of a parallel evaluation system occurring concurrently with the district's standard evaluation process. Because the principals had to do TAPS classroom assessments in addition to their "regular" evaluation system, extra demands were placed on the principals. They knew this was a pilot and results would not count so, in some cases, it was not taken as seriously or implemented with fidelity. The short time frame for implementation and the newness of the materials, standards, and process should also be taken into account.

Teachers, principals, and district leaders consistently agreed the ten TAPS standards were thorough and complete, and that they captured what teachers do, both in and out of the classroom. Participants largely believe TAPS are valuable and can contribute to student growth and academic achievement in classrooms, schools, districts, and across the state. They concur the focus on TAPS will strengthen instruction in classrooms that have not previously targeted student growth and academic achievement as the number one priority.

If the TAPS standards are to accomplish these instructional and academic changes in classrooms across the state, it is imperative that the rubrics be applied with greater fidelity. TAPS data from the pilot shows positively skewed ratings and it does not represent a balanced implementation process. School leaders need to have a stronger understanding of the standards, and apply them consistently. Principals and teachers must recognize that every teacher can improve in some capacity and truly exemplary teachers represent only a very small percentage of most faculties.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

3c. Was the student survey administrated as intended?

State level survey data was analyzed and compiled into grade level groups: K-2, 3-5, 6-8, and 9-12. The number of survey questions and participant responses to each question varied. There were fifteen questions on the K-2 and 3-5 surveys. The number of participant responses ranged from 21,920 to 23,428 at the K-2 level and 32,528 to 32,770 at the 3-5 grade level. Middle school and high school student surveys had 20 questions each. The number of middle school responses across the state ranged from 41,109 to 42,743 on the various questions, while high school reactions varied from 64,416 to 65,377.

Student survey data for K-2 across the state identified that students are positive about their teachers, with no areas scoring below 73%. The chart below indicates student responses:

State Results: K-2 Student Survey of Instructional Practice	
Survey Items	% Yes
I know what the rules are in my class.	92
My teacher knows a lot about what she is teaching.	89
My teacher is ready to teach every day.	85
My teacher explains things so I understand.	85
I learn new things in my class.	85
I can do the work my teacher gives me.	83
My teacher knows what I am good at.	80
My teacher lets my parents know how I am doing in school.	80
My teacher uses different ways to help me learn.	80
My teacher is eager to learn new things.	78
I can ask and answer questions in my class.	76
My teacher helps me when learning is hard.	75
I am happy when I am in class.	74
My teacher listens to me.	73
My teacher makes learning interesting.	73

The 3-5 grade survey reflects slightly different student perceptions. The state statistics indicate that once again students know the class rules and believe their teacher is knowledgeable, with ratings of 94% and 90%, respectively. A difference was noted between the K-2 and 3-5 surveys. Seven of the 15 items on the 3-5 surveys were rated below 70%. The chart below identifies the percent of student responses for all districts.

State Results: 3-5 Student Survey of Instructional Practice	
Survey Items	% Yes
I know what the rules are in my class.	94
My teacher knows a lot about what is taught.	90
My teacher makes it okay for me to ask questions when I do not understand something.	81
My teacher is prepared and ready for teaching every day.	79
My teacher explains things so I understand.	79

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

State Results: 3-5 Student Survey of Instructional Practice, continued	
Survey Items	% Yes
I learn new things in my class.	78
My teacher uses different ways to teach and help me learn	76
My teacher shows respect to all students.	76
My teacher is enthusiastic and eager to learn.	68
I am able to do the work my teacher gives me.	67
My teacher lets my parents know how I am doing in school.	67
My teacher listens to me.	65
My teacher allows me to show my learning in a variety of ways.	61
My teacher makes class interesting and challenging.	56
My teacher returns my work with helpful comments on it.	48

The 6-8 survey provided students with four response options of: strongly disagree, disagree, agree, and strongly agree. The highest percentage of positive responses was for, "My teacher knows a lot about what is taught," with 57% strongly agree and 37% agree. Other strong positive reactions were noted for, "I am able to do the work my teacher gives me," and, "I learn new things in my class," with 91% of students rating strongly agree and agree on both items. The chart below displays, in rank order, the agree and strongly agree responses to each statement.

State Results: 6-8 Student Survey of Instructional Practice			
Item	% Agree /Strongly Agree	% Agree	% Strongly Agree
My teacher knows a lot about what is taught.	94	37	57
I am able to do the work my teacher gives me.	91	51	40
I learn new things in my class.	91	44	47
My teacher sets high learning standards for the class.	88	43	45
My teacher explains things so I understand.	87	42	45
My teacher makes it okay for me to ask questions when I don't understand something.	86	40	46
My teacher plans activities that help me learn.	86	45	41
My teacher gives me help when I need it.	86	45	41
My teacher prepares materials in advance and has them ready to use.	86	47	39
My teacher gives clear instructions.	85	43	42
My teacher uses lots of different tests, quizzes, and assignments to find my strengths and where I need help.	83	41	42

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

State Results: 6-8 Student Survey of Instructional Practice, continued			
Item	% Agree /Strongly Agree	% Agree	% Strongly Agree
My teacher uses a variety of teaching practices during class.	83	49	34
My teacher is enthusiastic and eager to learn himself/herself.	82	43	39
My teacher uses different ways to teach and help me learn.	82	44	38
My teacher listens to me.	80	40	40
My teacher shows respect to all students.	79	36	43
My teacher handles classroom disruptions well.	77	40	37
My teacher makes class interesting and challenging.	76	40	36
My teacher allows me to demonstrate my learning in a variety of ways.	75	47	28
My teacher lets my parents know how I am doing in school.	72	41	31

The combined agree (38%) and strongly agree (54%) percentages on the high school student survey data had one item that ranked above 90%. Ninety-two percent of the students in grades 9-12 thought, "My teacher has deep knowledge about the subject he/she teaches." As with the middle school surveys, the lowest scoring item pertained to parent contact about student's learning. It was noted that the second lowest rated item dealt with teachers sharing feedback with students about their learning progress. The chart below provides the combined percentages for agree and strongly agree in rank order.

State Results: 9-12 Student Survey of Instructional Practice			
Item	% Agree /Strongly Agree	% Agree	% Strongly Agree
My teacher has deep knowledge about the subject he/she teaches.	92	38	54
My teacher communicates and maintains classroom rules, routines, and procedures.	89	45	44
My teacher is passionate about teaching and has a commitment for student learning.	88	44	44
My teacher shows respect to all students.	86	40	46
My teacher sets high learning standards for the class.	86	46	40
My teacher's lessons have clear and appropriate learning objectives for the subject area taught.	86	48	38
My teacher is approachable and listens to me.	84	42	42
My teacher is enthusiastic and eager to improve teaching practice.	83	44	39
My teacher creates well-organized and well-developed lessons.	83	47	36
My teacher uses a variety of assessments to determine what I have learned.	82	50	32

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

State Results: 9-12 Student Survey of Instructional Practice, continued			
Item	% Agree /Strongly Agree	% Agree	% Strongly Agree
My teacher uses a variety of teaching strategies during class.	81	46	35
My teacher modifies his/her teaching approaches when I don't understand.	80	48	32
My teacher recognizes us as individual learners with varying learning backgrounds, abilities, needs, and preferences.	79	44	35
My teacher uses assessment results to identify my strengths and the areas in which I need help.	79	48	31
My teacher uses a variety of activities and methods to engage me.	78	44	34
My teacher makes the learning experience challenging but rewarding.	78	46	32
My teacher allows me to demonstrate my learning in a variety of ways.	78	49	29
My teacher makes class interesting and challenging.	77	42	35
My teacher shares feedback about my learning progress with my parents and me.	68	42	26
My teacher contacts my parents to involve them in my learning.	63	38	25

3d. Were Student Learning Objectives (SLOs) administrated as intended?

Student Learning Objective (SLO) pilot data was available from 134 different courses/subjects, 757 participants, 160 schools, and 25 school systems across the state. There were a total of 28,151 assessments given among these participants. Overall scores of the participants were: "Did Not Meet," 8,711 (30.9%); "Meets," 5644 (20%); and "Exceeds," 3,111 (46.6%). Data was missing for 685 (2.4%) of the participants, which may be attributed to student absences and transfers.

Undertaking the work of building Student Learning Objectives across the RT3 districts was a massive task, thus the development and implementation of SLOs was conducted in phases.

Phase 1: The training and preliminary work on SLOs began in October 2011 with 37 training sessions, which included approximately 500 educators. The training was designed for each district to learn about the SLO development process and to begin writing SLOs using current district assessments. Each RT3 district chose 10 subjects for which to create an SLO. Courses were self-selected and primarily based on assessments/measures that districts were currently utilizing.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

The Teacher and Leader Effectiveness Division of GaDOE provided each district with an SLO spreadsheet to be completed by May 1, 2012. The spreadsheet requested the following information: (1) Teacher name, (2) Course number, (3) Student names with pre and post assessment scores and growth target, (4) Student performance levels (Did not meet, Meet, Exceed) on their class targets. As previously mentioned, the deadlines were relaxed by the Georgia Department of Education in order to provide districts sufficient time to collect accurate, quality data. Partial spring pilot data was received September 30, 2012.

From the SLO spreadsheet, class and school ratings (Exemplary, Proficient, Needs development, Ineffective) would be aggregated into a total rating on the SLO Evaluation Rubric. SLO scores from the pilot are in the process of being calculated using this rubric. Those SLO calculations will utilize the chart below to determine growth target ratings.

SLO Growth Target Calculations for Teachers and Leaders	
Exemplary	$\geq 90\%$ of students meet/exceed and $\geq 50\%$ exceed
Proficient Rating	$\geq 80\%$ of students meet/exceed
Needs Development	$> 50\%$ and $< 80\%$ of students meet/exceed
Ineffective	$\leq 50\%$ of students meet/exceed

All Phase I SLOs were posted on the RT3 SharePoint website by state course number and district. Well-designed SLOs were filed as exemplars on the SharePoint website, which was accessible only to districts piloting or implementing the Teacher Keys Effectiveness System.

Phase II: The second phase of SLO submissions included 53 additional courses, which were to be developed and submitted by July 2, 2012, and intended for implementation during the 2012-2013 school year. The Phase II courses/subjects included Pre-K and K-3 mathematics and reading/language arts, as well as 32 high school courses. For this phase, frequently taken high school courses needed for graduation were selected, since it would involve the greatest number of teachers of non-tested subjects.

During the implementation of Phase I, it was noted that districts faced significant challenges in developing SLOs, especially for courses where the districts did not have high quality, effective assessments. The RT3 SLO district contacts met in January 2012 and again in February 2012 to consider a collaborative SLO development process, which included collaborative development of these high quality, effective assessments. It was determined that each of the 26 RT3 districts would take the lead with two or three of the 53 Phase II SLOs and would develop the SLO with pre and post-assessments. The GaDOE staff provided three days of professional learning to 25 districts on the qualities of effective assessments. Several districts committed to collaborating with other districts on identified subject SLOs and agreed to involve higher education experts. The resulting SLOs and assessments became the Georgia Public Domain SLOs and are posted on SharePoint for all Race to the Top and piloting districts to consider. For the 2012-2013 school year, all RT3 districts were required to create their own SLOs, or adapt or adopt the public domain SLOs.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Phase III of the SLO development process will be initiated during the 2012-2013 school year and will reflect information learned during the data analysis of all completed SLO assessments.

Pilot data analysis of SLO frequency and percent by course identified the following variances:

- Seven courses [Advance Mixed Chorus (+9); Advance Placement English Language; CP Chemistry; Guitar Techniques (Grade 7); Honors Chemistry; Literature and Composition (Grade 10); and Theatre Arts (Grade 8)] only had scores in the "Exemplary" level.
- Visual Arts/Comprehensive had missing data for 31 students and no other available data.
- Eight courses [Grade 1 Mathematics (EIP), Grade 1 Social Studies, Grade 2 Reading (EIP), Grade 5 Mathematics, Grade 5 Theater Arts, Exploring Agriculture Education (G), Family and Consumer Science/ Grade and Visual Arts] had data in the "Does Not Meet" and "Exceeds" levels, but no data for the "Meets" level.
- Missing data were noted for 390 files that were not specifically tied to a course. Other missing data were found in the following courses:
 - Language Arts - British Literature/Composition (29); Language Arts, Grade 1 (2), Grade 2 (38), Grade 3 (48), and Kindergarten (3); 10th Grade Lit and Comp (5); World Lit and Comp (29),
 - Math - Mathematics (1), Math III Advance Algebra (3), Grade 1 (10), Grade 2 (24), Grade 3 (66), and Kindergarten (4).
 - Reading - Reading (2), Reading, Grade 1 (28), Grade 2 (23), Grade 3 (41), and Kindergarten (29)
 - Science - Science (1); Science, Grade 3 (14), Grade 4 (3), and Grade 5 (1)
 - Other missing data were found in MCAP (6); Music General, Grade 5 (1); Personal Fitness (39); Weight training (1); and Physical Education, Grade 4 (7), Grade 5 (2), Grade 6 (76), and Grade 7 (41); RCBM3 (31); Spanish 1 (19), Visual Arts / Comprehension (31), Computer Applications (1); Family and Consumer Science (18).
- The courses with the highest frequencies of SLO participation were in the lower elementary grades, K-3: Language Arts, (2136); Mathematics, (6801); and Reading, (5630).

The chart below provides student level SLO data by frequency and percentages of "Does Not Meet", "Meets", and "Exceeds." Overall, nine course titles were identified as being the same subject, but data had been collected separately because of a slight difference in the wording of the title (e.g., French I and FRENCH I). The data from these courses have been combined to reflect one course, rather than two or three.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

SLO Course	State SLO Data by Course and Rating					
	Did Not Meet		Meets		Exceeds	
	#	%	#	%	#	%
Access LA					5	100
Advance Mixed Chorus I (Grades 9-1					8	100
Advanced Placement English Language	23	31.1	44	59.5	7	9.5
Advanced Placement History of Art	8	40.0	4	20.0	8	40.0
Advanced Placement Music Theory	7	33.3	8	38.1	6	28.6
Advanced Placement Psychology	90	84.9	3	2.8	13	12.3
Advanced Placement World History	2	10.0	1	5.0	17	85.0
Advanced Weight Training	1	1.3	77	98.7		
American Government	3	3.5	3	3.5	80	93.0
American Government/Civics	28	25.9	42	38.9	38	35.2
AP FRENCH					6	100
Application of Therapeutic Service					25	100
Band/Grade 8	23	50.0	9	19.6	14	30.4
BEG CHORUS 6	7	3.3	198	92.5	9	4.2
Beginning Band I/Grades 9-12	5	14.3	2	5.7	28	80.0
Beginning Band/Grade 7	13	12.0	9	8.3	86	79.6
Beginning Choral Ensemble					17	100
Beginning Choral Ensemble I			5	21.7	18	78.3
Beginning Chorus (Grade 8)	7	26.9	6	23.1	13	50.0
Beginning Chorus/Grade 6	22	8.0	215	77.9	39	14.1
Beginning Orchestra	13	65.0	1	5.0	6	30.0
British Literature/Composition (29 Missing = 17.1%)	82	48.2	3	1.8	56	32.9
Business and Computer Science	35	92.1	3	7.9		
Business and Computer Science/Grad	24	17.3	19	13.7	96	69.1
Carpentry I	1	7.7	12	92.3		
Chemistry	1	2.0	2	4.0	47	94.0
Chemistry I	60	44.8	11	8.2	63	47.0
Computer Applications I (1 Missing = 2.4%)	7	16.7	2	4.8	32	76.2
CP Chemistry					28	100
Environmental Science	53	65.4	4	4.9	24	29.6
Examining the Teaching Profession	2	7.7	3	11.5	21	80.8
Exploring Agriculture Education (G	3	4.2			69	95.8
Family and Consumer Science/Grade (18 Missing =17.6%)	2	2.0			82	80.4
FOOD/NUTR LIFE	156	43.2	15	4.2	190	52.6
Forensic Science	18	16.8	56	52.3	33	30.8
French I (Missing 1 = 0.5%)	26	12.2	77	36.2	109	51.2
GERM I (Missing 7 = 4.7%)	9	6.0	27	18.1	106	71.1
Gifted Resource Class/Grade 3	7	13.5	1	1.9	44	84.6
Gifted Resource Class/Grade 5			6	75	2	25.0
GPS Pre-Calculus	3	2.2	15	10.9	120	87.0
Guitar Techniques/Grade 7					14	100
Health 6	166	50.0	19	5.7	150	44.8
HON FREN I	4	7.8	12	23.5	35	68.6
HON SPAN I	5	2.9	29	16.9	138	80.2

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

SLO Course	State SLO Data by Course and Rating, continued					
	Did Not Meet		Meets		Exceeds	
	#	%	#	%	#	%
Honor's Chemistry					16	100
Human Anatomy/Physiology	49	39.5	16	12.9	59	47.6
Invention and Innovation/Grade 7	7	12.5	4	7.1	46	80.4
Language Arts/Grad 1 (2 Missing = 0.2%)	241	34.8	68	9.8	382	55.1
Language Arts/Grade 2 (38 Missing = 7%)	220	40.5	67	12.3	218	40.1
Language Arts/Grade 3 (48 Missing = 10.6%)	189	41.7	57	12.6	159	35.1
Language Arts/Grade K (3 Missing = 0.7%)	141	30.6	67	14.5	250	54.2
Mathematics (1 Missing = 2.7%)	18	48.6	6	16.2	12	32.4
Mathematics III- Advanced Algebra/ (3 Missing=1.6%)	44	22.8	9	4.7	137	71.0
Mathematics IV-Pre-Calculus - Trig	3	4.4	35	51.5	28	41.2
Mathematics/Grade 1 (10 Missing = 0.5%)	641	32.4	494	25	833	42.1
Mathematics/Grade 1 EIP	3	25.0			9	75.0
Mathematics /Grade 2 (24 Missing = 1.1%)	751	36	379	18.2	930	44.6
Mathematics/Grade 2 EIP	3	23.1	1	7.7	9	69.2
Mathematics/Grade 3 (65 Missing = 2.7%)	792	33.4	225	9.5	1287	54.3
Mathematics/Grade 3 EIP			1	11.1	8	88.9
Mathematics/Grade 4			3	14.3	18	85.7
Mathematics/Grade 5	69	90.8			7	9.2
Mathematics/Grade 7			1	3.2	30	96.8
Mathematics/Grade 8	6	10.3	4	6.9	48	82.8
Mathematics/Grade K (4 Missing = 1%)	110	28.3	100	25.7	175	45.0
MCAP (6 Missing = 12.2%)	24	49.0	8	16.3	11	22.4
MUS/2	240	19.8	491	40.5	480	39.6
Music, General/Grade	3	17.6	7	41.2	7	41.2
Music, General/Grade 4			7	16.7	35	83.3
Music, General/Grade 5 (1 Missing = 1%)	56	56.6	20	20.2	22	22.2
Music, General/Grade K	9	9.1	18	18.2	72	72.7
Painting and Refinishing	6	22.2	8	29.6	13	48.1
PDP American Government	8	19.5	3	7.3	30	73.2
PDP Chemistry	3	8.1	1	2.7	33	89.2
Personal Fitness (39 Missing = 4.7%)	182	22.0	367	44.3	240	29.0
Physical Education/Grade 4 (7 Missing = 1%)	241	33.9	113	15.9	350	49.2
Physical Education/Grade 5 (2 Missing = 0.5%)	102	24.2	109	25.8	209	49.5
Physical Education/Grade 6 (76 Missing = 22.1%)	32	9.3	98	28.6	137	39.9
Physical Education/Grade 7 (41 Missing = 23.6%)	64	36.8	60	34.5	9	5.2
Physical Education/Grade 8	47	28.3	76	45.8	43	25.9
Physics I	221	82.5	12	4.5	35	13.1
Piano Techniques/Grade 8	1	4.2	5	20.8	18	75
RCBM3 (32 Missing = 20.8%)	90	58.4	2	1.3	30	19.5
Reading (2 Missing = 11.1%)	16	88.9				
Reading /Grade 1 (28 Missing = 1.6%)	510	29.3	233	13.4	968	55.7
Reading/Grade 2 (23 Missing = 1.8%)	509	40.5	121	9.6	605	48.1
Reading/Grade 2 EIP	10	83.3			2	16.7
Reading/Grade 3 (41 Missing = 2.3%)	713	39.5	100	5.5	949	52.6

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

SLO Course	State SLO Data by Course and Rating, continued					
	Did Not Meet		Meets		Exceeds	
	#	%	#	%	#	%
Reading/Grade 3 EIP	6	46.2	1	7.7	6	46.2
Reading/Grade K (29 Missing = 3.1%)	197	21.7	329	36.3	352	38.8
Science (1 Missing = 4.3%)	22	96.7				
Science/Grade 1	36	59.7	6	9.0	25	37.3
Science/Grade 3 (14 Missing = 3.2%)	189	43.2	98	22.4	136	31.1
Science/Grade 4 (3 Missing = 1.9%)	51	32.5	20	12.7	83	52.9
Science/Grade 5 (1 Missing = 0.6%)	61	36.7	25	14.6	84	49.1
Social Studies/Grade 1	18	38.3	10	21.3	19	40.4
Social Studies/Grade 3	3	15.0			17	85.0
Social Studies/Grade 5	4	4.8	4	4.8	76	90.5
SPAN I (19 Missing = 2.3%)	134	16.0	208	24.9	476	56.9
SPAN I GF			15	36.6	26	63.4
SPAN LV1 7	4	7.3	47	85.5	4	7.3
SPAN LV1 8	2	3.0	64	97.0		
Spanish II	69	19.7	215	61.4	66	18.9
Spanish III	24	25.3	9	9.5	62	65.3
Study Skills III					1	100
Tenth Grade Literature and Composi (5 Missing = 3.3%)	22	14.8	8	5.4	113	76.4
Theatre Arts (Grade 8)					27	100
Theatre Arts- Grade 5	3	5.6			51	94.4
TWO-D DESIGN	6	4.8	58	46.4	61	48.8
V ARTS 6	17	15.6	64	58.7	28	25.7
Visual Arts/Comprehensive (31 Missing = 100%)						
Visual Arts/Comprehensive II	5	16.7	24	80.0	1	3.3
Visual Arts/Fibers I	21	58.3	3	8.3	12	33.3
Visual Arts/Grade &	26	40.0			39	60.0
Visual Arts/Grade 2	2	3.4	7	11.9	50	84.7
Visual Arts/Grade 5	198	85.7	6	2.6	27	11.7
Visual Arts/Grade 8	13	22.0	7	11.9	39	66.1
Weight Training (1 Missing = 0.6%)	38	24.4	74	47.4	43	27.6
World History	121	33.7	75	20.9	163	45.4
World Literature/Composition (29 Missing = 55.8%)	21	40.4	1	1.9	1	1.9
Zoology	3	13.6	1	4.5	18	81.8
Other MISSING Data	112	28.7	136	34.9	142	36.4

Student Learning Objectives (SLOs) were one of the areas of great concern to district leaders, principals, and especially teachers. While there was an agreement from all groups that accountability should be equitable across all courses and disciplines, there was much trepidation about the SLO documents, targets, and process. Because the SLO assessments were primarily district-developed, there were concerns about quality, standardization, consistency, and confidentiality. District leaders and principals pointed out that state assessments (CRCT and EOCT) are standardized and consistent across the state, and teachers do not have prior access or knowledge of the questions. Quality was viewed as an issue because of the number of SLOs that had to be developed in a very short time period,

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

and the level of expertise among the developers was inconsistent. Another anxiety was the equitability across the SLOs and whether they would be defensible.

While the GaDOE reviewed and approved the SLOs, as well as identified exemplars, there were still comments on surveys, during site visits, and in the focus group interviews that referenced concerns about the SLO assessments. From the teachers' perspective, SLOs were most often viewed as additional work (e.g., SLO development, target setting, and in the documentation required for reporting formative and summative results). Many teachers appeared to see little value in the SLO assessments and did not feel the SLOs represented their course content.

For the teachers who were directly involved in the SLO pilot, there was a stronger belief that SLOs would show student growth, provide consistency of instruction across the district, and hold everyone accountable. Teachers not involved in student learning objectives (SLOs) during the pilot had little knowledge about them and were less likely to be positive toward them. Teachers who had SLOs for their courses were minimally aware of why they had them and how they are to be used. A lack of communication from the district level on SLOs was mentioned and does appear to be an area needing improvement. Teachers also requested more training on writing, interpreting, and using SLOs.

Teachers are concerned that SLOs are developed locally rather than by the state. Again, subjectivity was their main concern. They want to be assured that SLOs are equitable and some teachers believe the SLO targets are unreachable and unrealistic. There were questions about how SLOs will be used to measure the growth of high achieving students that teachers believe had little room for growth.

Several middle school teachers thought SLOs were applicable only to high school students. Others, at all educational levels, had no idea to whom SLOs would apply. Understanding SLOs, their intent, and how the data will be used needs clarification and explanation for all groups, but especially for teachers.

3e. Was LAPS administrated as intended?

LAPS data were available for over 220 participants in the state of Georgia. Formative and summative LAPS data was collected during the pilot on the eight LAPS Performance Standards.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

State LAPS Pilot Data - Formative Ratings								
Standard	Ineffective		Developing /Needs Improvement		Proficient		Exemplary	
	#	%	#	%	#	%	#	%
1. Instructional Leadership	2	0.88	23	10.13	150	66.08	52	22.91
2. School Climate	1	0.44	22	9.61	158	69.00	48	20.96
3. Planning and Assessment	1	0.44	30	13.22	151	66.52	45	19.82
4. Organizational Management			21	9.21	159	69.74	48	21.05
5. Human Resources Management	1	0.44	33	14.47	157	68.86	37	16.23
6. Teacher / Staff Evaluation	1	0.44	20	8.77	172	75.44	35	15.35
7. Professionalism			17	7.46	174	76.32	37	16.23
8. Communication and Community Relations	3	1.31	25	10.92	147	64.19	54	23.58
TOTAL	9	0.49	191	10.47	1268	69.52	356	19.52

Formative evaluations were predominantly positive, with 89.04% of the total ratings being proficient or exemplary. Data reflects that no ineffective ratings were given on Standard 4, Organizational Management, and Standard 7, Professionalism. These same standards also had the highest combined percentage of proficient and exemplary ratings.

State LAPS Pilot Data - Summative Ratings								
Standard	Ineffective		Developing /Needs Improvement		Proficient		Exemplary	
	#	%	#	%	#	%	#	%
1. Instructional Leadership	2	0.88%	19	8.33	148	64.91	59	25.88
2. School Climate	1	0.44%	21	9.17	152	66.38	55	24.02
3. Planning and Assessment	1	0.44%	28	12.28	150	65.79	49	21.49
4. Organizational Management			15	6.58	157	68.86	56	24.56
5. Human Resources Management			32	14.04	156	68.42	40	17.54
6. Teacher / Staff Evaluation	1	0.44%	16	7.02	172	75.44	39	17.11
7. Professionalism			9	3.95	177	77.63	42	18.42
8. Communication and Community Relations	2	0.87%	22	9.61	145	63.32	60	26.20
TOTAL	7	0.32	162	8.87	1257	68.84	400	21.90

Analyzing the amount of change from formative to summative data ratings provides an opportunity to look for areas of progress, as well as standards and processes needing improvement. Based on a decrease of ineffective and developing/needs improvement ratings and an increase in the proficient and exemplary evaluations, all eight standards showed improvement. Numerical and percentage results of the LAPS pilot data have been summarized in the two tables below.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Standard Numbers	LAPS Total by Numbers and Rating Category										
	Ineffective		Developing / Needs Improvement		Proficient		Exemplary		Total Numbers		
	Formative	Summative	Formative	Summative	Formative	Summative	Formative	Summative	Formative	Summative	
1	2	2	23	19	150	148	52	59	227	228	
2	1	1	22	21	158	152	48	55	229	229	
3	1	1	30	28	151	150	45	49	227	228	
4			21	15	159	157	48	56	228	228	
5	1		33	32	157	156	37	40	228	228	
6	1	1	20	16	172	172	35	39	228	228	
7			17	9	174	177	37	42	228	228	
8	3	2	25	22	147	145	54	60	229	229	
Sub -Total	9	7	191	162	1268	1257	356	400	1824	1826	
Total	16		353		2525		756		3650		

Standard Numbers	LAPS Total by Percent and Rating Category								
	Ineffective		Developing / Needs Improvement		Proficient		Exemplary		
	Formative	Summative	Formative	Summative	Formative	Summative	Formative	Summative	
1	0.88	0.88	10.13	8.33	66.08	64.91	22.91	25.88	
2	0.44	0.44	9.61	9.17	69.00	66.38	20.96	24.02	
3	0.44	0.44	13.22	12.28	66.52	65.79	19.82	21.49	
4			9.21	6.58	69.74	68.86	21.05	24.56	
5	0.44		14.47	14.04	68.86	68.42	16.23	17.54	
6	0.44	0.44	8.77	7.02	75.44	75.44	15.35	17.11	
7			7.46	3.95	76.32	77.63	16.23	18.42	
8	1.31	0.87	10.92	9.61	64.19	63.32	23.58	26.20	
Subtotal	0.49	0.38	10.47	8.87	69.52	68.84	19.52	21.90	
Total	0.44		9.67		69.18		20.71		

When comparing the numeric change between the formative and summative ratings in each level, a decrease was shown with ineffective ratings (-2), developing/needs improvement (-29), proficient (-11), while exemplary increased by 44 scores. Comparison of percentages between the formative and summative LAPS assessments indicates the largest improvement was found on Standard 7, Professionalism, with a decrease of 3.51% in the developing/needs improvement performance level. Standard 7 also had the highest combined proficient and exemplary scores. The total combined percent of proficient and exemplary ratings for the summative evaluations increased by 1.7% over the formative assessments.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

The combined overall formative and summative LAPS data from the 24 participating districts across the state that provided usable data shows little difference from the separate analysis of this data. The evaluations are overwhelmingly positive, with 89.8% of the summative ratings in the proficient and exemplary categories.

In reviewing the numeric and percentage breakdowns of the charts above, the following observations were noted.

- Decreases in the ineffective and developing/needs improvement categories ranged from two participants each on Standards 2 and 5 to eight participants on Standard 7.
- Standards 4 and 7 had no ineffective formative or summative ratings.
- Standard 7 had the highest formative and summative scores.
- Standard 3 and 5 had the lowest proficient and exemplary ratings.
- There were more than twice as many exemplary ratings given than ineffective and developing/needs improvement combined. (Ineffective + Developing/Needs Improvement = 369; Exemplary = 756).
- Standard 5 had no ineffective summative ratings.

During the focus group interviews, questions about the Leader Assessment on Performance Standards (LAPS) generated mixed comments from principals. They discussed the efficacy of the observation due to the limited amount of time observers spent in the building. Comments about LAPS ranged from “very positive” to “can’t make sense of the standards”.

Focus group responses from district leaders to LAPS questions were supportive and positive. They like the eight standards for the principals and believe they are representative of what school leaders need to do. They applaud the emphasis on the principal being an instructional leader. Their only concern was the amount of documentation required. District leaders suggested a need for refresher training and a coaching/mentor component.

Survey data completed by pilot districts had 42% of respondents identify LAPS as the component of the Leader Keys Effectiveness System (LKES) having the greatest impact on principal effectiveness as a leader. This was the highest percentage response, and was 18 percentage points above the next option.

The following concerns about LAPS were identified by principals and/or district leaders: (1) lack of comfort with information; (2) unsure of what documentation is necessary; (3) additional stress and added work; and (4) limited amount of time for an observer to capture everything a principal does. Suggestions for improvement included: (1) observers need more time to provide appropriate feedback, and (2) clarify the documentation and what is actually needed.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

3f. Was the instructional staff survey (climate survey) administered as intended?

State response rate on individual questions of the Teacher/Staff Survey (i.e., instructional staff survey, climate survey) ranged from 13,335 to 14,046 respondents. The item with the strongest positive responses by teachers for principals was, "Interested in quality school which provides quality education," with 95% agreeing or strongly agreeing. Other positive response items were, "Keeps current on education research/trends" (93%); and "Keeps teachers informed on communications from Superintendent" (92%).

None of the combined agree/strongly agree ratings were below 70%. The two items, "Treats all teachers fairly," and "Builds/maintains desirable morale level among teachers," had the lowest percentages as noted in the chart below.

Teacher / Staff Survey Results – All Districts			
Survey Question / Indicator	% Agree/ Strongly Agree	% Agree	% Strongly Agree
Interested in Quality School Which Provides Quality Education	95	30	65
Keeps Current on Educations Research/Trends	93	46	47
Keeps Teachers Informed of Communications from Superintendent	92	42	50
Carries Out Evaluation Program as Outlined	91	43	49
Procures Needed Materials/Equipment	91	47	44
Listens to Views of Parents	90	48	42
Keeps Classroom Interruptions to a Minimum	90	42	48
Displays Pleasant Disposition	89	40	49
Supports Teachers in Conference with Students/Parents	89	43	46
Balances Curricular Assignments/Duties	89	50	39
Provides Leadership in the Improvement of Instruction	88	44	44
Initiates Change for Good of Students	87	39	48
Involves Teachers in Developing Bi-Annual School Plan	87	48	39
Gives Constructive Criticism to Teachers in Private	87	44	43
Uses Judgment/Creativity/Logical Thinking	85	41	44
Seeks Teacher Recommendation for In-Service Programs	84	48	36
Maintains Open Lines of Communication	84	37	47
Makes Helpful Recommendation	83	46	37
Assists in Supervision of Students in Halls/Cafeteria	83	39	44
Visits My Classroom as Often as Should	82	44	38
Earns Respect from Teachers	82	38	44
Keeps Paperwork to a Minimum	81	48	33
Involves Teachers in Decision-making	80	42	38
Treats All Teachers Fairly	79	39	40
Builds/maintains Desirable Morale Level Among Teachers	75	38	37

Unlike teachers, who were concerned about the students' responses to the instructional surveys, principals were not anxious about the staff surveys. However, some consistent comments were noted during the focus group interviews when principals were asked how they would use the survey results. Positive comments included:

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

- Will use the survey for reflection.
- Will use it for growth.
- Use survey data for growth and improvement plans for myself.
- Already use a survey data about my performance.

Some comments from principals about how they intended to use the survey results were not as positive:

- Take with a grain of salt.
- Should not be part of the evaluation, just a tool for feedback.
- My job is not to make them happy; it is to improve student achievement.
- District personnel should have to be assessed.
- Principals are not making decisions. We are just doing what we are being told (by district).
- Only useful if you have some measure of autonomy, if not, then the district should be accountable.

Research Question #4: Are the student -teacher data links working as intended?

Information about the student-teacher data links and other technology platforms was collected during the 26 site visits conducted in spring 2012, as well as through pilot surveys. Responses to the questions about data links during the site visit often led to comments about various GaDOE technology platforms. Participants provided information on the SharePoint, TAPS Dashboard, Student Longitudinal Data System (SLDS), and local data initiatives.

During the site visits, very few comments were made about SharePoint and statements about SLDS were encouraging. However, TAPS Dashboard comments were less positive and indicated a need for “tweaks” and revisions:

- Lack of user friendliness, need to do a lot of scrolling, lots of repetition, a whole lot of hand movement.
- Contained names of people who didn't teach or hadn't taught for years.
- Tweaky, weirdness and kinks.
- Hard to get teachers to upload docs electronically.
- Electronic platform is really the only problem we have had with TKES.
- The current system needs to be tweaked as there were complications using the Dashboard.
- Because of the inability to upload documents to the dashboard, we had to create notebooks.
- Document Section of Dashboard: Can't be edited, not connected to performance standards, preloading documents cannot be done.
- We can't get in. User guide was NOT helpful or user friendly. Absurd that I am supposed to be training teachers on the dashboard, but I cannot even get in.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Overall, districts and principals expressed concern and frustration about electronic uploads, the difficulty and challenge of the site, and the inability to see the TAPS Dashboard. District leaders suggested the need for a training site on Dashboard. Principals stated that Dashboard was very challenging and needed improvement. Others liked the online capacity and have purchased iPads for principals to use with the evaluations. They were working within, and across, districts to get data systems “talking” to each other. Multiple respondents stated that integrating additional electronic methods of observation and data collection would be beneficial.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Pilot Verification and Validation

The final group of questions in TKES and LKES pilot evaluation plans relate to verification and validation. The primary purpose of this section of the report is to determine if the pilot was conducted as intended.

The spring 2012 pilot of the TKES and LKES provided an opportunity for the Georgia Department of Education Teacher and Leader Effectiveness Division to measure the soundness of the TKES and LKES instruments and processes, with the intention of using the pilot as an opportunity to learn from the participants and make revisions based on the various data collected. The TKES and LKES pilot involved the participants, providing them with the tools to increase their understanding of teacher and leader quality, while creating situations for improvement, and ultimately changing practices.

Some of questions in this section of the report cannot be answered at this time, or are still in the process of being answered. Factors affecting the current inability to answer these question are: final TEM and LEM for the pilot have not yet been calculated, Student Growth Percentile (SGP) and Achievement Gap Reduction (AGR) data are still being gathered, student achievement data (CRCT and EOCT) have not yet been factored into the teacher scores, and additional data on attendance, achievement gaps, internal consistency, etc. have yet to be calculated. As additional Teacher and Leader Effectiveness evaluation staff is put in place, the final data will be analyzed, and the comparisons will be completed. At that time an addendum to this report will be published.

Pilot data reported during and at the conclusion of the pilot will be used to address the following questions as the analyzed data becomes available.

Research Question #5: What is the distribution of TEM and LEM, as well as individual components?

Research Question #6: How does the distribution of TEM and LEM, and their individual components, vary by sampling method, district, locale, school size, school level, and content area?

Research Question #7: What is the relationship between individual components of the TKES (SGP, SLOs, TAPs, student survey) and LKES (SGP, LAPS, achievement gaps, instructional staff survey, attendance, retention of teachers) systems?

- A. **What is the relationship between the student achievement measures and the measure of professional practice?**
- B. **What is the relationship between SGP and SLO results for teachers that participated in both?**

Much of the data needed to address this question has already been reported in earlier sections of this report. TKES components have been analyzed and separated, with results

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

described on page 19-21 and 28-39. LKES components are discussed on pages 21-22 and 39-44. The key component missing in this analysis is the Student Growth Percentiles, which are in the process of being provided to the Teacher and Leader Effectiveness Division. Additional analysis for alignment across measures will be completed and provided in a future addendum to this report.

Attendance data and achievement gap data are also being collected. Retention of highly effective teachers data is unavailable until the state collects several years of teacher effectiveness measure data.

Two research questions identified in the evaluation plan will not be answered due to unavailable data. Research Question 8 involves a comparison of pilot TEM to other teacher evaluation systems (GTEP, CLASS Keys). GDOE does not have access to these other evaluation systems' data due to privacy laws. Data was not collected by GaDOE during the pilot to answer research question 9, concerning inter-rater reliability of TAPS ratings by principals and trainers. The data is being collected during the 2012-2013 implementation and pilot in order to address this question at the conclusion of the 2012-2013 school year. Research questions 8 and 9 are provided below.

Research Question #8: For schools implementing both systems, what is the relationship between TEM and scores on the existing evaluation system?

Research Question #9: What is the inter-rater reliability of the TAPS instrument? How do principals' ratings compare to trainers' ratings? How does that vary across districts? How does that vary between district and schools that have previously implemented CLASS and Leader Keys?

Partial data for questions 10 and 11 are available and have been analyzed. The responses to these two questions are found below.

Research Question #10: What is the internal consistency (Cronbach's alpha) of the student and instructional staff survey instruments by dimension? Are the intended dimensions demonstrated (confirmatory factor analysis)? What is the relationship between the dimensions?

The five dimensions/domains for TKES were previously discussed on page 29 of this report. They include: (1) Planning, (2) Instructional Delivery, (3) Assessment, (4) Learning Environment, and (5) Professionalism and Communication. Multiple questions on the student's instructional surveys combine to provide data on each dimension/domain. The number of questions associated with each influence the internal consistency. Two

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

dimensions, Planning and Professionalism/Communication are comprised of two questions each, while the Learning Environment has four questions for providing data on this domain.

Analysis using Cronbach measures internal consistency or alpha for Likert scale questionnaires or surveys such as the surveys associated with TKES and LKES by demonstrating how well a group of questions measure a single characteristic (i.e., dimension/domain). In the social sciences an alpha of at least 0.7 is considered acceptable with 0.8 being preferable.

The internal consistency of the student instructional surveys was highest for Grades 9-12 with the Cronbach alphas for all dimensions above 0.9. Surveys for Grades 6-8 had a strong internal consistency with all but one dimension being above 0.9. The lowest alpha was Assessment at 0.85. Based on Cronbach alpha analysis, student instructional surveys at the elementary level appear to have less internal consistency. K-2 had the lowest internal consistency rating with only one dimension, Instructional Delivery, above 0.7. On the 3-5 surveys, two domains, Instructional Delivery and Learning Environment, had alphas above 0.7. The charts below provide information on questions associated with each dimension and the corresponding alpha score.

TKES Instructional Survey Grades K-2		
Dimension	Questions	Cronbach Alpha
Planning	1, 2	0.58
Instructional Delivery	3, 4, 5, 6	0.76
Assessment	7, 8, 9	0.60
Learning Environment	10, 11, 12, 13	0.68
Professionalism & Communication	14, 15	0.59

Three dimensions, Planning, Professionalism and Communication, and Assessment had very close alphas on the K-2 instructional surveys. These represented the second, third and fourth lowest alphas in grades K-12, with the lowest for all grade levels on the 3-5 surveys. Planning in 3-5 had a 0.5 alpha, which was 0.08 less than this domain in K-2 and 0.14 lower than the next 3-5 domains of Professionalism and Communication. On both the K-2 and 3-5 surveys the two dimensions of Planning and Professionalism/Communication were the lowest alphas, while Instructional Delivery was the highest for both.

TKES Instructional Survey Grades 3-5		
Dimension	Questions	Cronbach Alpha
Planning	1, 2	0.50
Instructional Delivery	3, 4, 5, 6	0.77
Assessment	7, 8, 9	0.68
Learning Environment	10, 11, 12, 13	0.73
Professionalism & Communication	14, 15	0.64

Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report

Alpha comparisons for the grades 6-8 increase noticeably when compared to the K-2 and 3-5 surveys. The greatest increase was seen in the Planning dimension, with the 6-8 surveys 0.34 higher. The Assessment domain had the least change in alpha at 0.17.

TKES Instructional Survey Grades 6-8		
Dimension	Questions	Cronbach Alpha
Planning	1, 2	0.94
Instructional Delivery	3, 4, 5, 6	0.95
Assessment	7, 8, 9	0.85
Learning Environment	10, 11, 12, 13	0.93
Professionalism & Communication	14, 15	0.93

All dimensions at the 9-12 level were extremely strong with alphas ranging from 0.94 to 0.97. The lowest domain at the 9-12 level was Professionalism and Communication at 0.94.

TKES Instructional Survey Grades 9-12		
Dimension	Questions	Cronbach Alpha
Planning	1, 2	0.97
Instructional Delivery	3, 4, 5, 6	0.97
Assessment	7, 8, 9	0.96
Learning Environment	10, 11, 12, 13	0.95
Professionalism & Communication	14, 15	0.94

Based on the overall analysis of the TKES instructional survey completed by students for their teachers the dimensions in the upper grades (6-12) have robust internal consistency. The elementary grades (K-5) have lower Cronbach alphas. The variance in these scores may be related to the number of choices available for each question. The K-5 survey had three choices (No, Sometimes, Yes), while the 6-12 surveys had four options (Strongly Disagree, Disagree, Agree, Strongly Agree). Another difference in the surveys was the number of questions. K-2 surveys had 15 questions compared to 20 questions on the 6-12 survey. While these differences must be considered, it is important for GaDOE to explore the internal consistency of the surveys in order to increase the consistency of the alpha scores.

Cronbach alpha data reveal strong internal consistency for the LKES climate survey completed by teachers involved in the spring pilot. As mentioned with TKES, any alpha above 0.8 is considered to be satisfactory. The eight LKES professional standards had alphas ranging from 0.87 to 0.93. There were 25 total questions on the LKES survey and each professional standard was aligned to a varied number of survey questions. The professional standard Teacher/Staff Evaluation had three questions while the standards of Instructional Leadership, Organizational Management, and Professionalism had five each. Unlike the TKES survey questions that were specific with one domain, the LKES questions often provided information on multiple standards. Survey questions 3, 4, 5, 7, 10, 11, 18, 20, and 22 afford survey data relevant to two professional standards. The chart below offers information on the number of questions and alpha for each professional standard.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

LKES Climate Survey		
Professional Standard	Questions	Cronbach Alpha
Instructional Leadership	1, 3, 10, 13, 18	0.89
School Climate	7, 11, 14, 22	0.90
Planning and Assessment	6, 7, 10, 20	0.92
Organizational Management	3, 5, 8, 9, 16	0.88
Human Resources Management	4, 15, 18, 22	0.91
Teacher/Staff Evaluation	4, 5, 21	0.87
Professionalism	11, 12, 19, 24, 25	0.93
Communications and Community	2, 17, 20, 23	0.88

Research Question #11: How did the quality of the SLOs vary across districts?

There is not a large variation in how SLOs were written across districts, but there is variation in the rigor of growth targets. All districts received regional training on SLO development, which reduced the variation in how SLOs are written. Most of them are well written and clear to the evaluator. The GaDOE staff has compared the quality of the SLOs to examples from other states and found SLOs developed as part of the pilot were fairly good. Pilot SLOs had to be approved by GaDOE staff using a common rubric. When assessment parts were missing or there was ambiguity, the SLO was returned to the district for required or recommended revisions.

What does vary from district to district is the rigor of the growth targets. Some districts chose more realistic, rigorous growth targets, which increased expectations for students and teachers, while others chose to take a more conservative approach, to not threaten teachers and students.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Conclusion

The purpose of the Teacher Keys Effectiveness System (TKES) and Leader Keys Effectiveness System (LKES) pilot conducted by the Georgia Department of Education was to test TKES and LKES instruments, training, and administration procedures. The pilot allowed GaDOE to collect preliminary data to inform the revision of the tools, processes and practices prior to full implementation and create awareness about TKES and LKES among teachers, principals, and district office personnel. During the spring 2012 pilot, twenty-six Race to the Top (RT3) districts served as the sample group.

Data collected during the pilot included observational information on TKES and LKES training, comments from focus groups, survey statistics, analysis of assorted documents, and the review of reported data. Available data was analyzed and triangulated to address three key components of the pilot: training, implementation and validation. This report has focused on quantitative and qualitative state level data to answer the questions outlined in the evaluation plan.

TKES professional learning was provided to district and school level evaluators from all RT3 districts by GaDOE trainers during fall 2011. Workshop evaluations from the 56 sessions across the state were overwhelmingly positive. Workshop participants identified presenters' expertise, training material and the quality of learning as exceptional. Focus group interview results during May 2012 concur with the workshop ratings and identified similar strengths. Evidence from training data strongly supports that district leaders and principals had an appropriate understanding of TKES.

Training for evaluators was successful; however, the training for teachers lacked consistency across the sample group. Some district and school leaders responsible for redelivering the TKES training to the teachers in their districts and schools did not provide adequate training to teachers. According to focus group comments, the TKES orientation for the teachers ranged from "no training" to "excellent." The variation in training to classroom teachers is an area of concern.

Separate training and follow-up support were given on Student Learning Objectives (SLOs) to district leaders, principals, and schools. The training and support for SLOs reflect positive data. However, the pilot data regularly showed concerns and frustration with SLO expectations, improvement targets, and development.

TKES implementation received predominantly positive responses during the pilot. In analyzing the three key components of TKES, teachers and leaders strongly agree that the Teacher Assessment on Performance Standards (TAPS) is a good indicator of effective classroom instruction and processes. Numerous teachers use the standards and indicators for self-assessment and as a guide for improvement. The majority of teachers see great value in the TAPS process and potential impact on quality instruction. For districts previously implementing Classroom Analysis of State Standards (CLASS Keys), there were affirmative statements about the reduced number of TAPS standards and the continual focus on quality classrooms.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Encouraging responses to TAPS standards at the teacher, principal and district levels influenced successful implementation. Overall, the statewide implementation of TAPS was efficacious; yet, the implementation did not fully align with GaDOE expectations. TAPS data was skewed to the positive with 93.85% proficient/exemplary ratings, compared to 6.27% ineffective and developing/needs improvement. The unusually large percentage of highest ranked scores on TAPS may reflect a need for further training and/or increased monitoring by the GaDOE.

TKES instructional surveys received assorted reviews from teachers, leaders, and district personnel. Most teachers saw value in the student-completed instructional survey and recognize student feedback can assist them in improving instruction. However, there was apprehension among participants at all levels (classroom, school, and district) about the possible misuse by students. Concerns included students not taking the survey seriously, not reading the questions, or viewing it as an opportunity for reprisal.

Cronbach alpha data on TKES surveys found strong internal consistency for grades 6-12, however, grades K-5 had noticeably lower alpha scores. While the variance in the alpha scores across grade levels may be related to the number of response choices and the number of questions on the survey, it is important for GaDOE to reevaluate the surveys and make identifiable revisions to the lower grade surveys in order to increase the internal consistency for each dimension.

Another component of TKES, student growth percentiles (SGPs), had limited emphasis during the pilot. SGPs for tested subjects were still in the process of development by the assessment division of GaDOE and information about them was incomplete. The limited focus on SGPs during the pilot had minimal impact on TKES implementation. For teachers of tested subjects, the SGPs were viewed as a method of accountability similar to previous Adequate Yearly Progress (AYP) goals. For teachers of non-tested subjects, the ability to impact student learning was related to Student Learning Objective (SLO) growth. Again, SLO development, targets, and purpose were a source of frustration and concern for participants.

Leader Keys Effectiveness System (LKES) training began in January 2012. The results of the training were similar to TKES training evaluations in that the ratings and comments were extremely positive. Participants praised the trainers' expertise, approved of the LKES materials, and affirmed their level of understanding had greatly increased. LKES training appeared to be more consistently redelivered to principals than TKES training. Principals noted that some training sessions were rushed and contained too much information at one time. Similar to TKES, inadequate redelivery influenced participant understanding and support, as well as consistency of implementation. Evaluators' preparedness for orientation training at both the TKES and LKES level is an area that needs to be strengthened.

Effective LKES implementation was contingent on successful execution of the various components. Leader Assessment on Performance Standards (LAPS) was perceived

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

positively by district leaders/evaluators. They supported the eight standards for school leaders and the emphasis on being instructional leaders. Most principals acknowledged the purpose and benefit of the LAPS standards, and believe they emphasize the critical aspects of their job. Both district leaders and principals were concerned with the documentation requirements of LKES and the limited time evaluators were actually observing the principal.

LAPS data appeared to be consistently implemented across the participating districts, but the results were skewed to the positive. -Progress from the formative to the summative assessment was noted on all eight standards with final summative scores being 89.89% proficient/exemplary. Fidelity of implementation, as intended by GaDOE, appears to be a concern during the pilot.

Teachers completed LKES climate surveys during the pilot. School leaders were supportive of the survey, as most were accustomed to doing annual surveys with their staff. They also understood the value of using survey results for improvement. Some stated the survey was not relevant to all staff and a few questions addressed aspects viewed as not being within the principal's control. LKES climate survey data demonstrated strong internal consistency of each professional standard with Cronbach alpha scores ranging from 0.87 to 0.93.

Student Growth and Academic Achievement (SGP, SLO) make up the final component of LKES. Comments related to limited information about SGP when discussing TKES apply to LKES. The perspectives of the classroom teacher about SLOs were consistent with the viewpoint of the principals. Data associated with Achievement Gap Reduction are still being analyzed. This information will be included in an addendum to this report.

Based on data currently available for TKES and LKES components, the pilot was effective and affords significant results. Analyzed data offered evidence of successful training and implementation that provide validation for the pilot. While there are areas of concern with some components of TKES and LKES, the short time frame for the pilot, the inability to control the external components of 26 districts, and the inconsistencies resulting from diverse rural, suburban, and urban districts of varying size must also be considered. The implementation of parallel teacher and leader evaluations was another consideration. Principals and district leaders continued with their previous evaluation process during the pilot. Seventeen districts used GTEP (Georgia Teacher Evaluation Program) along with TKES and LKES, eight utilized CLASS Keys, and one district used a hybrid version of the two. Doing simultaneous evaluations increased the paperwork, time expectations, and amount of work to be done. All aspects of TKES and LKES were piloted and pilot data will help to guide revision and improvement.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

As GaDOE makes adjustments to TKES and LKES, the following recommendations need to be considered:

- **Strengthen the fidelity of implementation for TAPS and LAPS rubrics.** If the TAPS and LAPS standards are to accomplish the instructional changes in classrooms and schools across the state, it is imperative the rubrics are applied with consistency and greater balance. District and school leaders need a stronger understanding of the standards, rubrics and exemplars. Training for TAPS and LAPS needs to include a deeper focus on application of the rubric, and participants must reach a specific level of competence at the end of the training. Increasing inter-rater reliability among evaluators will be critical to consistency of implementation statewide.
- **Create consistency in TKES orientation training.** Those being evaluated, teachers, require a strong understanding of the TAP standards and process. To apply the standards without effective orientation to them is an improbable expectation. Orientation videos developed by GaDOE trainers may be one option to help standardize training.
- **Continue to focus on improving the development, application and use of Student Learning Objectives (SLOs).** Because the assessments are ultimately district developed or identified, there are concerns about quality, standardization, consistency, and confidentiality. The limited time period for development, and the inconsistency in the level of expertise across developers, influence the quality of the assessments. It is suggested that GaDOE use the current and new SLOs to develop an item bank of exemplars that can be used by districts. Districts and non-tested subject teachers need training and assistance on how to set appropriate growth targets and use pre-assessments to inform instruction.
- **Assist teachers and leaders with analysis and use of instructional survey data.** Professional learning may be beneficial for teachers and leaders on using the instructional surveys for improvement at the classroom, grade, school, and district level. Over time, teachers will likely become less anxious about students completing the instructional surveys. However, if teachers and leaders do not use the survey data effectively, the potential value is lost.
- **Provide data to districts and schools in a timely manner.** It is understandable that the data related to the pilot would take time to process and return to districts. As previously mentioned, pilot data was received later than anticipated because of the relaxation of initial deadlines and challenges experienced with data collection. Deciding to move the deadlines back was the correct decision for the pilot, schools, and districts; however, the delay in returning TKES and LKES data prevented its use for improvement planning. As processes and procedures are finalized, and as the SGP process is completed, data needs to be returned to the participants as soon as possible.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

The Georgia Department of Education's Teacher and Leader Effectiveness Division has undertaken the monumental task of developing, piloting, implementing, and evaluating TKES and LKES. The positive aspects of the training, materials, processes, and data far outweigh the minor problems that have been identified in this report. With slight revisions and continued analysis of progress, TKES and LKES have the potential to positively influence education in Georgia for years to come.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

References:

- Ferguson, R.F. (2010). Student perceptions of teacher effectiveness. *Harvard University, National Center for Teacher Effectiveness and the Achievement Gap Initiative.* Retrieved from http://www.achievementseminars.com/seminar_series_2008_2009/readings/ferguson_tripod_project.pdf
- Raudonis, L. (2012, November/December). Student surveys: Helpful or harmful? Page One. 32(2), 7-12.
- Reeves, D.B. (2010). Transforming professional development into student results. Alexandra, VA: ASCD.

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Appendix A

Topic Areas for Focus Groups

The moderator's guide questions addressed the following areas of the pilot process:

For Teachers:

1. Orientation
2. Teacher Assessment on Performance Standards (TAPS) Evaluation Process
3. Teacher Performance Standards
4. Teachers Performance Standards' Indicators
5. Value of TKES Standards and Indicators on Student Growth and Academic Achievement
6. Student Learning Objectives
7. TKES Assessment Rubric
8. Surveys of Instructional Practice

For Principals

1. Orientation
2. Leader Assessment on Performance Standards (LAPS) Evaluation Process
3. Leader Performance Standards
4. Leader Performance Standards' Indicators
5. Value of LKES Standards and Indicators on Student Growth and Academic Achievement
6. Impact of LKES Pilot Participation on Job Performance
7. Impact of TKES Pilot Participation on Teacher's Job Performance
8. LKES Assessment Rubric
9. Climate Surveys
10. Teacher Performance Standards' Indicators
11. Documentation Requirements for LKES

For Districts

1. Teacher Assessment on Performance Standards (TAPS)
2. Leader Assessment on Performance Standards (LAPS)
3. Teacher Performance Standards' Indicators
4. Leader Performance Standards' Indicators
5. Value of LKES Standards and Indicators on Student Growth and Academic Achievement
6. Surveys of Instructional Practice and Climate Surveys
7. Support and Assistance from Georgia Department of Education

**Georgia Department of Education
TKES and LKES 2012 Pilot Evaluation Report**

Appendix B: District Responses to the End of Pilot Survey

DISTRICT RESPONSES		
School System	Number of Responses	Percent of Total Responses
Atlanta Public Schools	195	8%
Ben Hill County Schools	1	0%
Bibb County Schools	5	0%
Burke County Schools	35	1%
Carrollton City Schools	10	0%
Chatham County Schools	239	9%
Cherokee County Schools	121	5%
Clayton County Schools	231	9%
Dade County Schools	13	1%
DeKalb County Schools	547	21%
Dougherty County Schools	75	3%
Gainesville City Schools	29	1%
Gwinnett County Schools	378	15%
Hall County Schools	1	0%
Henry County Schools	53	2%
Meriwether County Schools	19	1%
Muscogee County Schools	236	9%
Peach County Schools	31	1%
Pulaski County Schools	16	1%
Rabun County Schools	18	1%
Richmond County Schools	207	8%
Rockdale County Schools	5	0%
Spalding County Schools	18	1%
Treutlen County Schools	6	0%
Valdosta City Schools	60	2%
White County Schools	35	1%
TOTAL	2584	100%