

# SUMMARY FINDINGS: 2011 TENNESSEE EDUCATOR EVALUATION SURVEY

**OCTOBER 2011**

MATTHEW J. PEPPER

MARK C. DUNN

TONY PRATT

SUSAN FREEMAN BURNS

MATTHEW G. SPRINGER

Tennessee's Consortium on Research, Evaluation, and Development is responsible for carrying out a detailed, focused program of research as part of Tennessee's Race to the Top grant.

This research report provides a summary of findings from the 2011 Tennessee Educator Evaluation Survey. The views expressed herein do not necessarily reflect those of sponsoring agencies or individuals acknowledged. Any errors remain the sole responsibility of the authors.

Please visit [www.tnconsortium.org](http://www.tnconsortium.org) to learn more about our program of research and recent publications.

# EXECUTIVE SUMMARY

---

A large body of empirical research demonstrates teachers' effectiveness varies widely, perhaps nowhere more so than from classroom to classroom within the same school (Glazerman et al., 2010; Goldhaber et al., 1999; Nye et al., 2004; Sanders & Rivers, 1996; Wright et al., 1997). Differences in teacher quality have been shown to account for a larger share of the differences in student outcomes than any other observable school characteristic (Nye and Hedges, 2004). Despite these differences and their implications for student achievement, most existing systems of teacher evaluation do not effectively differentiate between effective and ineffective teachers. A 2009 report by the New Teacher Project found 99 percent of teachers received the highest rating in binary systems of evaluation, and in evaluations drawing from more than two categories, 94 percent of teachers received one of the top two ratings. Less than 1 percent were ranked as unsatisfactory (Weisberg et al., 2009). This disconnect suggests systems of teacher evaluation that better differentiate between teachers' effectiveness have the potential to become important levers for improving teacher performance, and, in turn, student achievement.

## The Reform Context in Tennessee

The Tennessee Consortium on Research, Evaluation, and Development (Consortium) recently completed an analysis of teacher evaluation models being used in Tennessee during the 2010-2011 school year. The purpose of this investigation was to provide feedback, grounded in relevant research, to inform the design and implementation of Tennessee's new teacher evaluation system established under the First to the Top legislation. With the assistance of state- and district-level administrators, Consortium researchers designed a Teacher Evaluation Field Test, which assessed educator and evaluator perceptions and experiences related to the state's existing model of teacher evaluation, the Framework, and a newly proposed, statewide model for teacher evaluation using the Teacher Advancement Program (TAP) rubric. Concurrently, three other evaluation systems were implemented during the 2010-2011 academic year. One system, TIGER, was implemented by schools in the Association of Independent and Municipal Schools (AIMS) districts, and a second, COACH, was implemented by the Hamilton County Department of Education (HCDE). Both of these alternative teacher evaluation systems were included within the Consortium's broader investigation. A third evaluation system was piloted in the Memphis City Schools, but was not examined for this report.

Requested by the Tennessee Department of Education (TDOE) to research educator perceptions and experiences relative to observation and evaluation generally, and educator evaluation in Tennessee specifically, Consortium researchers attended evaluation training sessions, conducted interviews, and administered the 2011 Tennessee Educator Evaluation Survey to teachers and evaluators. This survey gauged the experiences of educators using the existing state evaluation model, the model being piloted during the evaluation field test, and both the COACH and TIGER models being piloted by the aforementioned LEAs. Survey results and analyses from educators participating in these four teacher evaluation systems is the focus of this report.<sup>1</sup>

## Survey Findings

Survey outcomes suggest all four models were implemented with high levels of fidelity. Responses indicate teachers and administrators, in general, complied with expectations and conducted evaluation procedures in accordance with local and/or state guidelines. The impact of the evaluations on teacher workloads was varied, with field test treatment schools, or those using the Teacher Advancement Program (TAP) Rubric, reporting a slight decrease in workload for teachers, but an increased workload for evaluators. Both evaluators and teachers reported the burden of paperwork generally increased for evaluators in the TAP Rubric, TIGER, and COACH groups.

When asked about perceptions of the purpose of their most recent observation, teachers and evaluators across all four surveyed groups overwhelmingly report the observations were designed both to make judgments about performance and to give feedback to help improve teaching. Again across all four groups, a higher percentage of evaluators than teachers reported their observations were designed to accomplish both objectives.

Teachers were less in agreement concerning whether the evaluation system contributed to greater collegiality and professionalism, with between 51 percent and 61 percent responding 'Agree' or 'Strongly Agree' across all four models. Evaluators, in contrast, reported a more positive view of the impact of evaluations on teacher professionalism. The survey found a similar perception gap concerning teachers' comfort level with being evaluated: 66 percent of teachers claimed to be comfortable with classroom observations while 82 percent of evaluators indicated teacher comfort.

---

<sup>1</sup>The preliminary findings from this research were presented in April 2011 to the Teacher Evaluation Advisory Committee (TEAC), a group of education stakeholders established by the First to the Top legislation and charged with providing recommendations to the state board of education on issues related to the evaluation. Although for the purpose of the Field Test, examining the TAP Rubric group against those using the existing Framework was the Consortium's primary focus, alternative models were included in an effort to provide TEAC members with a broader basis for comparison. (A copy of this presentation is available upon request from the Consortium.)

## Recommendations

To begin, the Consortium recommends that TDOE continue to monitor the implementation of teacher evaluation programs through an annual survey. A number of critical issues have emerged from this survey that should be further examined as teacher evaluation moves forward.

Second, considering the increased workload on evaluators, TDOE should explore technological efficiencies that minimize the burden of both observations and evaluations such as a tablet or laptop evaluation program that automatically communicates with the proposed statewide, centralized system.

Third, the process used to train field test evaluators on utilizing the TAP Rubric should be maintained during statewide scale-up. Evaluators should also be periodically assessed on the consistency of their rating standards.

Fourth, in order to ensure that the reform momentum from the originating year of First to the Top continues, TDOE should make every effort to provide consistent and clear communications with Tennessee teachers and other stakeholders, particularly as it relates to educator evaluation.

TDOE should examine the evaluation model to ensure the TAP Rubric sufficiently allows for variations within teaching responsibilities, and that appropriate adjustments be considered for use with non-classroom positions such as librarians and instructional coaches.

Finally, the department should consider the diffusion of evaluator responsibilities, particularly in the observation process and promote greater engagement on the part of assistant principals, as well as central office leadership and other school-based leaders such as lead teachers or instructional coaches.

# INTRODUCTION

---

In recent years, teacher evaluation has drawn significant attention nationwide. While research demonstrates teacher impact on student achievement can vary considerably, existing systems of teacher evaluation tend to do a poor job of differentiating between teachers' relative effectiveness (Glazerman et al., 2010; Goldhaber et al., 1999; Nye et al., 2004; Sanders & Rivers, 1996; Weisberg et al., 2009; Wright et al., 1997). That these systems are broken is no secret. In 2010, Randi Weingarten, president of the American Federation of Teachers, stated, "with rare exceptions, teacher evaluation procedures are broken—cursory, perfunctory, superficial, and inconsistent" (Weingarten, 2010). United States Secretary of Education Arne Duncan has also sarcastically quipped, "today in our country, 99 percent of our teachers are above average" (Gabriel, 2010), referencing teacher evaluation systems that commonly give the vast majority of teachers favorable ratings.

As a result of the shortcomings of existing systems of teacher evaluation, principals, schools, and districts generally lack the in-depth information necessary for making important personnel decisions about teachers that could significantly impact student achievement. Improved systems of teacher evaluation have the potential to remedy this problem while also providing teachers with useful feedback to improve instruction.

**As a result of the shortcomings of existing systems of teacher evaluation, principals, schools, and districts generally lack the in-depth information necessary for making important personnel decisions about teachers that could significantly impact student achievement.**

## State and National Reform Context

In January 2010, the Tennessee General Assembly passed Senate Bill 5, also known as the First to the Top Act, reforming dozens of areas of state education policy. Spurred by the Federal Race to the Top competition, the ambitious reforms helped Tennessee win a \$501 million award to implement and institutionalize the changes statewide. One of the most ambitious, and contentious, provisions in the law requires all school personnel be evaluated annually, and personnel decisions be based, in part, on these evaluations. The law further specifies that 35 percent of all educators' evaluations be based on value-added student achievement data (as calculated using the Tennessee Value Added Assessment System, TVAAS), and 15 percent on "other measures of student achievement." For teachers, the remaining 50 percent of an evaluation must be based on qualitative measures including, but not limited to, a review of prior

evaluations, personal conferences, and classroom observations. While Tennessee pioneered the use of value-added assessment in the 1990s, and the state has received national praise for its efforts in educational data collection, legislative constraints limited the use of data to inform classroom practice or human capital decisions. Additionally, while the law requires all teachers be evaluated using the guidelines outlined above, only an approximate 40 percent of teachers in Tennessee teach in subjects and grades that produce TVAAS scores, raising a number of questions about the roughly 60 percent of teachers in non-tested grades and subjects.

In an effort to address the so-called qualitative aspect of the evaluation, explore strategies for teachers without value-added data, and give stakeholders a voice during the reform process, the First to the Top Act also created the Teacher Evaluation Advisory Committee (TEAC). Composed of fifteen members, including the Commissioner of Education, state legislators, public school teachers and principals, and other stakeholders, the TEAC's charge was to "develop and recommend to the [Tennessee State Board of Education] guidelines and criteria for the annual evaluation of all teachers and principals" (Senate Bill 5).

Tennessee's Consortium on Research, Evaluation, and Development is charged by Tennessee's Race to the Top grant with creating a detailed, focused program of research around key grant initiatives. As part of this work, and in an effort to advise the TEAC's deliberations, the Consortium worked with the Tennessee Department of Education (TDOE) to design and execute the Teacher Evaluation Field Test during the 2010-2011 school year. Preliminary findings from the evaluation field test and alternative models were presented to the TEAC in April 2011. More in-depth findings are now detailed in this report.

## A Review of Teacher Evaluation Literature

A large body of research suggests teachers are the single greatest school factor influencing student achievement (Glazerman et al., 2010; Goldhaber et al., 1999; Nye et al., 2004; Sanders & Rivers, 1996; Wright et al., 1997). Differences in teacher quality have been demonstrated to account for a larger share of differences in student outcomes than any other observable school characteristic (Nye and Hedges, 2004). At the same time, observable characteristics, such as whether teachers hold a master's degree or a specific certification, have been shown to be poor predictors of teacher effectiveness (Hanushek, 2006). Responding to these findings, and unprecedented funding opportunities such as Race to the Top, many state and local educational agencies have begun prioritizing teacher evaluation as a key policy reform area. With a dramatic increase in policy attention being paid to defining, measuring and identifying

effective teaching, the stakes for understanding what makes an evaluation system operational, valid, and accepted by educators and evaluators have never been higher.

Past research has found teacher evaluations are commonly characterized as “short and infrequent... conducted by administrators without extensive training, and influenced by powerful cultural forces” (Weisberg et al., 2009). To improve on the status quo—and with many schools failing to meet local and federal benchmarks for success—a handful of systems have begun using more rigorous, standards-based observation instruments in their teacher evaluation process. Many popular observation instruments, both older and more recent, are based on the work of Charlotte Danielson’s *Framework for Teaching* (1996). These instruments outline four distinct domains for effective teaching, with each domain broken into a series of individual indicators. A handful of studies (Gallagher, 2004; Kimball, et al., 2004; Milanowski, 2004) have investigated the criterion-related validity of observation instruments based on the Danielson framework and found a positive correlation between teacher scores on the observation rubric and student achievement scores.

Researchers have also investigated teacher and evaluator experiences, attitudes, and beliefs during the period in which new evaluation and observation instruments are implemented. By interviewing and surveying teachers and evaluators during and after implementation, researchers have identified several common themes. In particular, researchers have found teachers generally approve of new evaluation models and believe they are understandable and consistent with their own general views of teaching (Milanowski & Heneman, 2001; Kimball, 2002; Heneman and Milanowski, 2003; Heneman et al., 2006). Negative perceptions of individual program elements or the entire model tend to be related to issues of implementation and program design rather than the program model itself (Claudet, 1999; Heneman & Milanowski, 2003; Heneman et al., 2006). Teachers have generally viewed the observation instrument as *potentially* very helpful, and consider the quality and usefulness of feedback received from observations and portfolio reviews to be very important (Kimball, 2002; Heneman et al., 2006). Teachers have also indicated their evaluator’s content knowledge was critical to the usefulness of the feedback they received. Additionally, many negative reactions to instruments have indicated the feedback was focused too heavily on issues of classroom management (Milanowski & Heneman, 2001; Kimball, 2002; Heneman et al., 2006).

**With a dramatic increase in policy attention being paid to defining, measuring and identifying effective teaching, the stakes for understanding what makes an evaluation system operational, valid, and accepted by educators and evaluators have never been higher.**

Despite general approval of models being used for observation, perceptions of the overall program in which an instrument was utilized have been more mixed (Milanowski & Heneman, 2001; Heneman et al., 2006), or negative (Heneman & Milanowski, 2003). Negative reactions to the overall *experience* of using a specific instrument were most commonly related to perceptions of quality and timeliness of feedback and perceptions that feedback was not directly



applicable to improvement (Milanowski and Heneman, 2001; Heneman et al., 2006). Reactions were also linked to perceptions of administrators' ability to effectively implement the evaluation system (Milanowski and Heneman, 2001; Heneman and Milanowski, 2003). The majority of educator complaints clustered around issues associated with increased workloads, potential threats to teacher self-esteem, and a sense the potential benefits of the program were outweighed by the effort required to obtain them (Claudet, 1999; Milanowski and Heneman, 2001).

Amidst the growing public discussion regarding teacher evaluation, the National Institute for Excellence in Teaching (NIET) created the Teacher Advancement Program (TAP), a comprehensive model for evaluation, instruction, and personnel advancement. Beginning with the 2011-2012 school year, the TAP Rubric will be utilized by the majority of Tennessee districts.<sup>2</sup> Because the TAP system was designed to incorporate their observation rubric into a comprehensive personnel system that also involves performance-based compensation and differentiated teacher career paths, previous research has not addressed the effectiveness or validity of the TAP observation rubric independent of the program's other elements.

---

<sup>2</sup>The Hamilton County Department of Education (HCDE) and Memphis City Schools (MCS) as well as schools in districts included in the Association of Independent and Municipal Schools (AIMS) will use state-approved alternative evaluation models.

# METHODOLOGY

---

In an effort to collect information regarding teacher and evaluator perceptions and experiences related to observation and evaluation generally, and educator evaluation in Tennessee specifically, Tennessee's Consortium on Research, Evaluation, and Development assisted TDOE in conducting a field test of evaluation models. Findings from the field test served to inform TEAC discussions, and provide TDOE staff with information about experiences with teacher evaluation across the state. Using a randomized, controlled process, Consortium researchers divided a group of volunteer schools into a control group whose participants would continue to use the state's existing evaluation model, the *Framework for Evaluation and Professional Growth* (Framework). A treatment group whose participants would conduct teacher evaluations using the TAP observation rubric (TAP Rubric) was also identified. It is important to note that because participation was voluntary, schools participating in the field test may not have been representative of all schools statewide.

Concurrently, three other evaluation systems were implemented in Tennessee during the 2010-2011 school year. The Association of Independent and Municipal Schools (AIMS), an organization of 34 public city and special school districts, developed and piloted their own model, *Teacher Instructional Growth for Effectiveness and Results* (TIGER), based on the work of Charlotte Danielson. The Hamilton County Department of Education (HCDE), which oversees 78 schools including those within the city of Chattanooga, moved forward with implementation of Project COACH (COACH) based on the work of Kim Marshall. A third evaluation process was piloted in the Memphis City Schools, but was not examined for this report.

Prior to conducting a survey of teachers, school-level evaluators, and other educators, Consortium researchers conducted informational interviews with district- and state-level administrators responsible for designing and/or implementing each of the evaluation models being researched (Table 1); attended training sessions held for evaluators using the TAP Rubric and TIGER; and were provided with copies of official training materials and evaluation rubrics. Training and implementation for COACH occurred before the Consortium began its exploratory research, but researchers interviewed district-level administrators in HDCE. Researchers were also provided COACH promotional materials, rubrics, and other materials used in training principals and teachers.

## **Teacher evaluation systems piloted in Tennessee during the 2010-2011 school year, and examined for this report:**

1. Framework for Evaluation and Professional Growth (Framework)
2. Teacher Advancement Program observation rubric (TAP Rubric)
3. Teacher Instructional Growth for Effectiveness and Results (TIGER)
4. Project COACH

Though schools using the Framework received no additional training and operated under the same expectations as in prior years, Consortium researchers conducted informational interviews with state-level administrators to discuss the history and development of the Framework, as well as the regulations used for its current implementation. All three groups implementing a new model (TAP Rubric, TIGER, and COACH) continued to use the existing Framework for teachers who were eligible for tenure on the basis of their evaluation during the 2010-2011 school year.

Using the information indicated in the explanations above, Consortium researchers developed and administered a survey to teachers and administrators. Four specific groups were surveyed: the field test control group (Framework), the field test treatment group (TAP Rubric), and those using TIGER and Project COACH.

## Research Questions

After reviewing existing literature, considering interviews with state and district level administrators, and distilling information gathered when attending training sessions, Consortium researchers developed the following research questions:

- 1) To what extent did evaluators feel adequately trained to perform evaluations?
- 2) To what extent did teachers feel that the components and structures of the evaluation system were clearly communicated?
- 3) Are the reported frequencies and duration of observations in the ranges expected by the design of each program (fidelity of implementation)?
- 4) How frequent and prompt was observation feedback received, and in what form was it delivered (written/verbal)?
- 5) On whom has the responsibility of evaluation fallen, and to what extent do these individuals perceive evaluations to add significantly to their workload?
- 6) What do evaluators and teachers perceive to be the positive benefits accruing from the evaluations? What are the challenges?

Once existing research was examined, information regarding the various models was gathered, and research questions were created, survey questions were developed to examine issues of implementation and assess teacher and evaluator perceptions and experiences regarding teacher evaluation. Teachers and evaluators answered parallel questions, with the language altered to reflect whether the respondent conducted observations or was observed.

## Field Test Models

As explained previously, the evaluation field test examined experiences in participant groups using two different models. Teachers and administrators in the control group used the Framework, the model used by the majority of Tennessee teachers at the time. Treatment group participants used the Tap Rubric. Detailed explanations of these field test models are below.

### Framework for Evaluation and Professional Growth (Field Test control group)

The Framework for Evaluation and Professional Growth (Framework), based on the work of Charlotte Danielson, has been used statewide in Tennessee since 2000. Following initial approval in 1997 and a two-year pilot, the model remained unchanged until 2004 when an update added more content-related language and emphasized the use of data in teacher decision-making. The Framework requires evaluators to assign a rating of A, B, C, or *unsatisfactory*, with C being the highest possible rating. Consultants conducted training sessions across the state during the initial implementation phase in 2000, and additional refresher trainings have subsequently occurred throughout the state. Control group schools using the Framework conducted evaluations at the same frequency as previously mandated in state law prior to implementation of the First to the Top Act in July 2011.<sup>3</sup>

The Framework's formal evaluation procedure includes six distinct steps. Educators first complete a self-assessment that asks the teacher to describe his/her perceived strengths and weaknesses, as well as opportunities for improvement. Second, educators have an opportunity to provide an evaluator with a unit or lesson plan. Next, the teacher and evaluator complete at least one full observation cycle that includes a Planning Information Record, an announced, in-class, lesson-length observation, and a Reflection Information Record. A post-conference is then held to complete the Comprehensive Assessment Appraisal Record, which allows the administrator to provide feedback regarding the observation process. Following this observation cycle, the fourth step requires the educator to complete an Educator Information Record that asks questions about assessments, communication with stakeholders, collaboration, and professional development. In the fifth step, the evaluator gathers all of these documents to complete a summative report and determines a performance level (A, B, C, or *unsatisfactory*). Finally, every educator completes a Future Growth Plan that designates areas for improvement; first- and second-year apprentice teachers utilize this document to inform and guide the next year's evaluation.

---

<sup>3</sup>Schools using the Framework are required to adhere to a specific observation schedule and protocol. Teachers in their first or second year are observed a minimum of three times per year. Teachers in their third year of teaching are observed at least twice during that year. Teachers with a professional license (more than three years of experience) are observed at least twice every five years. For each observation, observers are required to provide feedback and accompanying documentation.

## Tap Rubric (field test treatment group)

Field test treatment group schools conducted evaluations using the TAP Rubric and evaluation procedures designed by NIET. To simulate the pace of observations required by state regulations and recommended by NIET, evaluators were asked to conduct educator observations during the eight week duration of the field test, simulating the rate required to evaluate all educators annually. To do so, principals first determined the number of lesson-length observations each evaluator would be required to complete per week at their school during full implementation (assuming 30 available weeks), and multiplied this number by eight to determine the minimum number of lesson-length observations expected for the purposes of the field test.

Over the course of an academic year, the TAP observation model, informed by recommendations from the TEAC, and later approved by the Tennessee State Board of Education, requires a minimum of four observations totaling at least sixty minutes for all teachers, and a minimum of six observations totaling ninety minutes for apprentice teachers. Half of these observations are required to last fifteen minutes and half must be lesson length. For both groups of teachers, only one of the formal lesson-length observations is announced, and this observation requires a pre-conference. For apprentice teachers, two of the three informal fifteen-minute observations must be announced; for professionally licensed/tenured teachers only one of the two fifteen-minute observations is announced. Lesson-length observations always require a formal post-conference; fifteen-minute observations require an informal post-conference.

To train evaluators in the TAP observation process, the state conducted two two-day training sessions to introduce field test participants to the model and explain expectations for evaluating educators during the field test. After the first meeting, conducted in fall 2010, evaluators were encouraged to complete at least one fifteen-minute observation using the TAP Rubric. After the second training session in January 2011, evaluators were required to complete evaluator certification, a process in which participants watch and score video-recorded lessons using TAP's five point scale. To complete the certification, an evaluator's total average score, and 75 percent of indicator scores, must be within one point of the scores determined by the TAP organization. Once certified, evaluators were asked to complete the requisite field test observations using the previously referenced simulation guidelines.

The TAP Rubric consists of four domains—three of which are evaluated during classroom observations. During an observation, an evaluator is required to score an educator on three indicators within the Planning domain, twelve indicators within the Instruction domain, and four

**The evaluation field test examined experiences in participant groups using two different models. Teachers and administrators in the control group used the Framework, the model used by the majority of Tennessee teachers at the time. Treatment group participants used the Tap Rubric.**

indicators within the Learning Environment domain. Evidence gathered outside of classroom observations is used to score the four indicators included in the Professionalism domain.

Prior to announced formal lesson-length observations, the teacher and evaluator complete a pre-conference conversation designed to obtain pertinent background information about the lesson and students involved. Teachers do not need to complete any documentation as part of the pre-conference unless the observation includes the Planning domain, in which case a lesson plan is expected. Prior to unannounced observations, a brief conversation may occur between teacher and observer. Following each lesson-length observation (announced or unannounced), teachers and evaluators participate in a post-conference during which an evaluator presents a teacher with a reinforcement plan designed to identify strengths and label reinforcement areas from the rubric, guide self-reflection, provide evidence from the observation, and recommend continued use of the indicators. Evaluators also provide teachers with a refinement plan during the conference. This plan identifies areas that need to be strengthened, addressing specific teaching standards, providing concrete suggestions related to the refinement area, offering examples of how to improve within the refinement area, and, as with the reinforcement plan, guiding the teacher's self-reflection.

## Other Models

In an effort to present a comprehensive picture of teacher and principal experiences with evaluation across Tennessee during the 2010-2011 school year, Consortium researchers also surveyed individuals using two alternative models: AIMS TIGER and COACH. Both models are described in detail below.

### AIMS TIGER Model

In January 2010, the AIMS consortium, comprising 34 school districts across Tennessee, initiated a discussion regarding the creation of a new educator evaluation model. After engaging external consultants, AIMS directors submitted a request to TDOE to pilot their newly created TIGER model, based on the work of Charlotte Danielson. Forty-seven schools signed on to participate in the 2010-2011 pilot year. The model creates separate observation processes for teachers at different levels of experience and effectiveness. During the pilot year, all non-tenured teachers were evaluated as Stage 1 Teachers and all tenured teachers were evaluated as Stage 2 Teachers. The model also identifies Stage 3 Teachers who are evaluated under the same procedures as a Stage 2 Teacher, but who also act as observation coaches for Stage 1 teachers. For the purposes of the pilot year, no teachers were identified as Stage 3 teach-

ers. Following a year of observation, teachers may remain in their stage, advance to a higher stage, or move to a lower stage.

Under the TIGER model, Stage 1 teachers are observed more often and spend a greater amount of time being evaluated. After a principal has identified a Stage 1 teacher, the principal schedules an announced observation and a formal pre-conference, asking the teacher to complete an Interview Protocol for Pre-conference Form prior to their meeting. The principal then conducts the announced observation and the teacher completes a Teacher Lesson Reflection Form for use in a formal post-conference with the principal. The teacher also gathers appropriate documents and artifacts for use in the post-conference. During the post-conference, the principal and teacher discuss the observation, completed forms, and other data, then the principal completes a Baseline Growth Plan for the teacher. Following the meeting, the teacher completes a separate Baseline Growth Plan and meets with a separate instructional coach, who completes his/her own Baseline Growth Plan for the teacher. Together, the teacher and coach complete an Action Plan, which is referred to during subsequent observations conducted by the coach. Over the course of the year, the coach conducts four ten-minute unannounced observations designed to assess a teacher's progress in addressing educator-specific areas. The coach determines if a teacher is making progress and may engage the principal if needed to work through next steps. During the second semester of the school year, the principal conducts an unannounced observation of the teacher, and both complete the post-conference using the same forms and procedures used during the first observation. At the end of the school year, the principal reviews all available evidence and completes a Summative Evaluation for the School Year Form, giving the Stage 1 teacher an overall proficiency level on a four-point scale.

A Stage 2 teacher is required to complete a Baseline Self-Assessment and Improvement Plan that is submitted to the principal. The principal reviews this document and provides written feedback to the teacher. At this stage a meeting between the teacher and principal is optional. The Stage 2 teacher receives this feedback and completes the third part of the Self-Assessment and Improvement Plan, which includes a Professional Development Plan. The principal and teacher meet to discuss a summary of the performance and improvement plan and discuss an observation schedule that ensures a minimum of sixty minutes of unannounced observation(s). The principal then conducts the observation(s) over the course of the academic year and the teacher submits an Individual Professional Development Log of Activities in the spring. Using this information, the principal reviews the available evidence and determines an overall proficiency level, relaying the appropriate information to the teacher during a summative conference.

## COACH Model (HCDE)

Policymakers in HCDE began discussing the possibility of creating a new educator evaluation system in the spring of 2010 following the First to the Top legislation and a meeting with Kim Marshall, a former principal and author of a *Rethinking Teacher Supervision and Evaluation*. Marshall, a coach and trainer with New Leaders for New Schools, has advocated for a teacher evaluation model that uses short unannounced “mini-observations” instead of lesson-length observations. Intrigued by Marshall’s work, county administrators quickly established an evaluation committee to design and refine a rubric that would build on Marshall’s work while addressing county-specific concerns and the new state requirements.

By utilizing mini-observations, county administrators hoped to design a system that would enable teachers to demonstrate authentic everyday teaching while allowing administrators to gain a better situational awareness of the larger school community. Additionally, by eliminating formal announced observations and the pre- and post-conferences associated with them, the county hoped to maximize instructional time while still allowing for the creation of a comprehensive set of data to be used during a year-end summative evaluation.

The COACH model requires a minimum of ten observations per teacher over the course of an academic year. All observations are unannounced and must last a minimum of five minutes. Additionally, all teachers are observed at least four times outside of the classroom in settings such as grade-level meetings, team planning sessions, or parent meetings. Following each observation, evaluators are required to deliver three to five minutes of verbal feedback to teachers within forty-eight hours. Evaluators are encouraged to provide specific actionable feedback delivered in a neutral setting (not the principal’s office). Following the verbal feedback, evaluators document the feedback in one or two sentences and upload this documentation to an online portal. Teachers are given access to the digital written feedback and are permitted to add their own brief comments as well.

During the observations, evaluators make use of the COACH rubric, which contains six distinct domains, each with ten indicators. All sixty indicators are weighted equally. Although evaluators are not asked to address all indicators in each mini-observation, over the course of the academic year, evaluators are expected to determine a score for all sixty indicators. Each indicator is scored on a 1-4 scale from “Does Not Meet Standards” to “Highly Effective.”

Evaluators complete a summative evaluation at the end of the year using the data collected from the mini-observations, and can choose to discuss the results of the evaluation directly with teachers, but a conference is not necessary. Similarly, evaluators may choose to meet with teachers at the beginning of the year for an initial self-assessment conference. The summative evaluation is used to generate a professional growth plan that identifies specific domains and indicators of focus and allows teachers and evaluators to discuss specific steps and timelines for addressing these areas.



Comparative characteristics of the two field test models and two additional, state-approved models are shown in Table 1.

**TABLE 1**  
**Model Characteristics – TDOE Field Test Models, AIMS TIGER, and HCDE COACH**

	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>HCDE COACH</b>
<b>Origin</b>	Created by TDOE; Statewide implementation 2000; updated 2004	Created by NIET for TAP; Selected by TDOE, field tested, spring 2011	Created by AIMS Consortium, piloted 2010-2011	Created by HCDE, spring 2010
<b>Number and Frequency of Observations</b>	Apprentice Teachers: Annual evaluations with 3 observations; Tenured Teachers: An evaluation every five years with 2 observations	All teachers annually. Apprentice Teachers: six observations, total >90 minutes; Tenured Teachers: four observations, total >60 minutes	All teachers annually Stage 1: A minimum of 6 observations; Stage 2 and 3: Varies, but at least 60 minutes of observation	All teachers annually. At least ten “mini-observations;” each observation should be a minimum of five minutes
<b>Announced or Unannounced</b>	Both	Both	Both	Unannounced
<b>Teacher Feedback</b>	Summary Report & Future Growth Plan	Informal Post-Observation Conference & Summary Rating/ Growth Plan	Written summative Baseline Growth Plan and Action Plan	Three to five minutes of verbal feedback, end-of-year written, summative feedback
<b>Scoring</b>	A, B, C, or “ineffective”; C is highest	1 through 5; 5 is highest	One of four levels for each of the four domains, can switch stages based on results	Six domains each with ten indicators, each on a 1-4 scale
<b>Based on the work of</b>	Charlotte Danielson	Charlotte Danielson	Charlotte Danielson	Kim Marshall

## Survey Administration

The 2011 Tennessee Educator Evaluation Survey (see Appendix A) was designed to gather information regarding implementation of the various teacher evaluation models, and educator perceptions of evaluation generally. It was administered to teachers and principals, or others designated as observers, through an online survey manager. Participants were contacted using e-mail addresses solicited directly from principals and district-level administrators by Consortium researchers and TDOE personnel. All schools within either the AIMS consortium or the HCDE were included in the survey if they had implemented the TIGER or COACH models during the 2010-2011 school year. This yielded forty-seven AIMS schools implementing TIGER and seventy-eight HCDE schools implementing COACH.

Field test participants were chosen by soliciting interest from district-level directors. Directors nominated schools, and with the principal’s agreement, one hundred thirty-four schools agreed to participate. Of those, fifty schools were randomly assigned to the treatment group. The remaining schools were assigned to the control group and instructed to conduct teacher observations using the existing Framework and standard evaluation procedures. Four field test treatment group schools declined to participate after being assigned, and one treatment group school failed to send appropriate contact information for conducting a survey. Nine field test control group schools declined to participate, two were excluded because they implemented the TIGER model, and twelve failed to send appropriate contact information for the survey. The resulting sample included forty-five treatment group schools and sixty control group schools.

Teachers and evaluators in participating schools were contacted via email with an active link to access the online survey, which was estimated to require ten minutes or less to complete. Following the initial email, reminder emails, including the same survey link, were sent on a regular basis. Consortium researchers also requested the help of principals and district-level administrators to promote participation among their schools. Summary statistics are shown in Table 2. It is important to note response rates are similar across all four groups.

**TABLE 2**  
**Survey Participation and Response Rates by Model**

	<b>Field Test Control (Framework)</b>	<b>Field Test Treat- ment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>HCDE COACH</b>
Districts	31	31	21	1
Schools	60	45	47	78
Invitations	2,373	1,822	1,619	2,596
Responses	1,272	1,048	928	1,412
Response Rate	54%	58%	57%	54%

## Analytical Methods

Following the completion of the survey period, Consortium researchers collected the data and began analysis. The survey was designed to “branch” respondents to one of two sets of questions on the basis of their answer to the first question—“What is the title of your current position at this school?” Those responding to the first question “Teacher” or “Other (Specify)” were directed towards questions designed for teachers or other educators being evaluated. Respondents who selected “other” specified a wide variety of positions, some of which were subject to evaluation (e.g., school counselor, librarian) while others were not (e.g., clerk, Title 1 coordinator). Due to the wide variation of the ways in which these individuals interact with

the evaluation system, data from respondents who selected “other” are excluded from these analyses.

Those responding “Principal,” “Assistant principal or vice principal,” or “Instructional coach, mentor, etc.” were directed towards questions designed for evaluators. Across all four models, over 96 percent of principals, assistant principals, or vice principals (266 of 277) indicated they were responsible for conducting formal evaluations. Only 15 percent of instructional coaches and mentors (18 of 119) indicated they were responsible for conducting teacher evaluations. Because of this sizeable difference, Consortium researchers limited the “evaluator” category to respondents indicating they acted as a principal, assistant principal, or vice principal. Due to their relatively small numbers within each group and the complicated and often dynamic role they played in the evaluation processes, responses from instructional coaches and mentors are not included in these results.

To summarize, teachers—the evaluatees—compose the first analytical group, and principals and assistant principals who also served as evaluators compose the second group. Instructional coaches, librarians, central-office evaluators, and other support staff represent a much smaller, varied proportion of respondents and are excluded from the analyses.

All respondents were asked multiple-choice questions about years and types of teacher/administrative experience, attitudes about First to the Top reforms, and benefits and challenges associated with their evaluation system. Questions related to the nature, duration, and timeline of observations were only presented to those respondents who indicated they either had been formerly observed or conducted formal observations during the 2010-2011 school year. Because of the branching rules, the number of respondents considered for each question varies significantly.

**To summarize, teachers—the evaluatees—compose the first analytical group, and principals and assistant principals who also served as evaluators compose the second group. Instructional coaches, librarians, central-office evaluators, and other support staff represent a much smaller, varied proportion of respondents and are excluded from the analyses.**

Consortium researchers also coded free response questions pertaining to benefits and challenges experienced with the various evaluation systems. Researchers developed a taxonomy for categorization based on existing literature, elements of the survey, and, in some cases, knowledge of the nature of the specific model being used. This coding was designed to illustrate relative frequencies, broad themes, and trends among the responses to the open-ended questions; it was not designed as an exhaustive qualitative analysis. Where applicable, results from these open-ended response questions are reported within the results section.

## Respondent Characteristics

With the two key analytical groups defined—teachers and evaluators—the analysis shifts to investigating proportionality between groups by model on observable characteristics. Respondents’ self-reported average years of experience, summarized in Table 3, were mostly similar across the four models. Administrators in the Tap Rubric group, on average, had fewer years of experience in administration than administrators in the Framework group. Teachers utilizing the COACH model had, on average, fewer years of teaching experience in their current district and school compared to teachers in the Framework group.

**TABLE 3**  
**Average Years of Experience**

Total years experience...		Field Test Control (Framework)	Field Test Treatment (TAP Rubric)	AIMS TIGER	HCDE COACH
Teachers	...in teaching	14.1	13.2	14.2	13.6
	...in present district	11.1	10.2*	10.8	10.2*
	...in present school	7.9	7.5	9.2**	6.7**
Evaluator	...as an administrator or coach	9.3	6.8*	8.1	9.7
	...as an administrator or coach in present district	8.6	5.9**	7.4	8.6
	...as an administrator or coach in present school	5.6	4.6	5.9	4.4

*Asterisks represent the result from a test of statistical significance between the TAP Rubric, AIMS TIGER, or COACH group compared with the Framework group. No asterisk indicates the TAP, AIMS, or COACH group is statistically similar to the Framework group; \*  $p < .05$ , \*\*  $p < .01$*

The proportion of tier-representation among respondents differed significantly, with the TIGER sample exhibiting a disproportionate percentage of elementary school teachers compared to Framework schools, and schools using the TAP Rubric exhibiting a disproportionate percentage of high school teachers compared to Framework schools—despite Tap Rubric and Framework groups being randomly assigned from the same pool. These tier differences are likely the cause of the disparities in subject taught, with TIGER respondents showing a significantly higher representation of teachers listing English Language Arts/Reading or Mathematics as a subject taught (See Table 4).

**TABLE 4**  
**Teachers by Schooling Level and Subject (% of Respondents)**

	<b>Field Test Control (Framework) (N= 1061)</b>	<b>Field Test Treatment (TAP Rubric) (N=854)</b>	<b>AIMS TIGER (N=752)</b>	<b>HCDE COACH (N=1211)</b>	
Level	Elementary	49%	37%	65%	47%
	Middle	22%	19%	17%	22%
	High	25%	39%	12%	26%
	Other	5%	6%	6%	5%
Subject	ELA/Reading	51%	45%	63%	46%
	Mathematics	45%	40%	57%	43%
	Performing Arts	7%	6%	5%	6%
	Physical Education	4%	4%	5%	4%
	Special Education	13%	11%	9%	13%

The statistical comparison between the Framework group and other three groups reveals significant differences in the years of experience of teachers and evaluators, tiered representation, and subjects taught. These differences, most notably the difference in the proportional representation of culturally distinct tiers, may have introduced bias into the survey results, and will be further discussed in the analytical section.

# STUDY FINDINGS<sup>4</sup>

## 1) To what extent did evaluators feel adequately trained to perform evaluations?

Evaluators were asked to rate the quality of the training they received as *Excellent*, *Good*, *Satisfactory*, or *Unsatisfactory*. An aggregation of these ratings is shown in Table 5. A fairly high percentage of evaluators from the Framework and Tap Rubric groups reported they had been adequately trained in their respective evaluation systems. Approximately four out of five perceived the training as *Excellent* or *Good*, one in five claimed their training was *Sufficient*, and less than two percent perceived their preparation as *Insufficient*.

Relative to the other three groups, TIGER evaluators reported they were the least prepared, with one in five claiming their training was *Insufficient*. Approximately two-thirds of evaluators using COACH reported their training to be *Excellent* or *Good*.

**TABLE 5**  
**Evaluator Rating on Quality of Training**

	Field Test Control (Framework) (N= 1061)	Field Test Treatment (TAP Rubric) (N=854)	AIMS TIGER (N=752)	HCDE COACH (N=1211)
<b>Quality of Training Was Excellent or Good</b>	77%	81%	58%	64%
<b>Quality of Training Was Satisfactory</b>	20%	17%	18%	31%
<b>Quality of Training Was Insufficient</b>	2%	1%	22%	4%

*Less than 3% of each evaluation group responded that they 'Did not receive training'*

<sup>4</sup>Detailed descriptive statistics from the 2011 Tennessee Educator Evaluation Survey can be found in Appendices B1 and B2.

## 2) To what extent did teachers feel that the components and structures of the evaluation system were clearly communicated?

Teachers were asked about their perceptions of communications concerning Tennessee’s plans for teacher evaluation; results are reported in Table 6.

**TABLE 6**  
**Teachers’ Level of Agreement with the Following Statement:**

Tennessee’s plans for teacher evaluation have been clearly communicated to me.

	<b>Field Test Control (Framework) (N= 1022)</b>	<b>Field Test Treatment (TAP Rubric) (N=825)</b>	<b>AIMS TIGER (N=738)</b>	<b>HCDE COACH (N=1193)</b>
<b>Strongly Agree or Agree</b>	50%	73%	74%	81%

*Alternative response options were ‘Disagree’ and ‘Strongly Disagree’*

In retrospect, the research team should have utilized a more precisely worded question, as it is unclear if teachers perceived the question’s subject to be about the evaluation model or the entire teacher evaluation system. The words “evaluation system” were utilized, but the question was also preceded and followed by inquiries about the specific evaluation model. Without this clear distinction between the evaluation model and the broader teacher evaluation system, results should be interpreted cautiously. Half of teachers utilizing the Framework indicated Tennessee’s plans for teacher evaluation had not been clearly communicated; perhaps a predictable finding considering modifications to the existing model had also occurred. Over eighty percent of teachers in HCDE schools indicated a clear understanding, and approximately three out of four teachers in the TAP Rubric and TIGER groups indicated they received clear communication. Again, it is impossible to know if this perception of clear communication applies to the evaluation model only, or the broader teacher evaluation system. What is clear is that there is room for better communication as the evaluation scales up statewide.

### 3) Are the reported frequencies and duration of observations in the ranges expected by the design of each program (fidelity of implementation)?

The survey asked respondents several questions about the nature and frequency of their classroom observations. Comparing these responses to the pre-survey research conducted on each model, Consortium researchers were able to gauge fidelity of implementation across models during the 2010-2011 school year. As shown in Table 7 below, the number of observations teachers reported at the time they completed the survey resonates with expectations about the number of observations required under each model. Operating under established expectations for evaluators, a larger percentage of teachers using the Framework reported not being evaluated. Because the TAP Rubric group implemented its evaluation model significantly later than AIMS or HCDE schools using COACH, their percentage of teachers reporting “none” at the time of the survey was significantly higher. The large percentage of COACH teachers reporting six or more observations at the time of the survey indicates many schools in HCDE were on track to complete the required number of mini-observations by the end of the school year. Similarly, the high percentage of teachers in the TAP Rubric group and the TIGER group reporting they had been observed at least once confirms that a majority of schools appeared to be implementing the model according to district or state-level expectations.

**TABLE 7**  
**Number of Observations Reported by Teachers (% of Respondents)**

Observations	Field Test Control (Framework) (N= 1061)	Field Test Treatment (TAP Rubric) (N=854)	AIMS TIGER (N=752)	HCDE COACH (N=1211)
<b>1 to 3</b>	53%	67%	71%	11%
<b>4 to 5</b>	6%	6%	11%	29%
<b>6 +</b>	3%	1%	3%	57%
<b>None</b>	37%	26%	15%	2%



When asked about the length of their most recent classroom observation, teacher responses also matched researchers' expectations (Table 8). Most observations that occurred within the COACH model were less than twenty minutes, while observations occurring within the TAP Rubric and TIGER models ranged between ten and forty minutes. The large percentage of observations lasting over forty minutes in the TAP Rubric group may be related to the over-representation of high schools, while the larger percentage of shorter observations among TIGER respondents is likely related to the over-representation of elementary schools. Responses from open-ended questions suggest lesson-length observations are often longer in high schools compared to elementary schools.

**TABLE 8**  
**Length of Most Recent Observation, Reported by Teachers (% of Respondents)**

<b>Length of Observations (minutes)</b>	<b>Field Test Control (Framework) (N= 1061)</b>	<b>Field Test Treatment (TAP Rubric) (N=854)</b>	<b>AIMS TIGER (N=752)</b>	<b>HCDE COACH (N=1211)</b>
<b>&lt; 10</b>	7%	3%	5%	22%
<b>10 to 15</b>	14%	6%	20%	48%
<b>15 to 20</b>	13%	11%	26%	18%
<b>20 to 30</b>	15%	18%	29%	7%
<b>30 to 40</b>	21%	24%	13%	2%
<b>40 +</b>	27%	36%	6%	2%
<b>Other</b>	3%	1%	1%	1%

One of the critiques of No Child Left Behind is that the accountability system, which determines success or failure based on math and reading/language tests only, skews school resources away from arts, foreign languages, science, and social studies. Researchers tested whether evaluators treated teachers of NCLB-tested subjects differently than teachers of non-NCLB subjects. The incentive structure built into NCLB may have resulted in evaluators spending more time in the classrooms of NCLB-tested subjects, or providing better or more feedback to these teachers. Also tested was a correlation between the average time of observation and teachers of any of the four TCAP subjects: math, reading/language, science and social studies. Tests of statistical significance found no evidence that teacher-reported evaluation periods were longer for either NCLB accountable subjects or TCAP tested subjects. Similarly, no statistically significant difference was found when comparing the type and swiftness of feedback received. Based on teacher reporting, evaluators did not treat evaluations of NCLB-tested or TCAP subjects differently.

#### 4) How frequent and prompt was observation feedback received, and in what form was it delivered (written/verbal)?

When asked about how much time passed before teachers received written or verbal feedback (Table 9), responses also confirmed expectations. Overwhelming majorities in all models reported receiving feedback in less than three days, although evaluators reported shorter time spans between observations and feedback than did teachers.

**TABLE 9**  
**Number of Days Between Observation and Verbal or Written Feedback**  
 Reported by Teachers and Evaluators (% of Respondents)

	Number of Days	Field Test Control (Framework)		Field Test Treatment (TAP Rubric)		AIMS TIGER		HCDE COACH	
		Teachers (N=1061)	Evaluators (N=89)	Teachers (N=854)	Evaluators (N=88)	Teachers (N=752)	Evaluators (N=71)	Teachers (N=1211)	Evaluators (N=116)
Verbal	<3	67%	77%	61%	82%	68%	65%	90%	100%
	3-10	24%	20%	33%	16%	27%	33%	9%	0%
	11-20	4%	1%	3%	0%	4%	2%	1%	0%
	>20	4%	1%	3%	2%	2%	0%	0%	0%
Written	<3	64%	66%	57%	84%	59%	66%	82%	79%
	3-10	27%	31%	37%	14%	34%	33%	16%	21%
	11-20	6%	2%	4%	0%	5%	2%	1%	0%
	>20	3%	2%	2%	2%	2%	0%	1%	0%

Taken together, the data suggest all four models produced regular verbal and written feedback. While results were similar across the TAP Rubric group, the Framework group, and the TIGER group, COACH teachers and evaluators reported receiving more timely verbal and written feedback.

## 5) On whom has the responsibility of evaluation fallen, and to what extent do these individuals perceive evaluations to add significantly to their workload?

Table 10 reports findings from an inquiry into how the evaluation model increased or decreased teacher and evaluator workloads relative to observations and evaluation. Evaluators for all four of the models reported an increased workload, with the most significant increase among those using the TAP Rubric and TIGER models. Teachers had more varied responses, with teachers using the COACH model indicating an average *decrease* in their workload related to observation and evaluation. This finding is likely due to the model’s procedural design, which requires no pre-observation work from teachers. Teachers utilizing the TAP Rubric report no significant change, while teachers using the TIGER model report an increased workload.

These findings are supported by the categorization of open-ended response items, in which the most common challenge for evaluators across all models was ‘Time Demands.’ Teacher responses were more nuanced, with teachers utilizing the TAP Rubric reporting the third most common benefit was ‘Less Paperwork’; however, they joined their colleagues utilizing the Framework and TIGER models with the most common challenge reported as ‘Time Demands.’ Teachers in the TAP Rubric group and COACH groups also cited the time demands imposed on their *administrators* as a challenge associated with each model. This finding helps confirm the earlier suggestion that the TAP Rubric and COACH models shifted time demands and responsibilities from teachers to administrators.

**TABLE 10**  
**How Has the Model Increased or Decreased Your Workload in Terms of Completing Observations and Evaluation-Related Tasks?**

	Field Test Control (Framework)	Field Test Treatment (TAP Rubric)	AIMS TIGER	HCDE COACH
<b>Teachers</b>	52% (991)	48% (803)	69% (734)	35% (1178)
<b>Evaluators</b>	66% (82)	86% (77)	84% (76)	73% (44)

*Figures represent the percent of respondents who selected Moderately increased or Significantly increased (alternative responses were No change, Moderately decreased, and Significantly decreased). Number of responses given in parentheses.*

An additional analysis investigated whether administrator responses were correlated with the number of years of administrator experience, as switching to a new evaluation system may be more or less challenging for a newer administrator. No evidence for such a difference was found.

Finally, Table 11 shows the average number of teacher-reported observers present at their most recent observation. Across all four models, teachers report that observations were generally completed by a single observer, most often the principal; observations by multiple observers appear to be rare.

**TABLE 11**  
**Teacher-Reported Average Number of Observers, Most Recent Observation**

	<b>Field Test Control (Framework)</b> (N= 1022)	<b>Field Test Treatment (TAP Rubric)</b> (N=825)	<b>AIMS TIGER</b> (N=738)	<b>HCDE COACH</b> (N=1193)
<b>Avg # of Evaluators</b>	1.1 (647)	1.2 (618)	1.1 (629)	1.1 (1168)

## 6) What do evaluators and teachers perceive to be the positive benefits accruing from the evaluations? What are the challenges?

Given the brisk pace of change surrounding educator evaluation and education policy in Tennessee generally, the survey asked a series of questions designed to gauge perceptions and opinions related to evaluation and accountability generally. When asked about their perceptions of the purpose of their most recent observation (Table 12), both teachers and evaluators overwhelmingly reported their observation was designed to give feedback to help improve teaching and make a judgment about performance. Across all four groups, a higher percentage of evaluators reported their observations were designed to accomplish both objectives. And in most groups, more teachers and evaluators indicated the observation was designed to provide feedback than it was to make a judgment about performance.

**TABLE 12**  
**Teacher and Evaluator Perceptions of Observation Purpose**

Was your most recent observation designed to...	Field Test Control (Framework)		Field Test Treatment (TAP Rubric)		AIMS TIGER		HCDE COACH	
	Teachers (N=1061)	Evaluators (N=89)	Teachers (N=854)	Evaluators (N=88)	Teachers (N=752)	Evaluators (N=71)	Teachers (N=1211)	Evaluators (N=116)
Give feedback to help improve teaching.	85%	93%	89%	99%	84%	92%	84%	98%
Make a judgment about performance.	74%	92%	75%	95%	70%	96%	66%	100%

*Percentage of Respondents Indicating “Yes”*

Teachers were less in agreement concerning whether the evaluation system contributed to greater collegiality and professionalism, with between 51 percent and 61 percent responding ‘Agree’ or ‘Strongly Agree’ across all four models. As reported in Table 13, evaluators reported a more optimistic view of the impact of evaluations on teacher professionalism. The COACH model shows the largest gap between teacher and principal perceptions, with the TAP Rubric and TIGER models showing a slightly lower teacher/principal perception gap.

**TABLE 13**  
**Teachers’ and Evaluators’ Level of Agreement with the Following Statement**

Teacher evaluation systems used at this school contribute to greater collegiality and professionalism among teachers

	Field Test Control (Framework) (N= 1022)	Field Test Treatment (TAP Rubric) (N=825)	AIMS TIGER (N=738)	HCDE COACH (N=1193)
<b>Teachers</b>	51% (995)	58% (802)	61% (724)	55% (1167)
<b>Evaluators</b>	50% (80)	74% (77)	79% (77)	87% (45)

*Alternative response options were ‘Disagree’ and ‘Strongly Disagree’*

Two additional survey questions probed teachers' perceptions of the challenges of the evaluation models, specifically pertaining to its fairness and comprehensiveness. Few teachers believed the teacher evaluation system was unfair given the characteristics of their students, and slightly less than half believed it omitted important aspects of teaching that should be considered.

**TABLE 14**  
**Teachers' Level of Agreement with the Listed Statement**

	<b>Field Test Control Framework</b>	<b>Field Test Treatment TAP Rubric</b>	<b>AIMS TIGER</b>	<b>HCDE COACH</b>
Teacher evaluation systems at my school are unfair because of the characteristics of the student population	30% (998)	25% (792)	26% (729)	20% (1173)
Teacher evaluation systems used at this school omit important aspects of teaching that should be considered	45% (992)	41% (796)	36% (733)	37% (1163)

*Percentage shown is the percent of respondents that responded 'Strongly agree' or 'Agree' to the following beliefs. Other answer options were 'Strongly disagree' and 'Disagree'.*

One final challenge cited consistently by both teachers and evaluators in all four groups was the applicability of the evaluation systems to educators in a variety of placements. An administrator in the TAP Rubric group expressed concern that “the new pilot that we are currently using does not work for all teachers. We have one model for one kind of teacher. It does not work for specials (e.g., music, librarian, guidance), special education, and speech and language teachers.” Teachers and administrators in all groups echoed these concerns and expressed a desire for modified versions of the evaluation systems that better addressed teachers in non-traditional placements.

## Group differences and bias

Having investigated the six research questions, this paper now revisits the differences in observable characteristics between groups; specifically, we look retrospectively at what results may have been related to these differences. Some of these observable differences are sizeable, although the lack of remarkable variation within survey results allows for more straightforward scrutiny.

First, as evidenced in Table 3, teachers utilizing the TAP Rubric have less experience within their present district by almost one year compared to teachers using the Framework, while teachers utilizing TIGER have been in their present school, on average, 1.7 years longer than Framework teachers. Teachers utilizing the COACH model have been in both their present school and district approximately one year less than Framework teachers. The most important teacher experience measure—the total number of years of teaching experience—is statistically similar across all four groups. Based on this final point, and the relative small sizes of the differences, the authors believe these differences can be dismissed.

One potentially more damaging threat to the validity of findings is the statistically significant, two-and-a-half year difference between the average total years of administrative experience when comparing evaluators using the Framework to evaluators using the TAP Rubric. This difference may bias survey results as administrators with fewer years of experience may find the new evaluation creates a greater workload than administrators with more experience. However, researchers explicitly tested for this, and found no evidence supporting this contention.

The final significant difference, differential representation of tiers across groups, was also explored by running differences by tier on questions on which educational theory would suggest the varying culture of tiers would disproportionately impact results. For example, one might suspect that the effect of observations on teacher collegiality and professionalism differed by tier. Little evidence was found supporting this and other tier-associated biases.

## Summary

To summarize, the 2011 Tennessee Educator Evaluation Survey found participants across all four groups indicated the models were faithfully implemented overall: the number of teacher observations were consistent with expectations, as were the length of observations and frequency of feedback. The impact of the evaluations on workloads was varied, with teachers using the TAP Rubric suggesting their workload decreased slightly, while evaluators using the same model reported an increased workload. Both teachers and evaluators using the three new models—TAP Rubric, TIGER, and COACH—reported an increased *evaluator* workload relative to the evaluator workload under the Framework.

When asked about perceptions of the purpose of their most recent observation (Table 12) teachers and evaluators across all four groups overwhelmingly report the observations were designed both to make judgments about performance, and also to give feedback to help improve teaching. Again across all four groups, a higher percentage of evaluators reported their observations were designed to accomplish both objectives. Finally, and again across all four groups, a higher percentage of both teachers and evaluators indicated the observations were designed to provide feedback more than they were to make judgments about performance.

# RECOMMENDATIONS AND CONCLUSION

---

Findings from the 2011 Tennessee Educator Evaluation Survey illuminated several key areas of potential interest to policy makers and practitioners. The knowledge gleaned from the survey and the research conducted in conjunction with it guides the following policy recommendations and avenues for future research.

First, TDOE should continue to monitor the implementation of teacher evaluation programs through an annual survey. In addition to tracking longitudinal changes in the implementation of the evaluation system, this survey will also assess how perceptions of observation instruments correlate with perceptions about other aspects of state education reforms including teacher evaluation generally, changes to laws regarding teacher tenure, the use of Common Core Standards, and school turnaround procedures.

Second, considering the increased workload on evaluators, TDOE should explore technological efficiencies that minimize the burden of both observations and evaluations such as a tablet or laptop evaluation program that automatically communicate with the proposed statewide, centralized system. An efficient and quality-controlled collection system that feeds an integrated database will allow the department to monitor the fidelity of evaluation implementation, especially the type, swiftness, and quality of evaluator feedback. Such a structure will also eventually allow important investigation into the intermediate and long-term impact of evaluations on student achievement.

The Consortium also recommends the process used to train field test evaluators on utilizing the TAP Rubric should be maintained during statewide scale-up. Evaluators should be periodically assessed on the consistency of their rating standards, with additional training and re-norming applied when appropriate.

A fourth recommendation addresses communication efforts. In order to ensure that the reform momentum from the originating year of First to the Top continues, TDOE should make every effort to provide consistent and clear communications with Tennessee teachers and other stakeholders, particularly as it relates to educator evaluation.

It is further recommended that TDOE examine the evaluation model to ensure the TAP Rubric sufficiently allows for variations within teaching responsibilities. Given the outcome of this examination, modified rubrics might be considered for use with non-classroom positions such as librarians and instructional coaches.



Finally, the department should utilize its position to consider the diffusion of evaluator responsibilities. While principals will certainly continue to shoulder the primary responsibility for observations and ultimately for evaluations, greater engagement on the part of assistant principals, as well as central office leadership and other school leaders—such as lead teachers or instructional coaches—may be warranted.

To conclude, the detailed survey results and analysis included herein address issues of implementation, resources and capacity, perceptions, and opinions about benefits and challenges. It is the Consortium's hope that the 2011 Tennessee Educator Evaluation Survey is the first of many efforts to review educator evaluation instruments in Tennessee, and that future research will be used to help refine evaluation systems as part of statewide reforms designed to improve teacher effectiveness, educational leadership, and student achievement.

# REFERENCES

---

- Claudet, J.G. (1999). An interpretive Analysis of Educator Change Processes in Response to Program Innovation: Implications for Personnel Evaluation. *Journal of Personnel Evaluation in Education*, 13(1), 47-69.
- Education Reforms: Exploring Teacher Quality Initiatives: Hearings before the House Education and the Workforce Committee*, 112th Cong. (2011). (testimony of Kevin Huffman).
- Gallagher, H. A. (2004). Vaughn Elementary's innovative teacher evaluation system: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education*, 79(4), 79-107.
- Gabriel, T. (2010). A celebratory road trip by education secretary, *New York Times*, September 22, 2010. p. A24.
- Glazerman, S., Loeb, S., Goldhaber, D., Staiger, D., Raudenbush, S., & Whitehurst, G. (2010). *Evaluating teachers: The important role of value-added*. Brown Center on Education Policy. Washington, DC: Brookings Institution.
- Goldhaber, D., & Anthony, E. (2004). Can teacher quality be effectively assessed? (Working Paper). Seattle, WA: Center on Reinventing Public Education. Retrieved from [http://www.crpe.org/cs/crpe/download/csr\\_files/wp\\_crpe6\\_nbptsoutcomes\\_apr04.pdf](http://www.crpe.org/cs/crpe/download/csr_files/wp_crpe6_nbptsoutcomes_apr04.pdf).
- Hanushek, E. A. and Rivkin, S. G. (2006). 'Teacher quality', in Handbook of the Economics of Education, eds E. A. Hanushek and F. Welch, Elsevier, Amsterdam.
- Heneman, H. G., III, & Milanowski, A. T. (1999). Teachers' attitudes about teacher bonuses under school-based performance award programs. *Journal of Personnel Evaluation in Education*, 12(4), 327-341.
- Heneman, H. G., III, & Milanowski, A. T. (2003). Continuing assessment of teacher reactions to a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education*, 17(2), 173-195.
- Heneman, H. G., Milanowski, A., Kimball, S. M., & Odden, A. (2006). *Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay* (CPRE Policy Briefs No. RB-45). Philadelphia: Consortium for Policy Research in Education.
- Kimball, S. M. (2002). Analysis of feedback, enabling conditions, and fairness perceptions of teachers in three school districts with new standards-based teacher evaluation systems. *Journal of Personnel Evaluation in Education*, 16(4), 241-268.

Kimball, S. M., White, B., Milanowski, A.T., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education*, 79(4), 54-78.

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement Evidence from Cincinnati. *Peabody Journal of Education*, 79(4), 33-53.

Milanowski, A. T., & Heneman, H. G., III. (2001). Assessment of teacher reactions to a standards-based teacher evaluation system: A pilot study. *Journal of Personnel Evaluation in Education*, 15(3), 193-212.

Nye, B., Konstantopoulos, S. & Hedges, L. How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237-257.

Sanders, W.L. & Rivers, J.C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research Center.

Solmon, L.C., White, J.T., Cohen, D., and Woo, D. (2007). *The Effectiveness of the Teacher Advancement Program*. Santa Monica, CA: Milken Family Foundation.

Springer, M.G., Ballou, D., and Peng, A. (2008). *Impact of the Teacher Advancement Program on Student Test Score Gains: Findings from an Independent Appraisal*. National Center on Performance Incentives Working Paper #2008-19. Nashville, TN: Vanderbilt University.

Weingarten, R. A new path forward for public education. *The Huffington Post*, January 15, 2010.

Weisberg, Daniel, Susan Sexton, Jennifer Mulhern, David Keeling. (2009). *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness*. Retrieved from widgeteffect.org.

Wright, S.P., Horn, S.P., & Sanders, W.L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, 57-67.

# APPENDIX A

## **Tennessee Educator Evaluation Survey** **Winter 2011**

**DEFINITION:** For the purpose of this survey, formal classroom observation involves an observer who observes and takes notes with the intention of some type of verbal or written feedback.

### **1. What is the title of your current position at this school?**

*Mark only one.*

- a. Principal
- b. Assistant principal or vice principal
- c. Instructional coach, mentor, etc.
- d. Teacher
- e. Other, specify: \_\_\_\_\_

SKIP TO QUESTION 24

SKIP TO QUESTION 24

SKIP TO QUESTION 23

### **TEACHER SURVEY FORM (Questions 2-22)**

### **2. Including this school year (2010-2011), how long have you worked as a teacher?**

- a. Total time employed as a teacher
  - b. Time employed in your present district
  - c. Time employed in your present school
- Years

### **3. Please indicate the grade(s) you currently teach at this school.**

*Mark all that apply.*

a. 1st grade	f. 6th grade	k. 11th grade
b. 2nd grade	g. 7th grade	l. 12th grade
c. 3rd grade	h. 8th grade	m. Kindergarten
d. 4th grade	i. 9th grade	n. Pre-kindergarten
e. 5th grade	j. 10th grade	o. Ungraded

**4. Which subjects do you currently teach?**

*Mark all that apply.*

- a. English / language arts / reading
- b. Mathematics
- c. Science
- d. Social studies or history
- e. Foreign language
- f. Special instruction for English Language Learners (ELL) or Limited English Proficient (LEP) students
- g. English as a Second Language (ESL)
- h. Visual or performing arts
- i. Special education
- j. Physical education
- k. Other, specify: \_\_\_\_\_

**5. How many TOTAL times during this school year (2010-2011) do you expect to be formally observed teaching? (Include observations that have already occurred)**

*Mark one answer only.*

- a. 1 to 3 times during this school year
- b. 4 to 5 times during this school year
- c. 6 to 7 times during this school year
- d. 8 to 9 times during this school year
- e. 10 or more times during this school year
- f. I do not expect to be observed during this school year
- g. I do not know

**6. During this school (2010-2011), how many times has someone formally observed you teaching **THUS FAR**?**

*Mark one answer only.*

- a. 1 to 3 times this school year
- b. 4 to 5 times this school year
- c. 6 to 7 times this school year
- d. 8 to 9 times this school year
- e. 10 or more times this school year
- f. I have not been observed this school year

**SKIP TO 18**

**7. How many school days have passed since your last observation?**

*Do not include weekends or holidays. Only count school days.*

|\_\_| School days

**8. How many minutes did the individual(s) conducting the observation stay in your classroom DURING YOUR MOST RECENT OBSERVATION?**

*Mark one answer only.*

- a. Less than 10 minutes
- b. 10 to 15 minutes
- c. 15 to 20 minutes
- d. 20 to 30 minutes
- e. 30 to 40 minutes
- f. More than 40 minutes

**9. Based on your experience during this school year (2010-2011), when you are observed, ON AVERAGE how much time does the observer spend in your classroom?**

*Mark one answer only.*

- a. Less than 10 minutes
- b. 10 to 15 minutes
- c. 15 to 20 minutes
- d. 20 to 30 minutes
- e. 30 to 40 minutes
- f. More than 40 minutes

**10. How would you describe the professional status of your classroom observer(s)?**

*Mark all that apply.*

- a. Principal(s)
- b. Assistant or vice principal(s)
- c. Department head(s)
- d. Math or literacy coach(es)
- e. Other senior teacher(s) from the school, such as a mentor, master, or lead teacher
- f. Observer(s) not working at your school
- g. Other, specify: \_\_\_\_\_

**11. How many individuals observed you in your classroom DURING YOUR MOST RECENT OBSERVATION?**

*Mark one answer only.*

- a. 1
- b. 2
- c. 3
- d. 4
- e. 5
- f. More than 5

**12. Thinking of your most recent observation, was it designed to give you feedback to help you improve your teaching?**

- a. Yes
- b. No

**13. Thinking of your most recent observation, was it designed to make a judgment about your performance?**

- a. Yes
- b. No

**14. Thinking of your most recent observation, did you receive or do you expect to receive VERBAL feedback regarding your observation?**

*Mark one answer only.*

- a. I have received verbal feedback
- b. I have not received verbal feedback, but I do expect to receive it
- c. I have not received verbal feedback, and I do not expect to receive it

SKIP TO 16  
SKIP TO 16

**15. How much time passed before you received VERBAL feedback regarding your observation?**

*Do not include weekends or holidays. Only count school days. Mark one answer only.*

- a. Less than 3 school days
- b. More than 3 school days, but less than 10 school days
- c. More than 10 school days, but less than 20 school days
- d. More than 20 school days

**16. Thinking of your most recent observation, did you receive or do you expect to receive WRITTEN feedback regarding your observation?**

*Mark one answer only.*

- a. I have received written feedback
- b. I have not received written feedback, but I do expect to receive one
- c. I have not received written feedback, and I do not expect to receive one

SKIP TO 18  
SKIP TO 18

**17. How much time passed before you received WRITTEN feedback regarding your observation?**

*Do not include weekends or holidays. Only count school days. Mark one answer only.*

- a. Less than 3 school days
- b. More than 3 school days, but less than 10 school days
- c. More than 10 school days, but less than 20 school days
- d. More than 20 school days

**18. Please indicate the model currently being used for observations in your school.**

*Mark only one answer.*

- a. Framework for Evaluation and Professional Growth
- b. State of Tennessee Field Test (TAP rubric)
- c. AIMS TIGER Process
- d. COACH Model
- e. I do not know
- f. Other, specify: \_\_\_\_\_

**19. Please indicate how the model being used for observation and evaluation has increased or decreased your WORKLOAD in the following areas during the current school year (2010-2011).**

*Select one response for each row.*

	<b>Significantly decreased</b>	<b>Moderately decreased</b>	<b>No Change</b>	<b>Moderately increased</b>	<b>Significantly increased</b>
a. Interacting with other teachers to improve instruction	1	2	3	4	5
b. Guidance and discipline of students	1	2	3	4	5
c. Completing tasks relative to observations and evaluation	1	2	3	4	5
d. Interacting with parents	1	2	3	4	5
e. Grading papers	1	2	3	4	5
f. Developing lesson plans	1	2	3	4	5



**20. Please indicate the extent to which you agree or disagree with each statement about the new teacher evaluation requirements under the First to the Top reforms. Questions should be answered based on your experience during the current school year (2010-2011).**

*Select one response for each row.*

	<b>Strongly agree</b>	<b>Agree</b>	<b>Disagree</b>	<b>Strongly disagree</b>
a. Tennessee's plans for teacher evaluation have been clearly communicated to me	1	2	3	4
b. Teacher evaluation systems at my school are unfair because of the characteristics of the student population	1	2	3	4
c. Teacher evaluation systems used at this school omit important aspects of teaching that should be considered	1	2	3	4
d. Teacher evaluation systems used at this school contribute to greater collegiality and professionalism among teachers	1	2	3	4
e. Teachers at my school are comfortable with being evaluated on classroom observations	1	2	3	4
f. Parents and this school community believe more rigorous teacher evaluation systems are important	1	2	3	4
g. Teacher evaluation systems are likely to continue for the foreseeable future	1	2	3	4

**21. Generally speaking, what BENEFITS have you encountered with the observation system being used at your school this year (2010-2011)?**

**22. Generally speaking, what CHALLENGES have you encountered with the observation system being used at your school this year (2010-2011)?**

**END OF TEACHER SURVEY. ALL ANSWERING QUESTION 22, END SURVEY HERE.**

**OBSERVER/ADMINISTRATOR QUESTIONS (23-47)**

**23. Including this school year (2010-2011), how long have you served in your current position?**

- a. Total years in your current position
  - b. Total years in your current position in your present district
  - c. Total years in your current position in your present school
- Years

**All ANSWERING QUESTION 23: SKIP TO 25**

**24. Including this school year (2010-2011), how many years have you been employed in an administrative position?**

- a. Total years employed as an administrator
  - b. Total years employed as an administrator in your present district
  - c. Total years employed as an administrator in your present school
- Years

**25. Prior to moving into your current position, were you ever employed as a teacher?**

- a. Yes
- b. No

**SKIP TO 29**

**26. Prior to moving into your current position, how many years did you work as a teacher?**

- a. Total years as a teacher
  - b. Total years as a teacher in your present district
- Years

**27. Which subjects did you teach?**

*Mark all that apply.*

- a. English/language arts/reading
- b. Mathematics
- c. Science
- d. Social studies or history
- e. Foreign language
- f. Special instruction for English Language Learners (ELL) or Limited English Proficient (LEP) students
- g. English as a Second Language (ESL)
- h. Visual or performing arts
- i. Special education
- j. Physical education
- k. Other, specify: \_\_\_\_\_

**28. Which grades did you teach?**

Mark all that apply.

a. 1st grade	f. 6th grade	k. 11th grade
b. 2nd grade	g. 7th grade	l. 12th grade
c. 3rd grade	h. 8th grade	m. Kindergarten
d. 4th grade	i. 9th grade	n. Pre-kindergarten
e. 5th grade	j. 10th grade	o. Ungraded

**29. Are classroom observations used to evaluate teachers' performance in your school?**

- a. Yes
- b. No

SKIP TO 45

**30. Please indicate the model being used for observations in your school.**

Mark only one answer.

- a. Framework for Evaluation and Professional Growth
- b. State of Tennessee Field Test (TAP rubric)
- c. AIMS TIGER Process
- d. COACH Model
- e. I do not know
- f. Other, specify: \_\_\_\_\_

**31. How would you rate the quality of the TRAINING you received to implement this model?**

Mark only one answer.

- a. Excellent
- b. Good
- c. Satisfactory
- d. Insufficient
- e. I did not receive training

**32. How would you describe the professional status of the other classroom observer(s) in this school?**

Mark all that apply.

- a. Principal(s)
- b. Assistant or vice principal(s)
- c. Department head(s)
- d. Math or literacy coach(es)
- e. Other senior teacher(s) from the school, such as a mentor, master, or lead teacher
- f. Observer(s) not working at your school
- g. I am the only one
- h. Other, specify: \_\_\_\_\_

**33. How often are teachers typically observed during the school year?**

*Mark only one answer.*

- a. 1 to 3 times per year
- b. 4 to 5 times per year
- c. 6 to 7 times per year
- d. 8 to 9 times per year
- e. 10 or more times per year
- f. I do not know

**34. During this school year (2010-2011), are you responsible for conducting formal classroom observations of teachers?**

- a. Yes
- b. No

**SKIP TO 44**

**35. How many times during this school year (2010-2011) have YOU formally observed teachers?**

*Mark only one answer.*

- a. Zero (0)
- b. Fewer than 10 times
- c. 10 to 25 times
- d. 26 to 40 times
- e. 41 to 55 times
- f. 56 to 70 times
- g. More than 70 times

**SKIP TO 44**

**36. How much time has passed since you completed your last observation?**

*Do not include weekends or holidays. Only count school days. Mark one answer only.*

- a. Less than 3 school days
- b. 3 to 10 school days
- c. 11 to 20 school days
- d. More than 20 school days

**37. How many minutes did you stay in the classroom during the most recent observation you completed?**

*Mark one answer only.*

- a. Less than 10 minutes
- b. 10 to 15 minutes
- c. 15 to 20 minutes
- d. 20 to 30 minutes
- e. 30 to 40 minutes
- f. More than 40 minutes

**38. Thinking of the most recent observation you completed, did it provide useful information to help the teacher improve his/her teaching?**

- 1. Yes
- 2. No

**39. Thinking of the most recent observation you completed, did it provide useful information to make a judgment about the teacher's performance?**

1. Yes
2. No

**40. Thinking of the most recent observation you completed, did you provide or do you expect to provide VERBAL feedback to the teacher regarding your observation?**

- a. I have provided verbal feedback
- b. I have not provided verbal feedback, but I expect to
- c. I have not provided verbal feedback, and I do not expect to

SKIP TO 42  
SKIP TO 42

**41. Thinking of the most recent observation you completed, how much time passed before you provided VERBAL feedback to the teacher regarding your observation?**

*Do not include weekends or holidays. Only count school days. Mark one answer only.*

- a. Less than 3 school days
- b. More than 3 school days, but less than 10 school days
- c. More than 10 school days, but less than 20 school days
- d. More than 20 school days

**42. Thinking of the most recent observation you completed, did you provide or do you expect to provide WRITTEN feedback to the teacher regarding your observation?**

- a. I have provided formal written feedback
- b. I have not provided formal written feedback, but I do expect to
- c. I have not provided formal written feedback, and I do not expect to

SKIP TO 44  
SKIP TO 44

**43. Thinking of the most recent observation you completed, how much time passed before you provided WRITTEN feedback to the teacher regarding your observation?**

*Do not include weekends or holidays. Only count school days. Mark one answer only.*

- a. Less than 3 school days
- b. 3 to 10 school days
- c. 11 to 20 school days
- d. More than 20 school days

**44. Please indicate how the model being used for observation and evaluation has increased or decreased your WORKLOAD in the following areas during the current school year (2010-2011).**

*Select one response for each row.*

	<b>Significantly decreased</b>	<b>Moderately decreased</b>	<b>No change</b>	<b>Moderately increased</b>	<b>Significantly increased</b>
a. Interacting with teachers to improve instruction	1	2	3	4	5
b. Guidance and discipline of students	1	2	3	4	5
c. Completing tasks relative to teacher evaluation	1	2	3	4	5
d. Interacting with parents	1	2	3	4	5
e. Completing paperwork	1	2	3	4	5
f. Dealing with tasks related to management of the school	1	2	3	4	5

**45. Please indicate the extent to which you agree or disagree with each statement about the new teacher evaluation requirements under the First to the Top reforms. Questions should be answered based on your experience during the current school year (2010-2011).**

*Select one response for each row.*

	<b>Strongly agree</b>	<b>Agree</b>	<b>Disagree</b>	<b>Strongly disagree</b>
a. Tennessee's plans for teacher evaluation have been clearly communicated to me	1	2	3	4
b. Teacher evaluation systems at my school are unfair because of the characteristics of the student population	1	2	3	4
c. Teacher evaluation systems used at this school omit important aspects of teaching that should be considered	1	2	3	4
d. Teacher evaluation systems used at this school contribute to greater collegiality and professionalism among teachers	1	2	3	4
e. Teachers at my school are comfortable with being evaluated on classroom observations	1	2	3	4
f. Parents and this school community believe more rigorous teacher evaluation systems are important	1	2	3	4
g. Teacher evaluation systems are likely to continue for the foreseeable future	1	2	3	4

**46. Generally speaking, what BENEFITS have you encountered with the observation system being used at your school this year (2010-2011)?**

**47. Generally speaking, what CHALLENGES have you encountered with the observation system being used at your school this year (2010-2011)?**

# APPENDIX B1

<b>Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey</b>					
<b>Teacher Responses</b>					
<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
1	Current Position, Principal	4% (53)	3% (35)	5% (43)	3% (42)
1	Current Position, Assistant or Vice Principal	3% (34)	4% (39)	3% (29)	0% (2)
1	Current Position, Instructional coach, mentor, etc	2% (31)	3% (28)	3% (24)	3% (36)
1	Current Position, Teacher	84% (1061)	82% (854)	82% (752)	86% (1211)
1	Current Position, Other	7% (89)	8% (88)	8% (71)	8% (116)
Total # of Teachers, the remaining rows below report results from this population.		1061	854	752	1211
2 (a)	Average years teaching experience	14.1	13.2	14.2	13.6
2 (b)	Teaching experience within present district	11.1	10.2	10.8	10.2
2 (c)	Teaching experience within present school	7.9	7.5	9.2	6.7
3	Percent teaching within elementary (K-5)	48.5%	36.5%	64.8%	46.9%
3	Percent teaching within middle (5-8)	22.1%	19.1%	17.2%	21.7%
3	Percent teaching within high (9-12)	24.8%	38.6%	12.1%	26.0%
3	Percent other	4.5%	5.7%	6.0%	5.4%
Q4: More than one answer per respondent possible					
4	Teacher of English/ Language Arts/ Reading	51%	45%	63%	46%
4	Teacher of Mathematics	45%	40%	57%	43%



**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Teacher Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
4	Teacher of Science	40%	35%	52%	40%
4	Teacher of Social Studies	40%	34%	51%	39%
4	Teacher of Foreign Language	3%	3%	1%	3%
4	Instructor for ELL Students	2%	1%	1%	1%
4	English as a Second Language	2%	1%	2%	1%
4	Teacher of Visual or Performing Arts	7%	6%	5%	6%
4	Special Education Teacher	13%	11%	9%	13%
4	Physical Education Teacher	4%	4%	5%	4%
4	Teacher, Other	11%	16%	8%	11%
5	Expect to be observed 1 to 3 times during year	54% (573)	59% (506)	47% (356)	4% (53)
5	Expect to be observed 4 to 5 times during year	10% (107)	20% (174)	33% (248)	3% (35)
5	Expect to be observed 6 or greater times during year	5% (58)	7% (58)	14% (103)	88% (1065)
5	I do not expect to be formally observed	27% (284)	10% (86)	3% (22)	2% (30)
5	Don't know/ No answer	4% (39)	4% (30)	3% (23)	2% (28)
6	Have been observed 1 to 3 times so far	53% (567)	67% (569)	71% (531)	11% (131)
6	Have been observed 4 to 5 times so far	6% (61)	6% (48)	11% (84)	29% (354)
6	Have been observed 6 or greater times so far	3% (34)	1% (9)	3% (19)	57% (695)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Teacher Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
6	Have not been observed	37% (388)	26% (223)	15% (115)	2% (27)
7	Average # of school days since last observation	25 (528)	24 (528)	20 (515)	13 (1008)
8	Length of last evaluation <10 minutes	5% (58)	3% (22)	5% (35)	24% (289)
8	Length of last evaluation 10-15 minutes	9% (99)	7% (58)	18% (133)	46% (552)
8	Length of last evaluation 15-20 minutes	8% (86)	6% (53)	21% (158)	15% (184)
8	Length of last evaluation 20-30 minutes	8% (89)	9% (77)	22% (166)	7% (82)
8	Length of last evaluation 30-40 minutes	10% (106)	13% (111)	11% (80)	3% (41)
8	Length of last evaluation >40 minutes	21% (218)	36% (305)	8% (62)	2% (30)
8	Length of last evaluation, no response	38% (405)	27% (228)	16% (118)	3% (33)
9	Average time of observation <10 minutes	5% (50)	2% (20)	5% (34)	22% (264)
9	Average time of observation 10-15 minutes	9% (94)	5% (40)	17% (129)	47% (571)
9	Average time of observation 15-20 minutes	8% (87)	8% (68)	22% (164)	18% (214)
9	Average time of observation 20-30 minutes	9% (99)	13% (114)	24% (182)	7% (79)
9	Average time of observation 30-40 minutes	13% (142)	18% (152)	11% (82)	2% (23)
9	Average time of observation >40 minutes	17% (183)	27% (228)	5% (41)	2% (25)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Teacher Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
9	Average time of observation, no re-sponse	38% (406)	27% (232)	16% (120)	3% (35)
Q10: More than one answer per respondent possible					
10	Observed by principal	52% (547)	53% (449)	66% (493)	90% (1094)
10	Observed by assistant or vice principal	36% (387)	43% (369)	41% (309)	77% (934)
10	Observed by department head	5% (51)	9% (78)	13% (97)	4% (53)
10	Observed by math or literacy coach	7% (79)	10% (85)	9% (65)	8% (102)
10	Observed by other senior teacher	4% (42)	5% (44)	10% (77)	3% (37)
10	Observed by someone not from your school	8% (81)	12% (100)	7% (56)	10% (117)
10	Observed by Other	7% (73)	7% (59)	8% (59)	4% (46)
11	Avg # observers, most recent observation	1.1 (647)	1.2 (618)	1.1 (629)	1.1 (1168)
12	Feedback from most recent observation designed to improve your teaching = 'Yes'	85% (557)	89% (549)	84% (531)	84% (976)
12	Feedback from most recent observation designed to improve your teaching = 'No'	15% (95)	11% (67)	16% (98)	16% (182)
13	Feedback from most recent observation makes a judgment about performance = 'Yes'	74% (487)	75% (461)	70% (437)	66% (772)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Teacher Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
13	Feedback from most recent observation makes a judgment about performance = 'No'	26% (171)	25% (156)	30% (187)	34% (389)
14	Verbal feedback received	76% (497)	72% (452)	55% (349)	76% (898)
14	Verbal feedback not received but expected	13% (86)	19% (119)	31% (195)	10% (123)
14	Verbal feedback not received, not expected	11% (73)	9% (55)	14% (87)	13% (155)
16	Written feedback received	71% (464)	53% (333)	55% (344)	79% (929)
16	Written feedback not received but expected	16% (103)	28% (177)	32% (203)	14% (169)
16	Written feedback not received, not expected	13% (82)	18% (115)	13% (82)	6% (71)
15	Verbal feedback received in <3 school days	67% (329)	61% (275)	68% (235)	90% (805)
15	Verbal feedback received in 3-10 school days	24% (119)	33% (149)	27% (93)	9% (84)
15	Verbal feedback received in 11-20 school days	4% (22)	3% (15)	4% (13)	1% (5)
15	Verbal feedback received in >20 school days	4% (22)	3% (12)	2% (7)	0% (1)
17	Written feedback received in <3 school days	64% (295)	57% (188)	59% (200)	82% (754)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Teacher Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
17	Written feedback received in 3-10 school days	27% (126)	37% (121)	34% (115)	16% (151)
17	Written feedback received in 11-20 school days	6% (27)	4% (13)	5% (17)	1% (12)
17	Written feedback received in >20 school days	3% (15)	2% (8)	2% (7)	1% (7)
18	Teacher reports current model is Framework	70% (709)	27% (216)	6% (41)	11% (127)
18	Teacher reports current model is Field Test	2% (25)	50% (404)	3% (24)	4% (43)
18	Teacher reports current model is AIMS TIGER	1% (8)	0% (3)	83% (599)	0% (0)
18	Teacher reports current model is COACH	0% (5)	0% (1)	1% (4)	71% (814)
18	Teacher reports current model other	26% (269)	23% (182)	7% (54)	14% (163)
<p>Percentage shown is the percent of respondents that responded that the model being utilized had 'Moderately increased' or 'Significantly increased' their workload in the following area. Other answer options were 'Significantly decreased', 'Moderately decreased', and 'No change'.</p>					
19 (a)	Workload change because of interacting with other teachers to improve instruction	32% (997)	38% (807)	43% (736)	30% (1183)
19(b)	Workload change because of guidance and discipline of students	21% (991)	19% (805)	18% (734)	20% (1176)
19 (c)	Workload change because of completing tasks relative to observations and evaluation	52% (991)	48% (803)	69% (734)	35% (1178)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Teacher Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
19 (d)	Workload change because of interacting with parents	19% (989)	15% (805)	13% (731)	14% (1176)
19 (e)	Workload change because of grading assignments	20% (988)	19% (803)	17% (733)	14% (1175)
19 (f)	Workload change because of developing lesson plans	44% (992)	53% (805)	45% (733)	32% (1178)
Percentage shown is the percent of respondents that responded 'Strongly agree' or 'Agree' to the following beliefs. Other answer options were 'Strongly disagree' and 'Disagree'.					
20 (a)	Belief: Tennessee's plans for teacher evaluation have been clearly communicated to me	50% (1022)	73% (825)	74% (738)	81% (1193)
20 (b)	Belief: Teacher evaluation systems at my school are unfair because of the characteristics of the student population	30% (998)	25% (792)	26% (729)	20% (1173)
20 (c)	Belief: Teacher evaluation systems used at this school omit important aspects of teaching that should be considered	45% (992)	41% (796)	36% (733)	37% (1163)
20 (d)	Belief: Teacher evaluation systems used at this school contribute to greater collegiality and professionalism among teachers	51% (995)	58% (802)	61% (724)	55% (1167)
20 (e)	Belief: Teachers at my school are comfortable with being evaluated on classroom observations	76% (1011)	66% (808)	78% (735)	76% (1174)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Teacher Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
20 (f)	Belief: Parents at this school community believe more rigorous teacher evaluation systems are important	37% (972)	40% (779)	46% (711)	52% (1139)
20 (g)	Belief: Teacher evaluation systems are likely to continue for the foreseeable future	96% (1015)	97% (820)	97% (734)	97% (1177)

# APPENDIX B2

<b>Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey</b>					
<b>Evaluator Responses</b>					
<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
1	Current Position, Principal	4% (53)	3% (35)	5% (43)	3% (42)
1	Current Position, Assistant or Vice Principal	3% (34)	4% (39)	3% (29)	0% (2)
1	Current Position, Instructional coach, mentor, etc	2% (31)	3% (28)	3% (24)	3% (36)
1	Current Position, Teacher	84% (1061)	82% (854)	82% (752)	86% (1211)
1	Current Position, Other	7% (89)	8% (88)	8% (71)	8% (116)
34	Did you serve as an evaluator? Principals	100% (53)	100% (35)	100% (43)	95% (40)
34	Did you serve as an evaluator? Asst Principals	91% (31)	92% (36)	90% (26)	100% (2)
34	Did you serve as an evaluator? Coaches/Mentor	0% (0)	25% (7)	33% (8)	8% (3)
Total # of Evaluators, the remaining rows below report results from this population.		84	78	77	45
23/4 (a)	Average # years as administrator/coach	9.3	6.8	8.1	9.7
23/4 (b)	# of years administrator/coach, present district	8.6	5.9	7.4	8.6
23/4 (c)	# of years administrator/coach, present school	5.6	4.6	5.9	4.4
25	Has previous teaching experience	94% (79)	96% (75)	95% (73)	91% (41)
26 (a)	Average years teaching experience	15.2	13.2	13.9	12.0
26 (b)	Teaching experience, present district	13.5	12.7	11.1	10.6



<b>Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey</b>					
<b>Evaluator Responses</b>					
<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
Q27: More than one answer per respondent possible					
27	Taught English/ Language Arts/ Reading	56%	44%	58%	53%
27	Taught Mathematics	58%	45%	55%	58%
27	Taught Science	55%	36%	61%	53%
27	Taught Social Studies	50%	40%	60%	51%
27	Taught Foreign Language	2%	0%	0%	0%
27	Instructed ELL Students	0%	1%	0%	0%
27	English as a Second Language	2%	1%	0%	0%
27	Taught Visual or Performing Arts	4%	5%	4%	2%
27	Taught Special Education	10%	9%	10%	16%
27	Taught Physical Education	10%	17%	18%	11%
27	Taught Other Subject	17%	19%	14%	11%
28	Percent taught within elementary (K-5)	26%	12%	22%	42%
28	Percent taught within middle (5-8)	13%	26%	23%	7%
28	Percent taught within high (9-12)	14%	28%	20%	18%
28	Percent multiple/other	46%	35%	35%	33%
31	Quality of training on model used Excellent	23% (19)	26% (20)	16% (12)	22% (10)
31	Quality of training on model used Good	54% (45)	55% (43)	42% (32)	42% (19)
31	Quality of training on model used Satisfactory	20% (17)	17% (13)	18% (14)	31% (14)
31	Quality of training on model used Insufficient	2% (2)	1% (1)	22% (17)	4% (2)
31	I did not receive training	0% (0)	1% (1)	3% (2)	0% (0)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Evaluator Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
Q32: More than one answer per respondent possible					
32	In my school, Principal used as evaluator	96% (81)	92% (72)	94% (72)	98% (44)
32	In my school, AP(s) used as evaluators	77% (65)	80% (62)	69% (53)	71% (32)
32	In my school, Dept Heads used as evaluators	8% (7)	8% (6)	17% (13)	4% (2)
32	In my school, Mth/Lit Coach used as evaluators	1% (1)	8% (6)	18% (14)	7% (3)
32	In my school, Lead Teacher(s) used as evaluators	5% (4)	10% (8)	16% (12)	2% (1)
32	In my school, External Observers used as evaluators	16% (13)	10% (8)	10% (8)	13% (6)
32	In my school, I am the only evaluator	5% (4)	3% (2)	5% (4)	2% (1)
32	In my school, Other used as evaluators	7% (6)	14% (11)	9% (7)	2% (1)
33	My teachers typical are evaluated 1-3 times	77% (65)	53% (41)	34% (26)	13% (6)
33	My teachers typical are evaluated 4-5 times	7% (6)	31% (24)	38% (29)	2% (1)
33	My teachers typical are evaluated >6 times	15% (13)	17% (13)	27% (21)	84% (38)
35	I have completed fewer than 10 tch evaluations	24% (20)	33% (26)	30% (23)	20% (9)
35	I have completed 10-25 teacher evaluations	55% (46)	46% (36)	42% (32)	9% (4)
35	I have completed 26-40 teacher evaluations	18% (15)	12% (9)	18% (14)	7% (3)
35	I have completed more than 40 tch evaluations	2% (2)	9% (7)	10% (8)	64% (29)
36	I completed my last observation < 3 days ago	32% (27)	40% (31)	47% (36)	71% (32)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Evaluator Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
36	I completed my last observation 3-10 days ago	36% (30)	35% (27)	34% (26)	24% (11)
36	I completed my last observation 11-20 days ago	21% (18)	14% (11)	13% (10)	2% (1)
36	I completed my last observation > 20 days ago	10% (8)	12% (9)	4% (3)	2% (1)
37	Average time of observation <10 minutes	0% (0)	1% (1)	0% (0)	42% (19)
37	Average time of observation 10-15 minutes	2% (2)	1% (1)	9% (7)	38% (17)
37	Average time of observation 15-20 minutes	4% (3)	5% (4)	30% (23)	4% (2)
37	Average time of observation 20-30 minutes	6% (5)	5% (4)	21% (16)	4% (2)
37	Average time of observation 30-40 minutes	29% (24)	12% (9)	25% (19)	0% (0)
37	Average time of observation >40 minutes	58% (49)	76% (59)	13% (10)	9% (4)
38	Feedback from most recent observation designed to improve your teaching = 'Yes'	93% (77)	99% (77)	92% (68)	98% (41)
39	Feedback from most recent observation makes a judgment about performance = 'Yes'	92% (76)	95% (74)	96% (72)	100% (44)
40	Verbal feedback provided	83% (69)	72% (56)	65% (49)	89% (39)
40	Verbal feedback not provided but I expect to	17% (14)	27% (21)	31% (23)	11% (5)
40	Verbal feedback not provided, not expected	0% (0)	1% (1)	4% (3)	0% (0)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

<b>Evaluator Responses</b>					
<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
42	Written feedback provided	78% (65)	57% (43)	55% (41)	77% (34)
42	Written feedback not provided but I expect to	19% (16)	38% (29)	41% (30)	20% (9)
42	Written feedback not provided, not expected	2% (2)	5% (4)	4% (3)	2% (1)
41	Verbal feedback provided in <3 school days	77% (53)	82% (46)	65% (32)	100% (39)
41	Verbal feedback provided in 3-10 school days	20% (14)	16% (9)	33% (16)	0% (0)
41	Verbal feedback provided in 11-20 school days	1% (1)	0% (0)	2% (1)	0% (0)
41	Verbal feedback provided in >20 school days	1% (1)	2% (1)	0% (0)	0% (0)
43	Written feedback provided in <3 school days	66% (43)	84% (36)	66% (27)	79% (27)
43	Written feedback provided in 3-10 school days	31% (20)	14% (6)	32% (13)	21% (7)
43	Written feedback provided in 11-20 school days	2% (1)	0% (0)	2% (1)	0% (0)
43	Written feedback provided in >20 school days	2% (1)	2% (1)	0% (0)	0% (0)
Percentage shown is the percent of respondents that responded that the model being utilized had 'Moderately increased' or 'Significantly increased' their workload in the following area. Other answer options were 'Significantly decreased,' 'Moderately decreased,' and 'No change.'					
44 (a)	Workload change because of interacting with other teachers to improve instruction	57% (82)	74% (78)	78% (77)	87% (45)
44(b)	Workload change because of guidance and discipline of students	26% (82)	18% (78)	18% (76)	23% (44)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Evaluator Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
44 (c)	Workload change because of completing tasks relative to teacher evaluation	66% (82)	86% (77)	84% (76)	73% (44)
44 (d)	Workload change because of interacting with parents	17% (81)	12% (78)	5% (77)	11% (44)
44 (e)	Workload change because of completing paperwork	74% (82)	78% (77)	86% (77)	53% (45)
44 (f)	Workload change because of dealing with tasks related to management of the school	51% (81)	45% (78)	36% (76)	29% (45)
Percentage shown is the percent of respondents that responded 'Strongly agree' or 'Agree' to the following beliefs. Other answer options were 'Strongly disagree' and 'Disagree.'					
45 (a)	Belief: Tennessee's plans for teacher evaluation have been clearly communicated to me	36% (81)	75% (77)	78% (77)	59% (44)
45 (b)	Belief: Teacher evaluation systems at my school are unfair because of the characteristics of the student population	9% (82)	4% (77)	5% (77)	7% (44)
45 (c)	Belief: Teacher evaluation systems used at this school omit important aspects of teaching that should be considered	44% (81)	22% (77)	17% (77)	9% (45)
45 (d)	Belief: Teacher evaluation systems used at this school contribute to greater collegiality and professionalism among teachers	50% (80)	74% (77)	79% (77)	87% (45)
45 (e)	Belief: Teachers at my school are comfortable with being evaluated on classroom observations	88% (80)	82% (77)	97% (77)	91% (45)

**Descriptive Statistics from the 2011 Tennessee Educator Evaluation Survey**

**Evaluator Responses**

<b>Ques #</b>	<b>Question Description</b>	<b>Field Test Control (Framework)</b>	<b>Field Test Treatment (TAP Rubric)</b>	<b>AIMS TIGER</b>	<b>COACH Model</b>
45 (f)	Belief: Parents at this school community believe more rigorous teacher evaluation systems are important	46% (78)	69% (75)	76% (76)	70% (43)
45 (g)	Belief: Teacher evaluation systems are likely to continue for the foreseeable future	98% (81)	97% (77)	100% (77)	98% (45)

# Tennessee Consortium on Research, Evaluation, & Development

Peabody #44 | 230 Appleton Place | Nashville, Tennessee 37203

Phone 615-322-5538 | Fax 615-322-6018

[www.tnconsortium.org](http://www.tnconsortium.org)

TENNESSEE  
**CONSORTIUM**  
ON RESEARCH, EVALUATION & DEVELOPMENT

 **VANDERBILT**  
PEABODY COLLEGE