



An Analysis of Teacher Evaluation Data and Teacher Characteristics

Amanda Klafehn, Associate Legislative Research Analyst
(615) 401-7859 / Amanda.Klafehn@cot.tn.gov

February 2015

Key Points

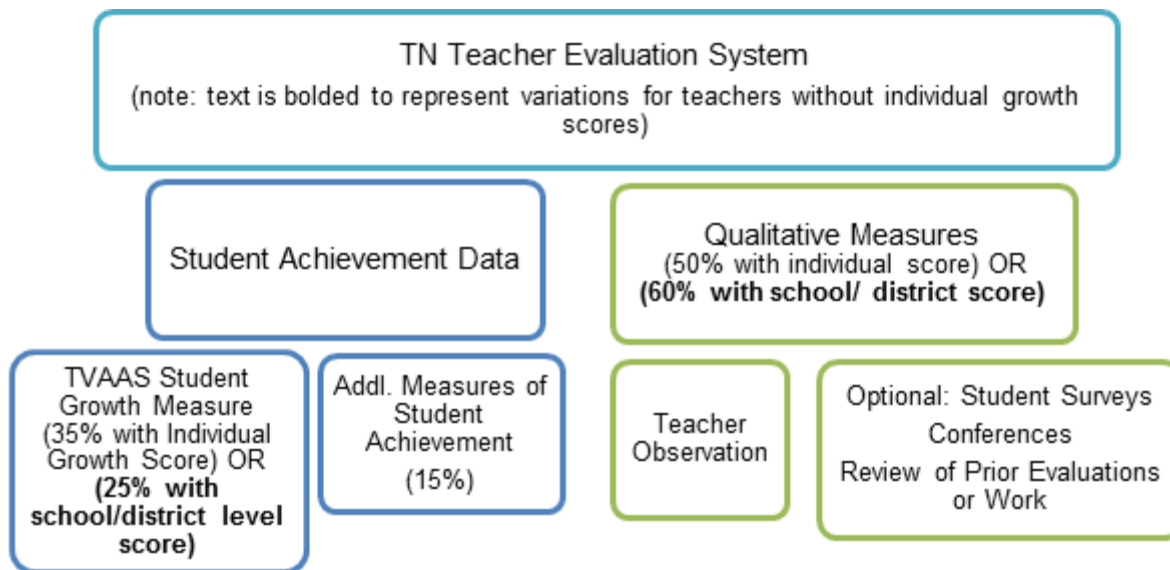
- This brief analyzes individual value-added and qualitative composite scores relative to several teacher characteristics (race, gender, level of education, age, years of experience, and salary). Data is from the 2011-12 and 2012-13 academic years.
- Across both years, approximately 75 percent of teachers received a level 3 or above for the individual value-added component of their evaluation, and approximately 97 percent of teachers received a level 3 or above on their qualitative composite (which includes observation, surveys, etc.). A level 3 indicates the teacher is “At Expectations.”
- In 2012, approximately 31 percent of the teaching population received individual growth scores. In 2013, this number increased, and roughly 44 percent of the teaching population received individual growth scores.

Introduction

As requested by the Tennessee House Education Committee, the Comptroller's Offices of Research and Education Accountability (OREA) examined the distribution of teacher evaluation scores relative to teacher characteristics (race, gender, level of education, age, years of experience, and salary). Variations in the distribution of evaluation scores across teacher characteristics are illustrated, but the analysis does not consider possible reasons for any identified variations.

Tennessee's teacher evaluation system comprises two parts: student achievement measures and qualitative measures. Student achievement measures are divided into two sub-categories: student growth, as represented by the Tennessee Value-Added Assessment System (TVAAS), and additional achievement measures (e.g., ACT/SAT assessments or graduation rate), which are selected by teachers and their evaluators from a list approved by the State Board of Education.¹ The qualitative component of the evaluation primarily comprises classroom observations, but may also include student surveys, personal conferences between the teacher and evaluator, and prior evaluations.^{2,3} [Exhibit 1](#) provides specific weights for the metrics used for teacher evaluation in Tennessee. These metrics and weights were in place for the 2012-13 academic year.⁴ For more information on TVAAS and the evaluation model in Tennessee see [Use of Value Added in Teacher Evaluation: Key Concepts and State Profiles](#).

Exhibit 1: Metrics and Weighting in the State Teacher Evaluation Model



The State Board of Education (SBOE) has approved the statewide Tennessee Educator Acceleration Model (TEAM) and four additional locale-specific models: Project COACH, the Teacher Effectiveness Model (TEM), the Teacher Instructional Growth for Effectiveness and Results (TIGER), and the Achievement Framework for Excellent Teaching (AFET).⁵ Each model is subject to the same student achievement weights; however, the qualitative measures (e.g., classroom observations, student perception surveys, personal conferences, and/or review of prior evaluations) that are used and the weight assigned to each measure can vary.

Exhibit 2 provides more detail on the statewide teacher evaluation model and the four approved alternative models. The primary differences among the models are their observation processes and the rubric used.⁶ Alternative models may also incorporate different qualitative measures, such as student surveys; the TEM model, used in Shelby County Schools, includes student surveys and teacher professionalism as part of the qualitative component. The TDOE permits some local evaluation decisions, including the use of TRIPOD student perception surveys as 5 percent of the overall teacher evaluation score under TEAM flexibility.⁷ The number of districts using student surveys in their teacher evaluation model has grown: in 2013-14, 17 districts piloted the use of student surveys, and in 2014-15, 19 districts will use student surveys as part of the qualitative component.^{8,9} The Achievement School District (ASD), which seeks to improve some of the lowest performing schools in Tennessee, and uses the AFET model, assigns the highest weight to student surveys, at 15 percent, and, correspondingly, lowers the observation score weight to 35 percent, which is the lowest of the models, for teachers with individual value added data (SY 2013-14).

Quantitative and Qualitative Measures

Throughout this memo, there are references to the Tennessee Value-Added Assessment System (TVAAS) and to qualitative composite measures. TVAAS is a statistical estimation based on student test score data, and is a quantitative measure. The qualitative composite measure comprises observation scores, primarily, but may also include student surveys or conferences. These measures are obtained via interviews and direct observation and are considered qualitative.

For the remainder of this analysis, the only models that will be referenced are TEAM, TEM, Project COACH, and TIGER. The AFET Model is not included because the model was approved for use in August 2013, and has no data for the time period (school years 2011-12 and 2012-13) covered in the analysis.¹⁰

Exhibit 2: Teacher Evaluation Models^(a)

Teacher Evaluation Model and Approval Year	Locale	Number of Teachers ^(d)	Student Achievement Weights	Qualitative Measures	Differences in Observation Process and Rubric	Notable Characteristics/Theory
TEAM (2011)	Statewide model	2011-12 55,600	<p>Individual Growth Scores: -35% TVAAS growth -15% Achievement Measure</p> <p>Without Individual Growth Scores: -25% TVAAS Growth (at school/district level) -15% Achievement Measure</p>	<p>Individual Growth Scores: 50% observation using TAP Teaching Standards</p> <p>Without Individual Growth Scores: 60% observation using TAP Teaching Standards</p>	Scoring on 1-5 scale; frequent observations with feedback through pre- and post-conferences	Statewide model established under Tennessee First To the Top Act (2010) that was to include student achievement data in teacher evaluations
		2012-13 55,256				
Project COACH (2011)	Hamilton County Schools and Bradley County Schools	2011-12 3,797	Same	<p>Individual Growth Scores: 50% observation</p> <p>Without Individual Growth Scores: 60% observation</p>	Mini-observations of a minimum of 10 minutes throughout the year, followed with communication and actionable feedback between administrator and teacher; scoring on a 1-4 scale ^(b)	Many mini-observations and feedback loops to provide instructional support to teachers
		2012-13 4,173				
TEM 3.0 (2011)	Shelby County Schools- (former Memphis City Schools)	2011-12 6,430	Same	<p>Individual Growth Scores: -40% observation -5% student surveys (if survey is unavailable, observation weight is 45%)</p> <p>Without Individual Growth Scores: -5% professionalism</p>	Observations and post-conferences are conducted to determine appropriate next steps; scoring on a 1-5 scale	Includes student surveys and a professionalism metric, in addition to observation data
		2012-13 5,664				

Teacher Evaluation Model and Approval Year	Locale	Number of Teachers ^(d)	Student Achievement Weights	Qualitative Measures	Differences in Observation Process and Rubric	Notable Characteristics/Theory
TIGER (2011)	12 TN School Districts	2011-12 1,634	Same	Individual Growth Scores: 50% observation	Scoring on 1-4 scale; coaching and walkthroughs to support teachers at varying stages ^(c)	Designed to promote teacher growth and ensure teachers provide quality instruction; includes a coaching component for teachers at various levels for improvement
		2012-13 1,666		Without Individual Growth Scores: 60% observation		
The Achievement Framework for Excellent Teaching (AFET) (2013)	ASD Schools	No Data Available	Same	Individual Growth Scores: -35% observation -15% student surveys Without Individual Growth Scores: -40% observation -20% student surveys	Using the ASD's Framework for Excellent Teaching, teachers will be rated on a 1-4 scale; differential weighting of evaluation based upon chronology of evaluations (last two rounds weighted more); minimum of 4 observations	An approach designed to be supportive of evaluators and teachers in meaningful conversation on teacher performance

Notes: (a) Appendix 1 provides a brief analysis of the evaluation model types and distribution of the qualitative composite scores.

(b) The Project COACH model uses a derived version of Marshall's model for a rubric; Marshall's Model comprises four parts: mini-observations, team curriculum planning, a team interim assessment, and end of the year rubric. Project COACH spent time modifying and adjusting the indicators for Marshall's model to make them their own.

(c) The TIGER model uses a derived version of Danielson's model. Charlotte Danielson developed a rubric that TIGER creators adapted and modified for their teacher evaluation.

(d) The data included in the Number of Teachers column is drawn from Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Sources: Tennessee State Board of Education, *Teacher and Principal Evaluation Policy 5.201*, effective July 25, 2014, <http://www.tn.gov/sbe/> (accessed Dec. 17, 2014); State Collaborative on Reforming Education, *Supporting Effective Teaching in Tennessee: Listening and Gathering Feedback on Tennessee's Teacher Evaluations*, 2012, p. 3; Tennessee Department of Education, "Teacher Evaluation," <http://team-tn.org/>; Angela Lawson, "Quick Facts Overview", April 2012, <http://www.bradleyschools.org/> (accessed Dec. 16, 2014); Jim Scales and Connie Cloud Atkins, *Hamilton County Department of Education: Rethinking Teacher Evaluation through Project COACH*, District Management Council, 2011, pp. 9-10; Teacher Effectiveness Initiative, "Update on Teacher Effectiveness Measure (TEM) for 2013-14"; Shelby County Schools, "Teacher Effectiveness Measure (TEM3): Student Growth & Achievement FAQs," <http://www.scsk12.org/uf/>; Teacher Instructional Growth for Effectiveness and Results, "About Tiger," 2014, <http://tigermode.net/about/>; The TIGER Model for Teacher Evaluation: *Experiences from the Field: LEAD Conference*, Oct. 27, 2014, <http://tn.gov/education/> (accessed 12/16/2014); Teacher Instructional Growth for Effectiveness and Results, "TIGER Frequently Asked Questions," p. 9; Tennessee Achievement School District, "TN ASD Teacher Evaluation Proposed System Revisions," pp. 1-4.

Methodology

The dataset for this analysis was provided to OREA by the Tennessee Department of Education (TDOE) and consists of teacher evaluation data from the 2011-2012 and 2012-2013 school years. OREA analyzed the data by comparing teacher evaluation score levels, which range from 1 to 5, relative to teacher characteristic variables (race, gender, level of education, age, years of experience, and salary). See [Appendix 2](#) for greater detail on the methodology.

Teacher Evaluation Data Source

The dataset provided to OREA was a compilation of several sources from the TDOE. The dataset included all evaluation components, age, years of experience, salary, race, gender, and level of education variables for each unidentifiable case over the 2011-12 and 2012-13 school years. See Appendix 2 for more information. Throughout the remainder of the analysis, the dataset will be sourced as Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received August 22, 2014.

This method of analysis shows the distribution of teacher evaluation scores across characteristics while also accounting for significant population differences among demographic groupings (e.g., approximately 80 percent of the Tennessee teaching population is female). TDOE has employed similar tactics to compare teacher evaluation scores with teacher characteristics. In 2013, the department reviewed the distribution of value-added scores relative to the following teacher characteristics: years of experience, education level, and salary.¹¹

Each teacher characteristic variable was reviewed against individual value-added scores and qualitative composite scores,^A both of which can be directly attributed to an individual teacher. School or district level value-added and school achievement measures, by contrast, cannot be directly attributed to individual teachers or their characteristics; because of this, there is no analysis of school or district level growth scores and teacher characteristics.

Individual value-added and observation score levels have the following designations:^B

- 5 – Significantly Above Expectations
- 4 – Above Expectations
- 3 – At Expectations
- 2 – Below Expectations
- 1 – Significantly Below Expectations

^A Tony Pratt, Deputy Assistant Commissioner of Data and Research, Tennessee Department of Education, e-mail, Oct. 1, 2014. The qualitative composite score combines observation scores, surveys, personal conferences, etc., into one variable.

^B The dataset provided by TDOE converted the observation scores in other models to be comparable to the statewide model (1-5 scoring range).

Teacher Evaluation Data

Exhibit 3 provides some descriptive information on the Tennessee teaching population across this time period.

Exhibit 3: Descriptive Information from TDOE Dataset (calculations exclude missing data)

	2011-12	2012-13
Total Number of Individuals Evaluated as Teachers^(a)	67,461	66,759
Percent Female	79.9%	80.0%
Percent White	87.3%	88.0%
Percent with Master's Degrees	49.6%	52.2%
Average^(b) Age/ Median^(c) Age	42 years/ 41 years	42 years/ 41 years
Average Years of Experience/ Median Years of Experience	14 years/ 12 years	12 years/ 10 years
Average Salary	\$45,857	\$47,778

Notes: (a) The count of individuals evaluated as teachers could include additional individuals, beyond the traditional teacher, such as school counselors, social workers, or instructional coaches who teach part time.

(b) Average: the average age is calculated by totaling each teacher's age and dividing that by the total number of teachers (the sample for this statistic includes only teachers who have age data, n=60,117 (2012) and n=59,380 (2013)).

(c) Median: the median age is the point at which 50% of the cases fall above that age and 50% of the cases fall below that age or the middle value.

Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

In 2012, approximately 31 percent (n= 20,595) of the teaching population (see Exhibit 3, Note a) received individual growth scores; 100 percent of the teaching population received final qualitative scores. In 2013, the percent of the teaching population that received individual growth scores rose to 44 percent (n=29,664) and 100 percent of the same population received final qualitative scores.¹² TDOE staff indicate the increase in individual value-added scores between the two years may be attributed to the introduction of new end-of-course (EOC) exams, the use of SAT-10 assessments in grades K-2 (district option), and/or the introduction of portfolio models in World Languages and Fine Arts (district option).¹³ Teachers who did not have individual growth scores were not included in the value-added portion of this analysis.

Individual Value-Added Growth Scores

TVAAS is a statistical estimation of a teacher's impact on student growth in one academic year, which is converted to an evaluation score level of 1 to 5. Only teachers who teach in subject areas and grade levels with statewide assessments (TCAP and EOC), teachers who teach in grades K-2 whose district has approved the use of SAT-10 tests, or teachers in subject areas with approved portfolio models, receive individual growth scores.^c A review of the distribution of individual value-added scores indicates that a larger proportion of teachers receive either levels 3 or 5.

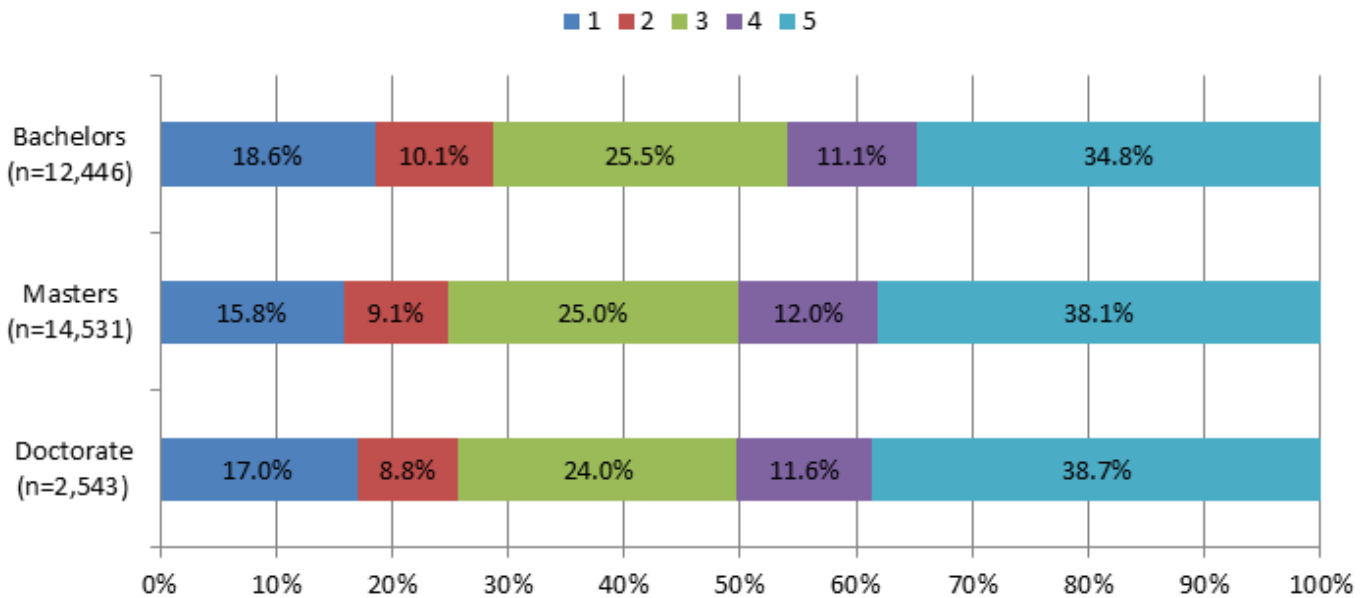
^c This could include: SAT-10 tests in grades K-2; TCAP tests in grades 3-8; EOCs in middle and high school grades; alternative portfolio models in world languages, fine arts, or physical education.

Exhibit 4: Distribution of All Individual Value-Added Scores

Individual Value-Added Distribution (Overall)	% with level 1	% with level 2	% with level 3	% with level 4	% with level 5
2011-12	17.1%	8.4%	24.5%	11.8%	38.3%
2012-13	17.1%	9.5%	25.1%	11.6%	36.8%

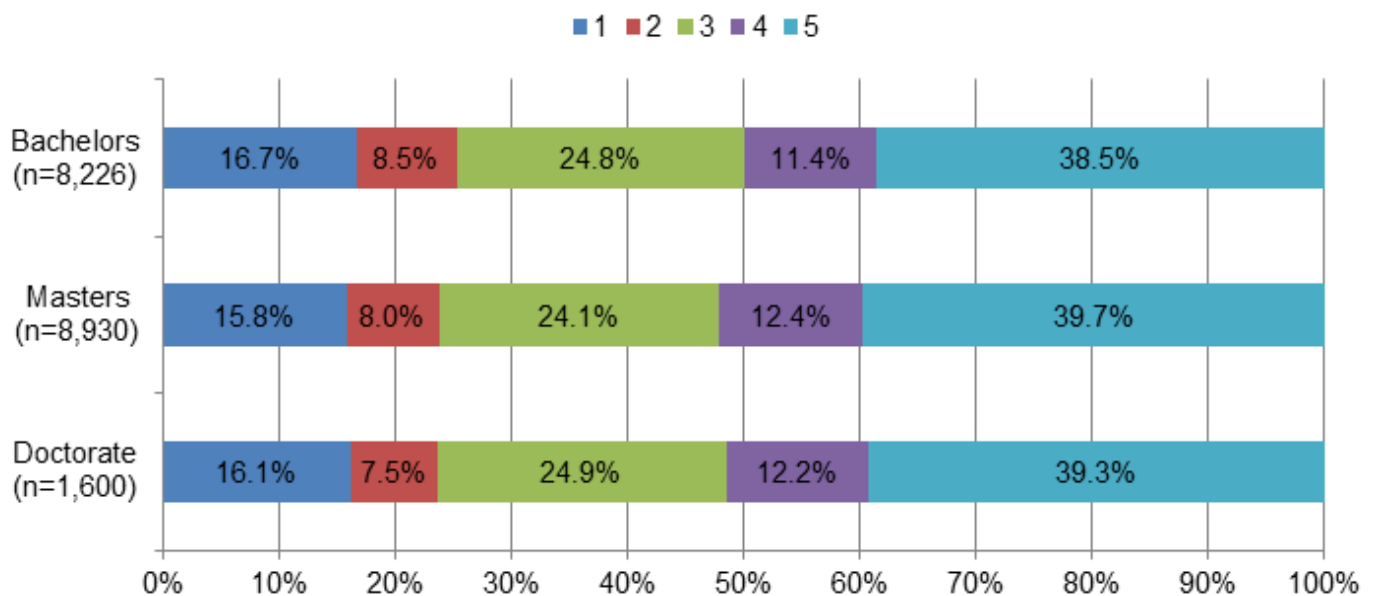
Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Exhibit 5: Individual Growth Score by Level of Education, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Exhibit 6: Individual Growth Score by Level of Education, 2011-12**

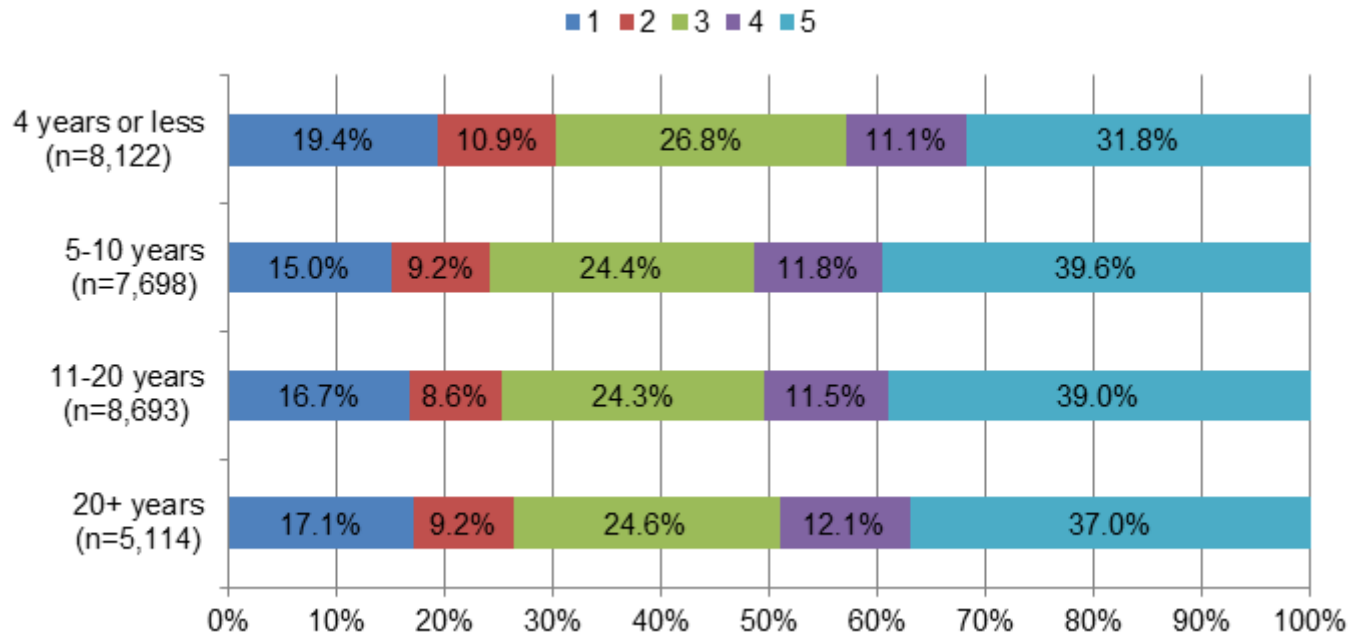


** Data on the level of education for teachers under the TEM evaluation model was missing for 2012. This makes up approximately 8.3% of teachers with individual growth scores.

Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Approximately 39 percent of teachers within the Bachelors, Masters, and Doctoral degree categories received a level 5 for their individual growth score in 2012. Additionally, approximately 16 percent of teachers across each grouping received a level 1. The findings for 2013 follow a similar distribution, with approximately 37 percent of teachers receiving a level 5 across each degree category and roughly 17 percent of teachers across each category receiving a level 1.

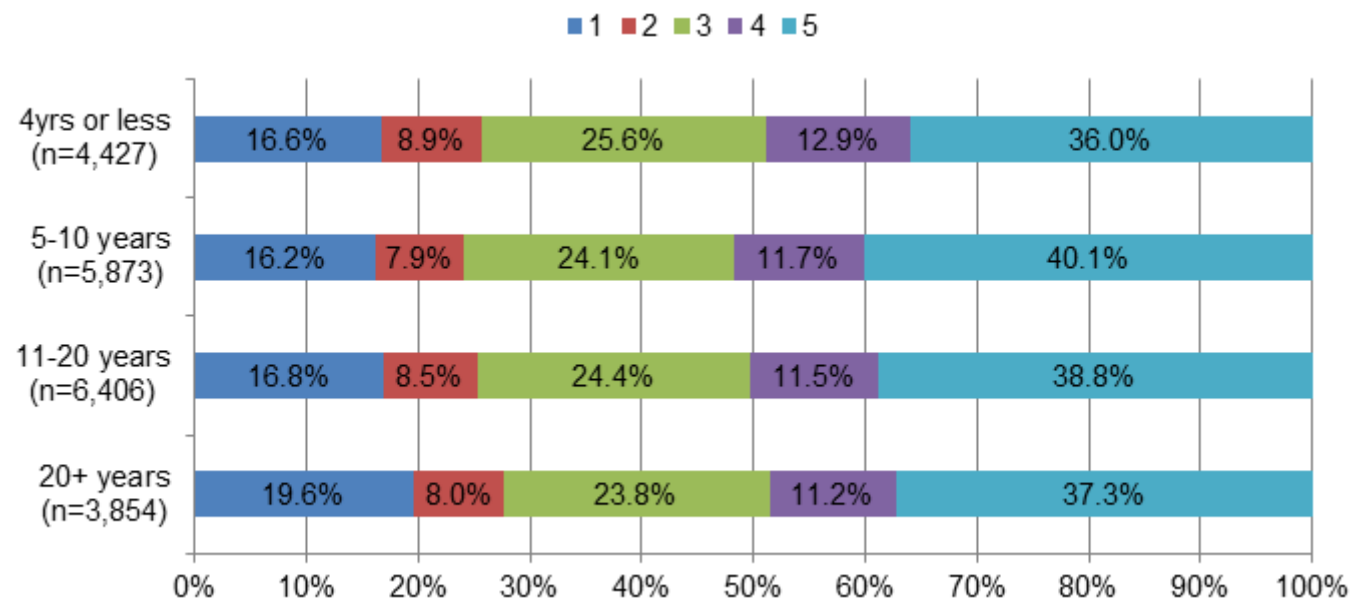
Exhibit 7: Individual Growth Score by Years of Experience, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

In 2013, the proportion of teachers across each category receiving individual growth score level 5 is approximately 37 percent, whereas, the proportion of teachers receiving a level 1 is 17 percent across each years of experience category.

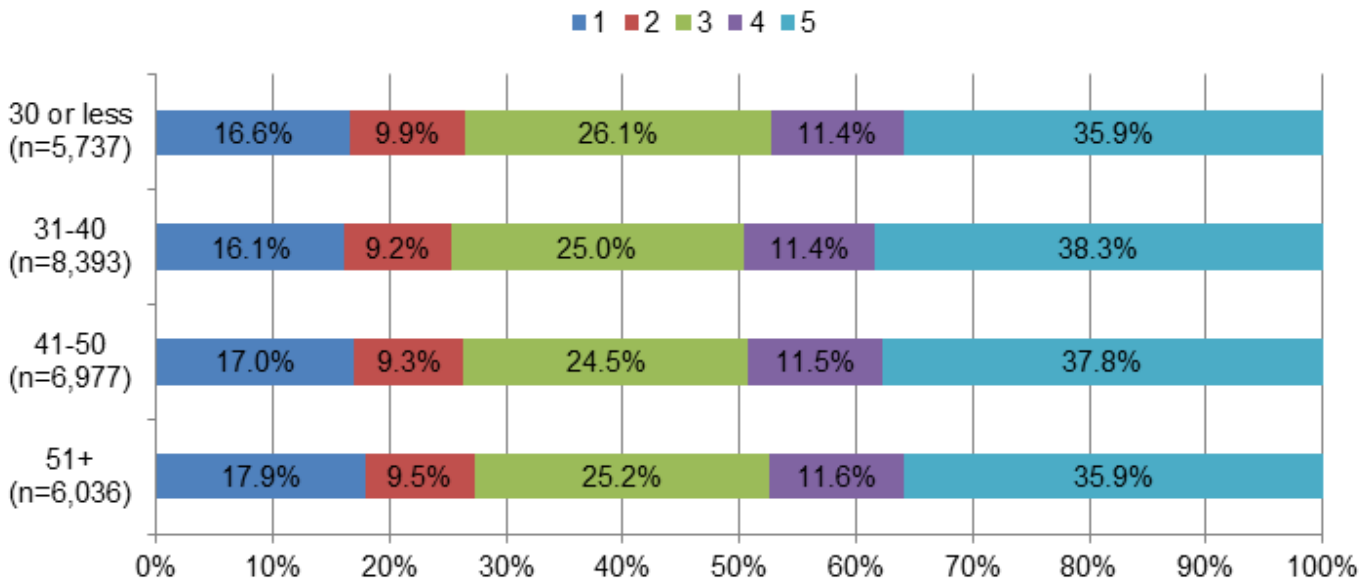
Exhibit 8: Individual Growth Score by Years of Experience, 2011-12



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

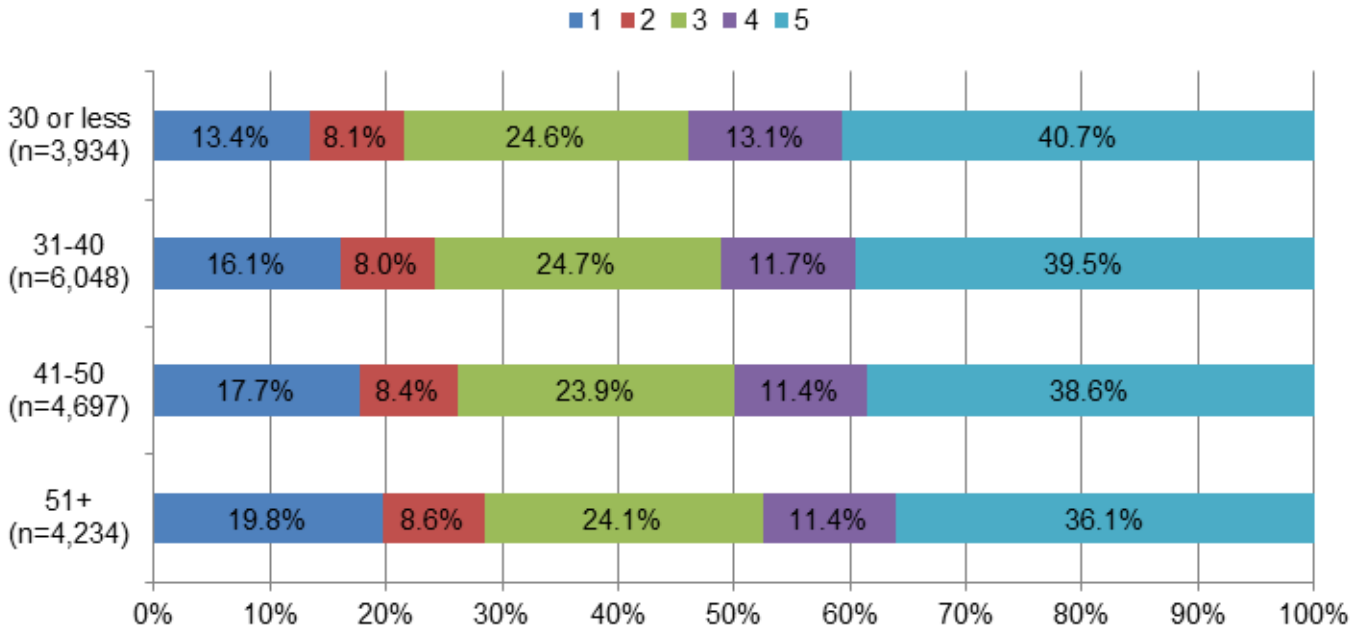
In 2012, the proportion of teachers receiving an individual growth score level 5 across each category is approximately 38 percent. Correspondingly, the proportion of teachers across each category receiving a level 1 for their individual growth score is approximately 17 percent.

Exhibit 9: Individual Growth Score by Age, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

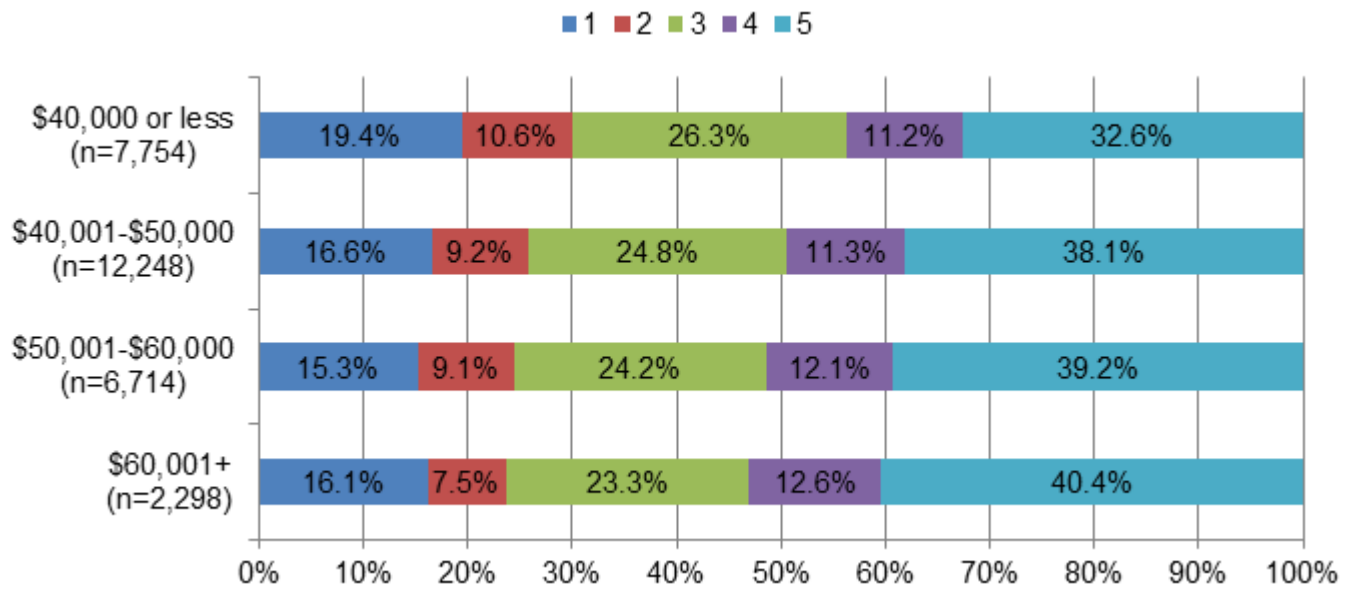
Exhibit 10: Individual Growth Score by Age, 2011-12



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

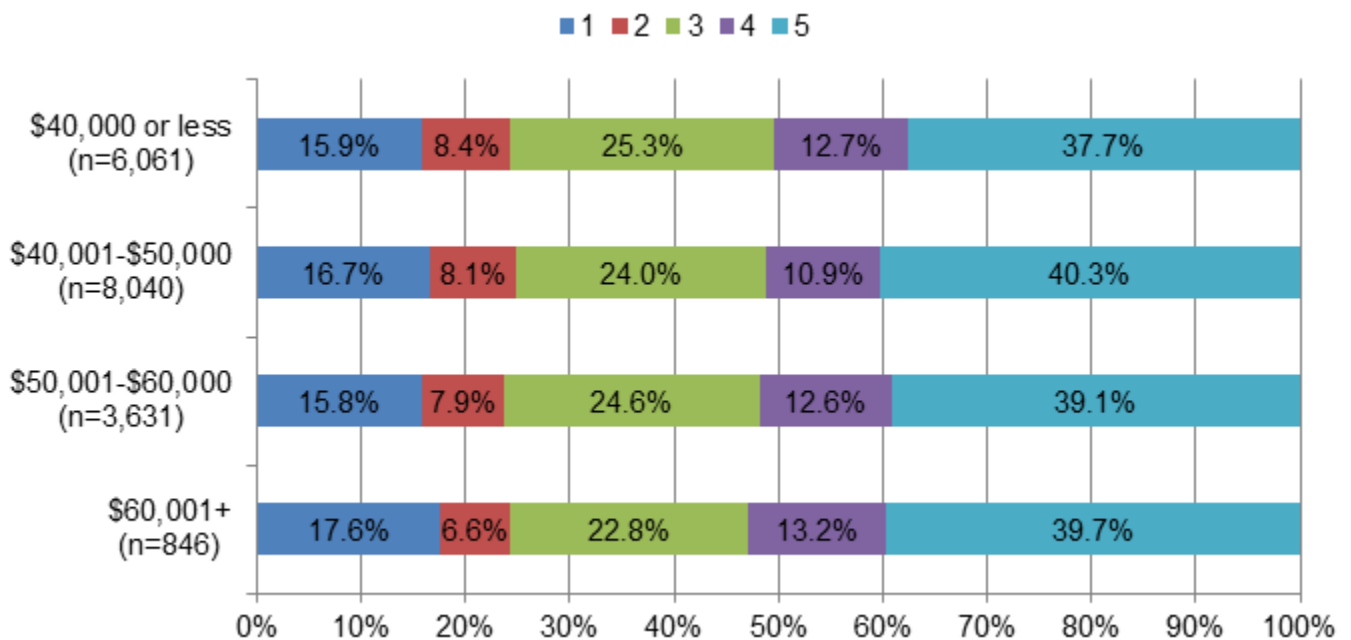
Teachers were grouped into four age categories: 30 years or less; 31-40 years; 41-50 years; and 51+ years. In 2013, approximately 37 percent of teachers across each age group received a level 5 for their individual growth score; correspondingly, approximately 17 percent of teachers received a level 1 across each age group. The distribution is similar for teachers' age and individual growth score in 2012.

Exhibit 11: Individual Growth Score by Salary, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Exhibit 12: Individual Growth Score by Salary, 2011-12**

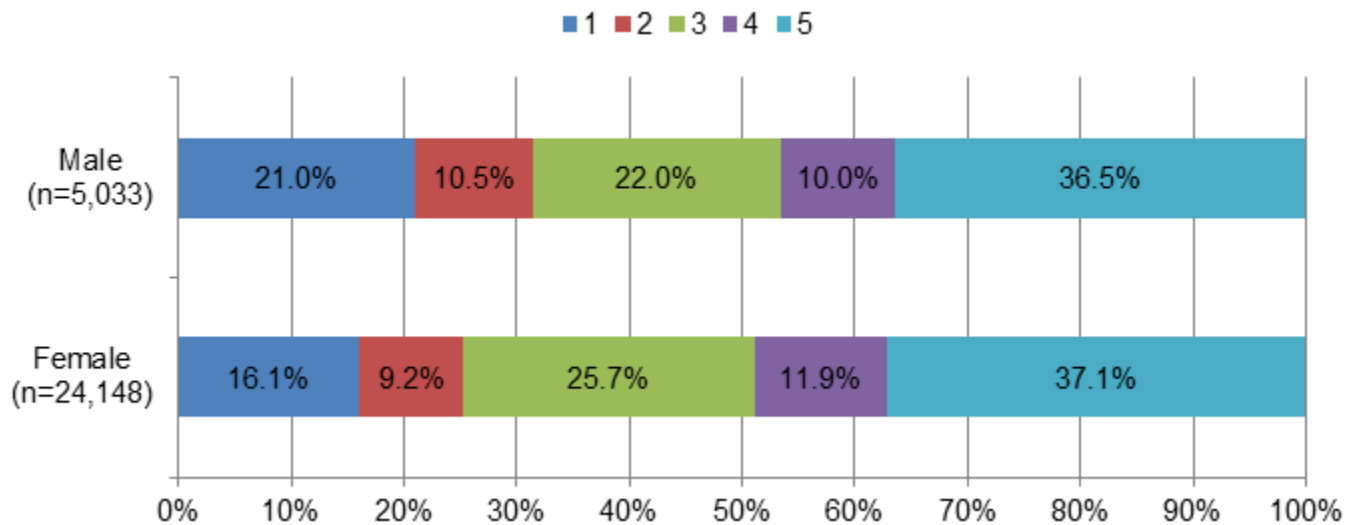


** Data on the salary for teachers under the TEM evaluation model was missing for 2012. This makes up approximately 8.3% of teachers with individual growth scores.

Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Teachers' salaries were grouped into four categories: \$40,000 or less; \$40,001-\$50,000; \$50,001-\$60,000; and +\$60,001. In 2012, approximately 39 percent of all teachers with salary data received a level 5 for their individual growth score. This corresponds with approximately 16 percent of teachers receiving a level 1. Similarly, in 2013, findings illustrate that roughly 37 percent of all teachers with salary data received a level 5, while approximately 17 percent of teachers received a level 1 for their individual growth score.

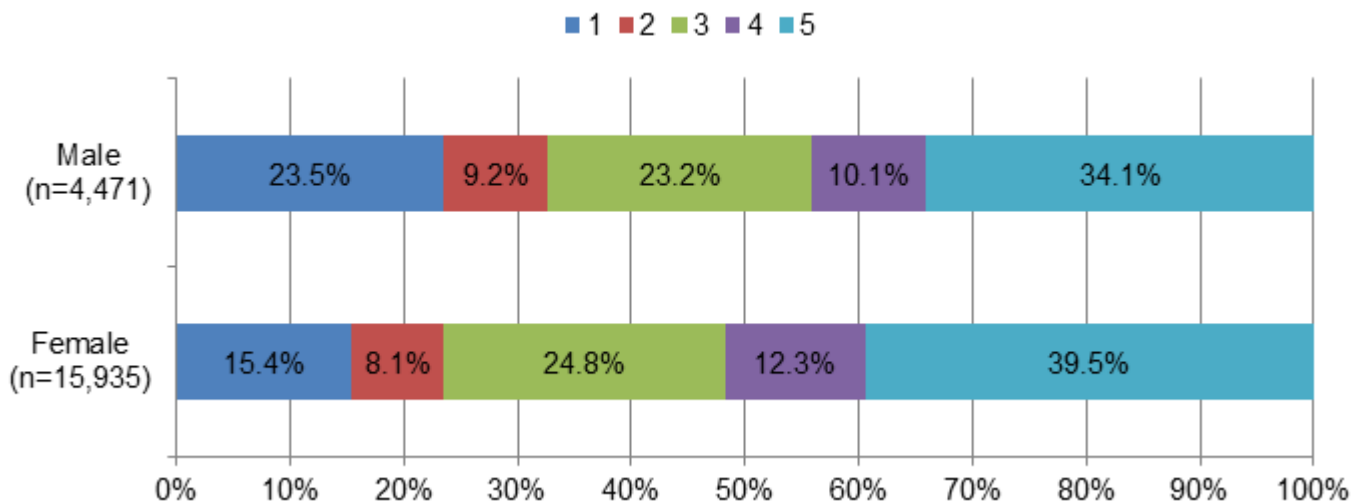
Exhibit 13: Individual Growth Score by Gender, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

As noted in Exhibit 13, approximately 37 percent of the male and female teaching populations each received level 5 for their individual growth scores. Correspondingly, 21 percent of the male teaching population received a level 1, while 16 percent of the female teaching population received a level 1. The distribution of individual growth scores by gender is similar across the 2012 and 2013 school years. In 2013, nearly 83 percent of teachers with individual growth scores identified as female; 17 percent of teachers identified as male.

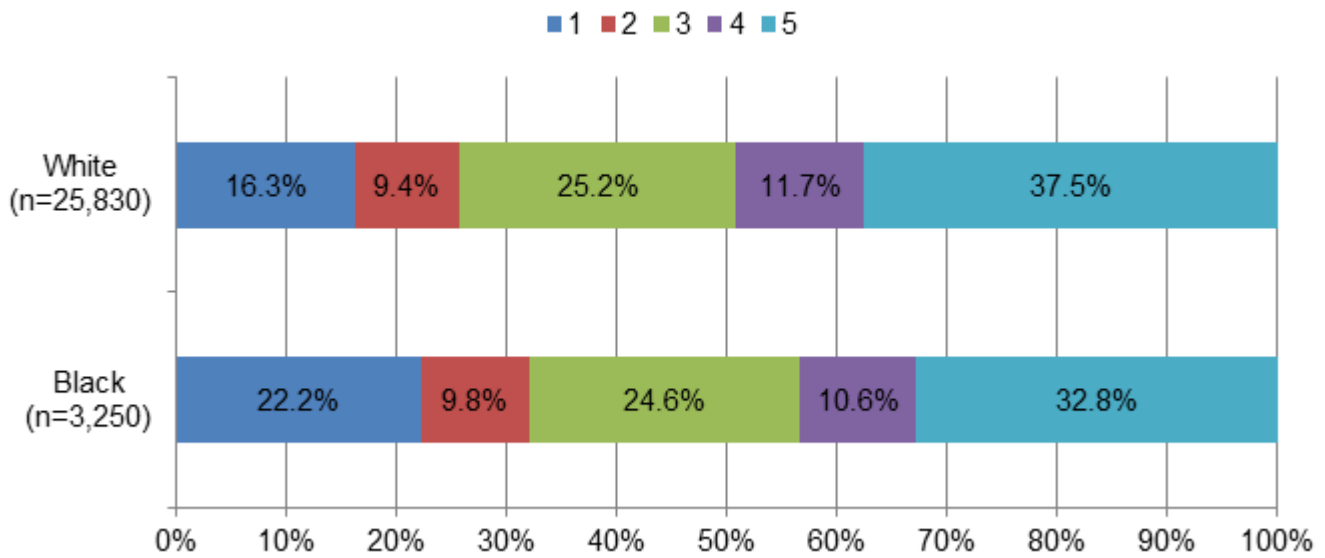
Exhibit 14: Individual Growth Score by Gender, 2011-12



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

In 2012, approximately 78 percent of teachers identified as female, while 22 percent of teachers with individual growth scores identified as male. In 2012, approximately 34 percent of male teachers received a level 5 and 40 percent of female teachers received a level 5 for their individual growth scores. Roughly 15 percent of female teachers received a level 1 and nearly 24 percent of the male teaching population received a level 1. The large population differences are important to consider alongside the distribution of individual growth scores.

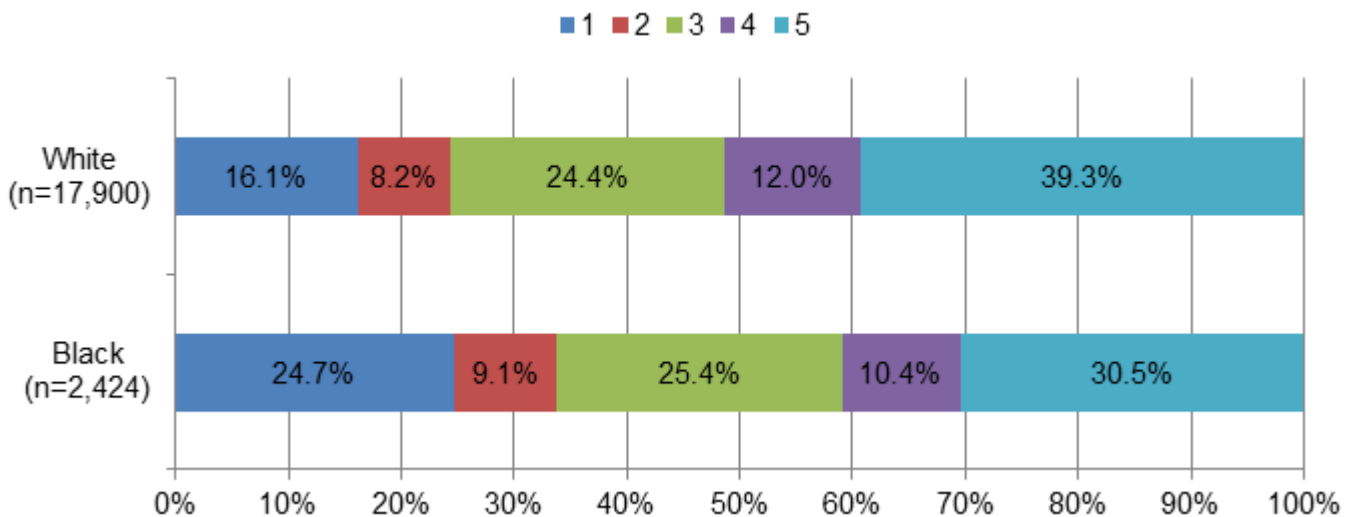
Exhibit 15: Individual Growth Score by Race, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

In 2013, of teachers with individual growth scores, approximately 89 percent were White, 11 percent identified as Black (African-American), and 0.4 percent identified as Asian, Hispanic, American Indian/Alaskan Native, or Native Hawaiian or Other Pacific Islander. Due to the extreme differences in population size, OREA limited its analysis to the White and Black teaching populations.[Ⓓ] Exhibit 15 illustrates that nearly 38 percent of white teachers received a level 5 and 33 percent of Black teachers received a level 5 for their individual growth score. Approximately 22 percent of Black teachers received a level 1 and 16 percent of White teachers received a level 1 for their individual growth score.

Exhibit 16: Individual Growth Score by Race, 2011-12



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

[Ⓓ] The limiting of the analysis to White and Black teaching populations is for both the 2011-12 and 2012-13 academic years and across both the individual growth score/qualitative composite analyses.

In 2012, 17,900 (88 percent) teachers with individual growth scores identified as White, while 2,424 (12 percent) identified as Black (African-American), and 84 (0.41 percent) teachers identified as Asian, Hispanic, American Indian/Alaskan Native, or Native Hawaiian or Other Pacific Islander. Similar to 2013, the “other” category was excluded from the analysis. [Exhibit 16](#) illustrates that approximately 39 percent of the White teaching population and 31 percent of the Black teaching population received a level 5. Roughly 16 percent of the White teaching population received a level 1, while 25 percent of the Black teaching population received a level 1.

Qualitative Measures^E

The qualitative component of teacher evaluation is largely composed of observation ratings. (See [Exhibit 2](#) for a more detailed breakdown of evaluation models and qualitative metrics and weights.)^F The distribution of qualitative scores indicates that approximately 97 percent of the teaching population received marks of ‘At Expectations’ or above. [Exhibit 17](#) illustrates that the majority of teachers received scores at the 3, 4, and 5 levels.

Exhibit 17: Distribution of Qualitative Scores (Overall)

Qualitative Score Distribution (Overall)	% with level 1	% with level 2	% with level 3	% with level 4	% with level 5
2011-12	0.17%	2.3%	21.8%	53.1%	22.8%
2012-13	0.26%	2.9%	22.9%	44.1%	29.9%

Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

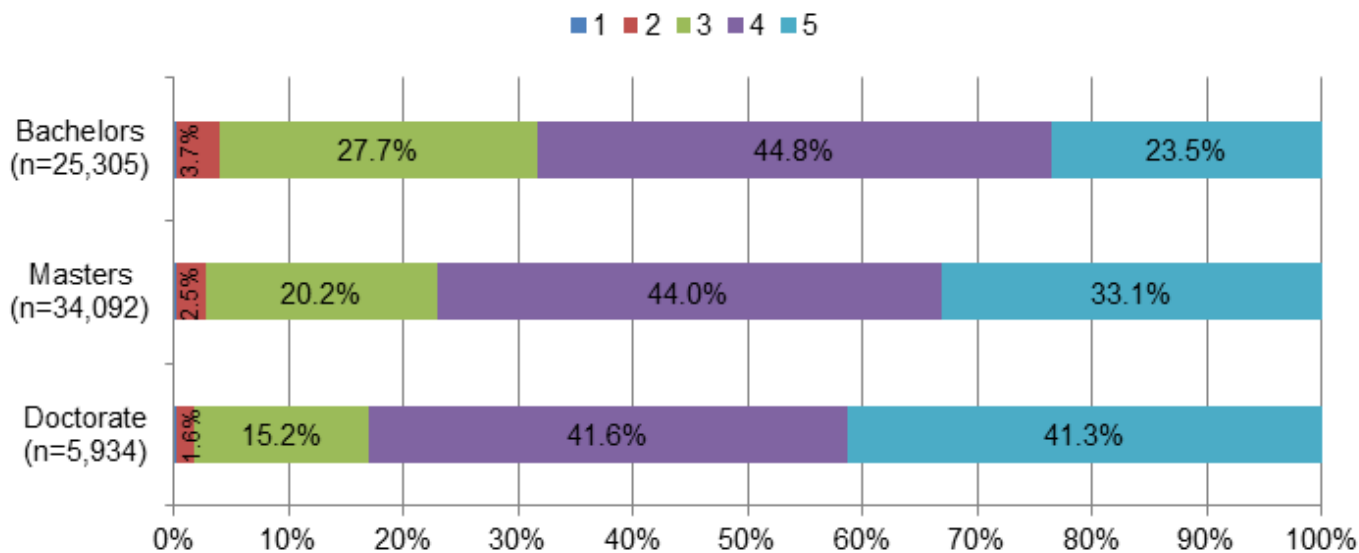
Across the 2011-12 and 2012-13 school years, the rate at which Tennessee teachers were not meeting expectations on the qualitative portion of their evaluation is 2.5 percent and 3.2 percent, respectively. The finding that very few teachers do not meet expectations in their observations (the largest component to the qualitative composite metric) is consistent with findings in other states and districts.¹⁴

The number of teachers receiving a rating of 1 is small and overall accounts for 0.2 percent (n=112) and 0.3 percent (n=173) of teachers in 2011-12 and 2012-13, respectively. For this reason, the analysis does not provide percentages for a 1 rating on each exhibit in the qualitative composite section.

^E Note that throughout the exhibits in this portion of the memo, the percentage of various teacher groupings receiving a rating of 1 is excluded because it is very small.

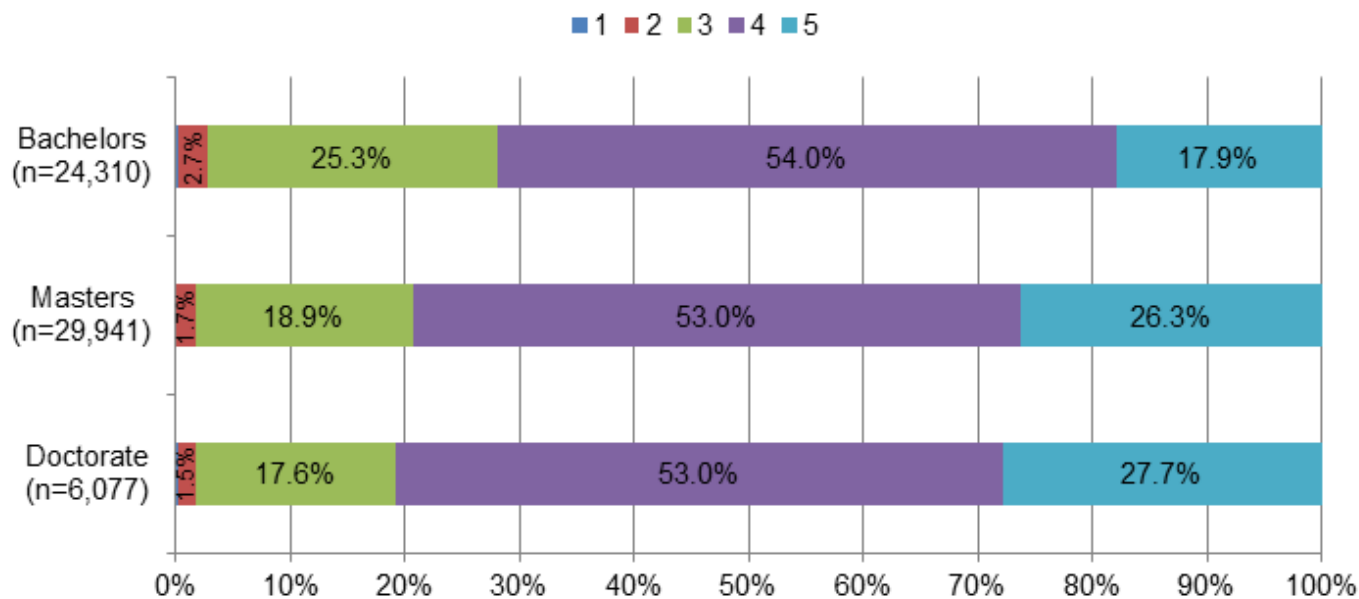
^F Tony Pratt, Deputy Assistant Commissioner of Data and Research, Tennessee Department of Education, e-mail, Oct. 1, 2014. This metric is the composite of qualitative measures. For example, the TEM observation score includes the student survey and knowledge/professionalism score and weights them appropriately. Therefore, this measure is not necessarily just observation scores, but can be attributed to a single teacher and is considered the 50% or 60% (dependent upon individual or school/district level growth score) qualitative metric for teacher evaluation.

Exhibit 18: Qualitative Measures Distributed by Level of Education, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Exhibit 19: Qualitative Measures Distributed by Level of Education, 2011-12**



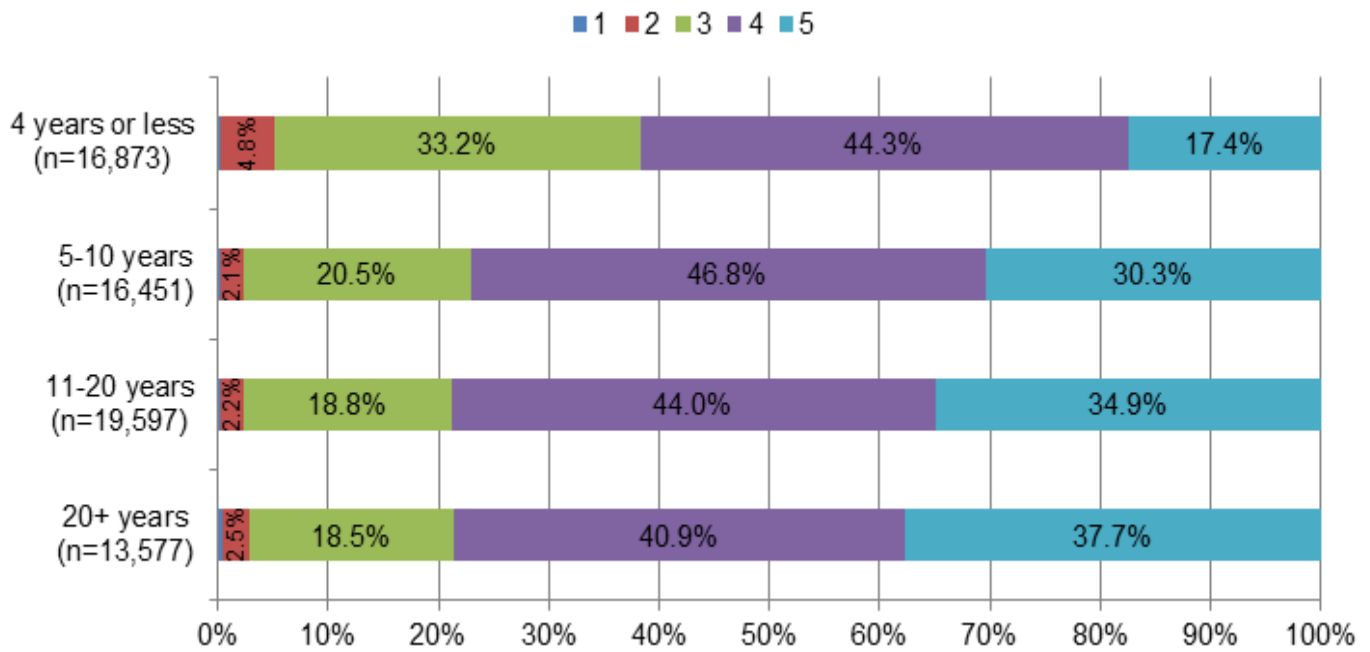
** Data on the level of education for teachers under the TEM evaluation model was missing for 2012. This makes up approximately 9.5% of teachers with qualitative composite scores.

Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Exhibit 18 (2013) illustrates the distribution of qualitative composites by teacher education level. Approximately 24 percent of the teaching population with Bachelor’s degrees received a level 5, while nearly 33 percent of teachers with Masters and 41 percent of teachers with Doctoral degrees received a level 5. A study and its review of the literature found mixed conclusions concerning the effect of a teacher’s level of education on student performance (test scores).¹⁵ However, there is a lack of solid research on the link between a teacher’s level of education and observation outcomes.^G

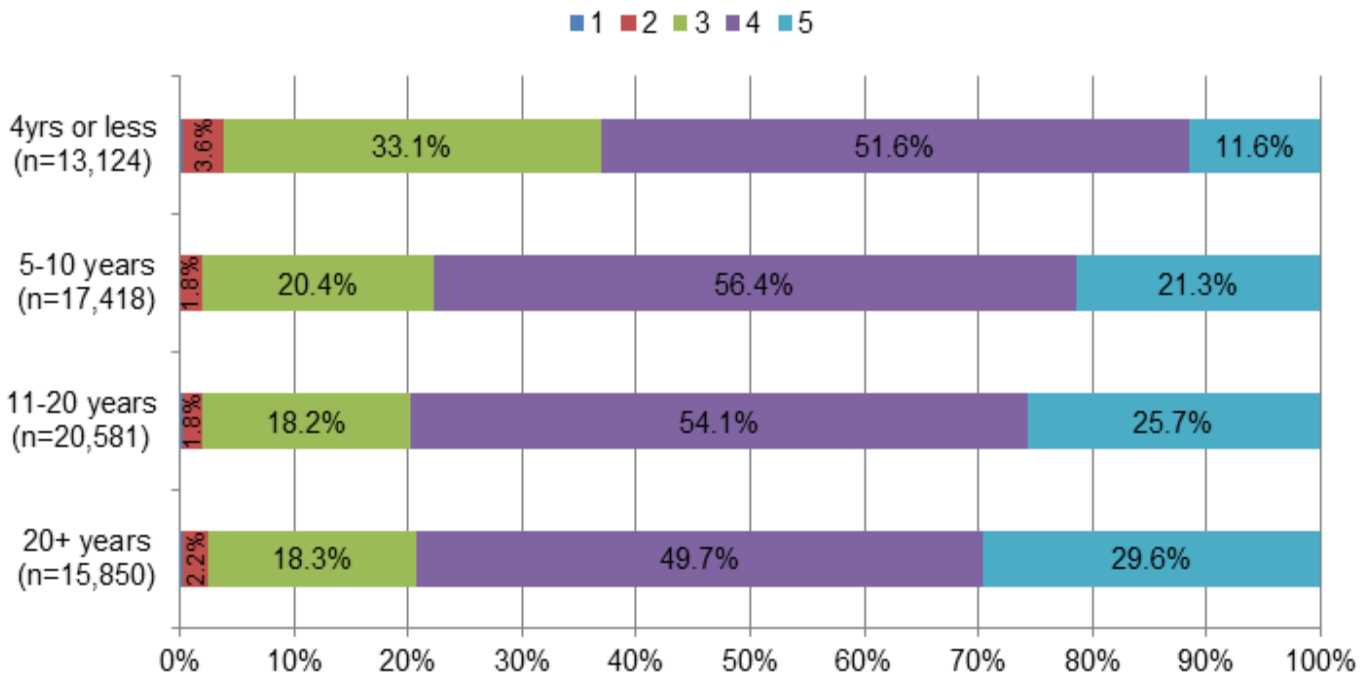
^G An OREA search did not find any relevant articles.

Exhibit 20: Qualitative Measures Distributed by Years of Experience, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Exhibit 21: Qualitative Measures Distributed by Years of Experience, 2011-12

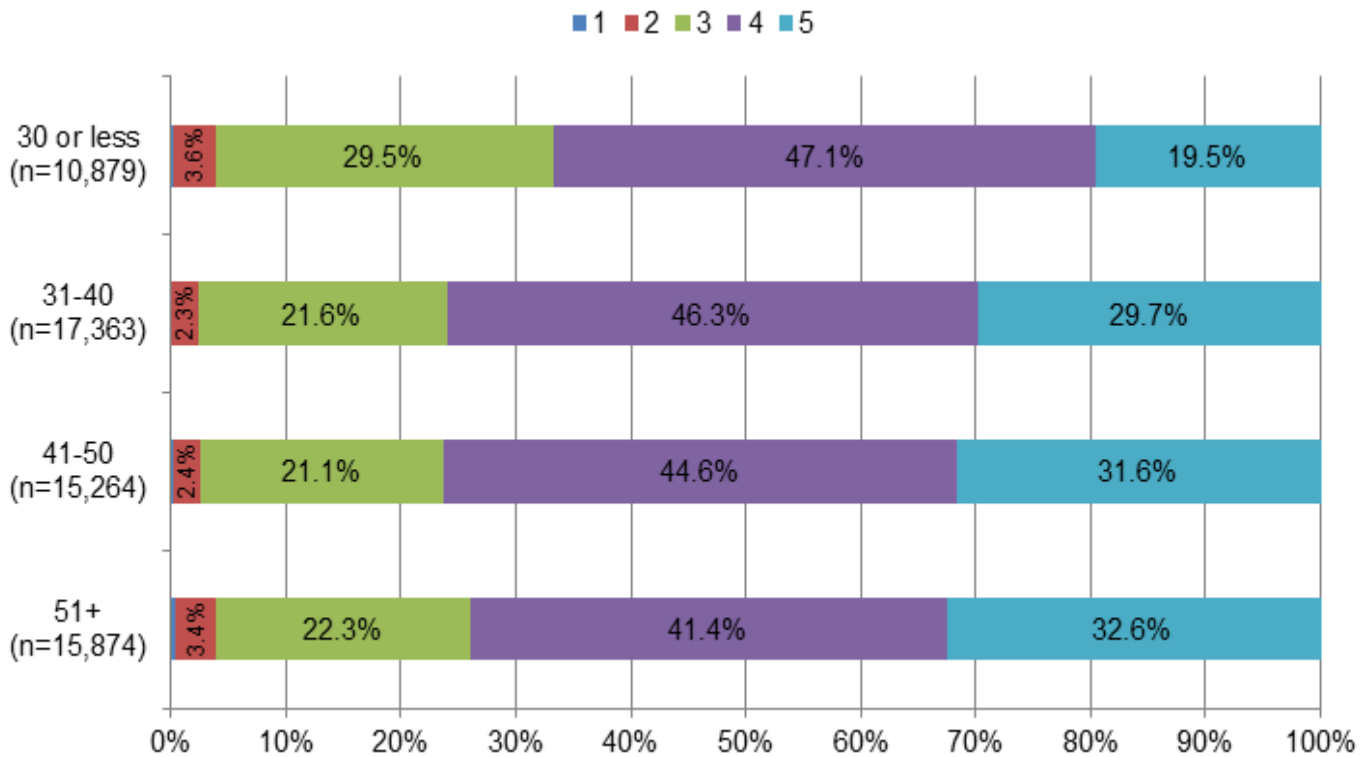


Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

A review of the distribution of the qualitative composite measures by years of experience shows that teachers with four years or less of experience, proportionally, receive fewer level 5s. Approximately 17 percent of teachers with four years or less of experience received a level 5, while 30 percent of teachers with 5-10 years of experience received a level 5. A study of North Carolina’s evaluation system found that beginning teachers with the same teacher effect score as their more experienced

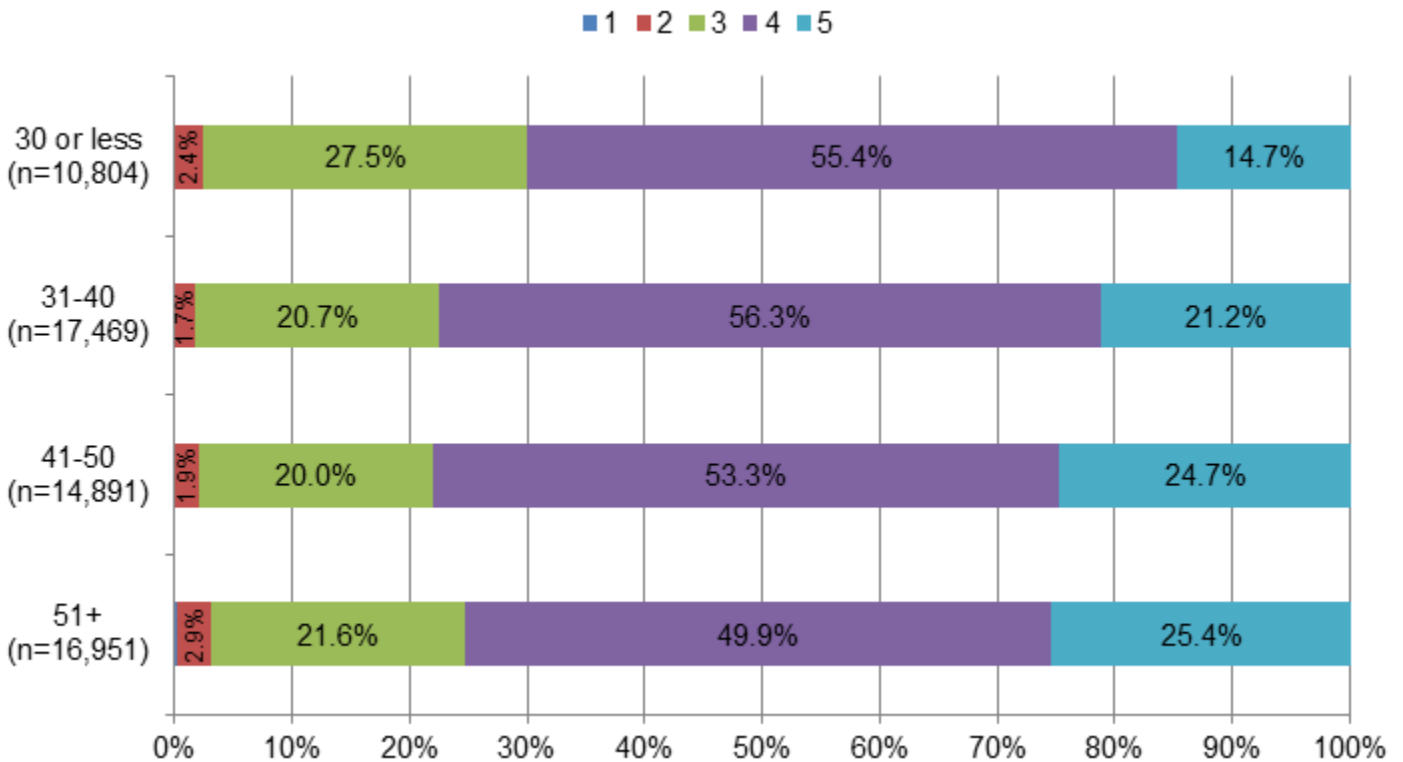
counterparts still tended to receive lower evaluation ratings.¹⁶ A study reviewing teacher effectiveness and principal perception of teacher effectiveness found mixed results regarding teachers' years of experience and observation rating.¹⁷

Exhibit 22: Qualitative Composite Measure Distributed by Age, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

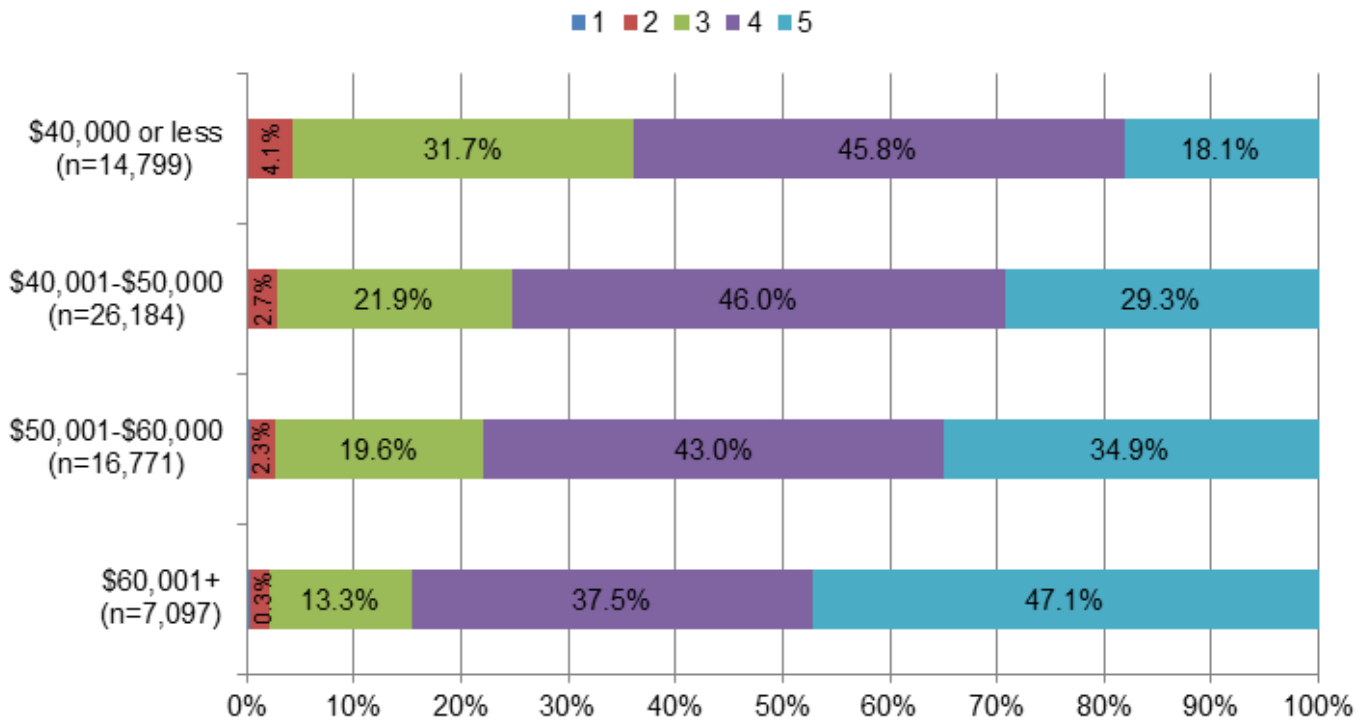
Exhibit 23: Qualitative Composite Measure Distributed by Age, 2011-12



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Exhibits 22 and 23 illustrate that younger teachers receive, proportionally, fewer level 5s than their older counterparts. For example, in 2013, approximately 20 percent of teachers in the 30 years or less category received a level 5, while nearly 30 percent of teachers in the 31-40 years category received a level 5. The percent of teachers receiving a level 5 is similar among the 31-40, 41-50, and 50+ years' categories. The percent of teachers in each age group that receives levels of 1 and 2 is very low.

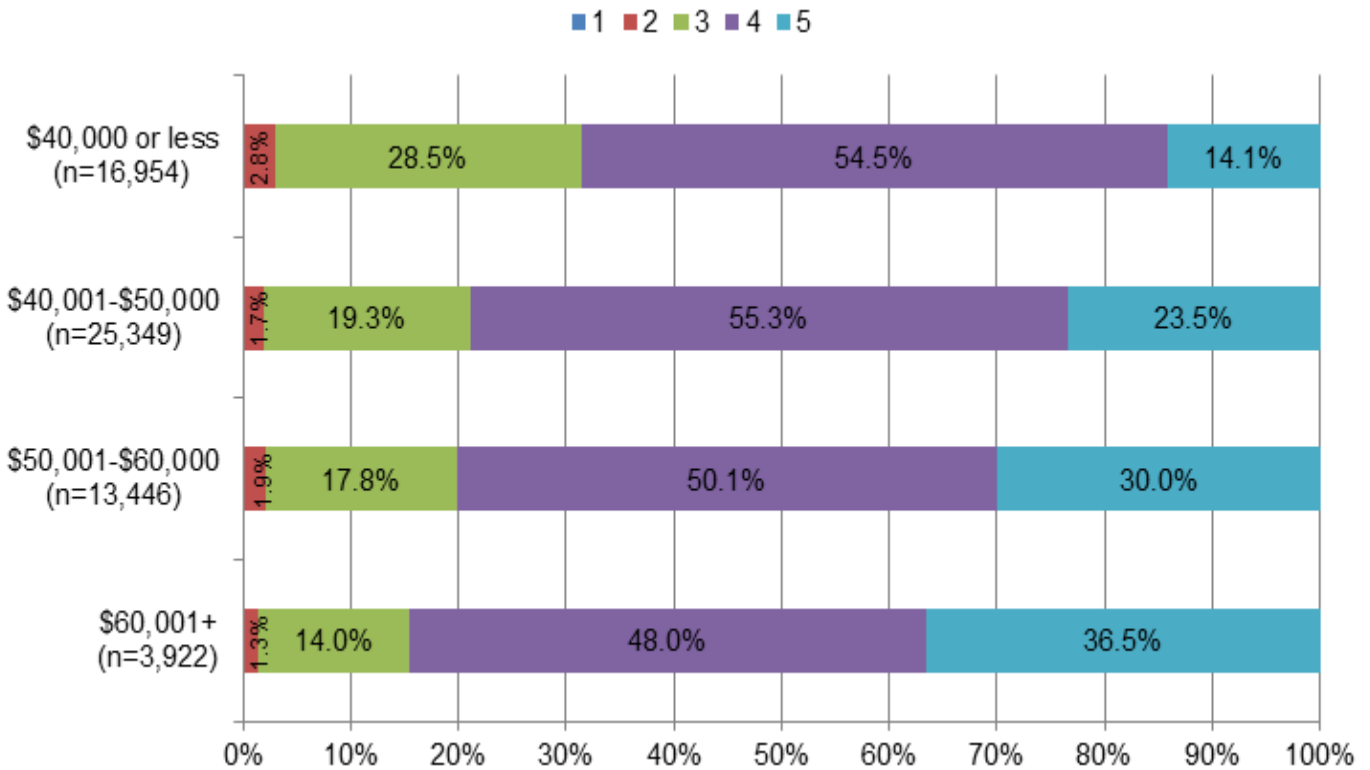
Exhibit 24: Qualitative Composite Measure Distributed by Salary Range, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

In 2013, approximately 47 percent of teachers making \$60,001+ received a level 5 on their qualitative composite, while 18 percent of teachers in the \$40,000 or less category received a level 5. Exhibit 25 shows that, in 2012, approximately 37 percent of teachers in the \$60,001+ category and 14 percent of teachers in the \$40,000 or less category received a level 5. As previously mentioned, salary is correlated with years of experience and age, and this is evident in the similar distribution of the qualitative composite metrics across these three variables. Very few teachers across each salary range category receive ratings of 1 or 2 for their qualitative composite score.

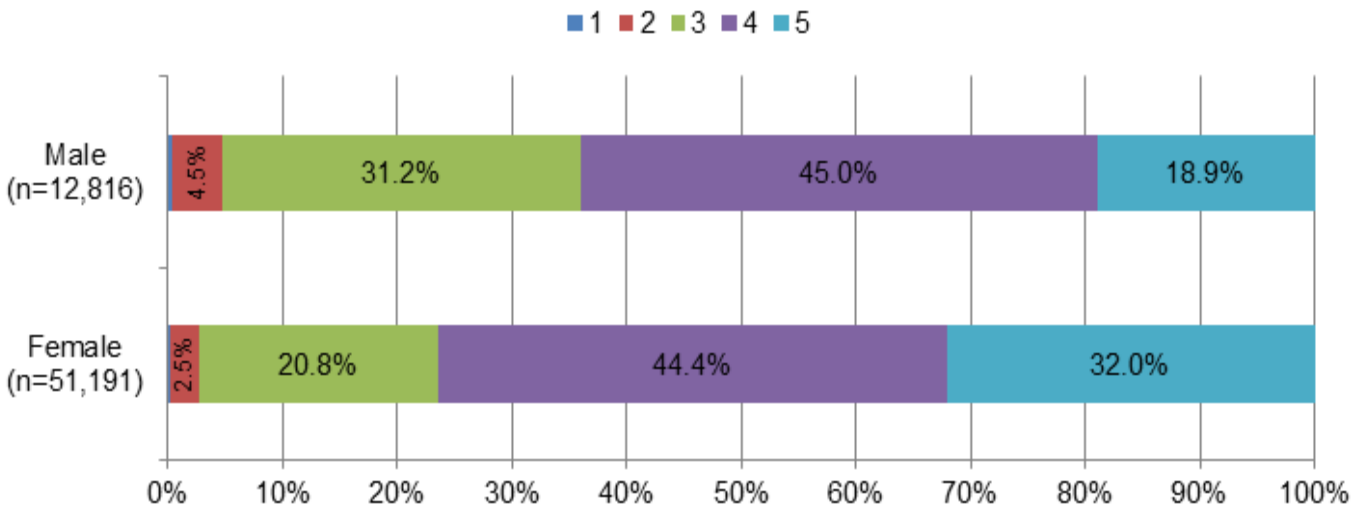
Exhibit 25: Qualitative Composite Measure Distributed by Salary Range, 2011-12**



** Data on the salary for teachers under the TEM evaluation model was missing for 2012. This makes up approximately 9.5% of teachers with qualitative composite scores.

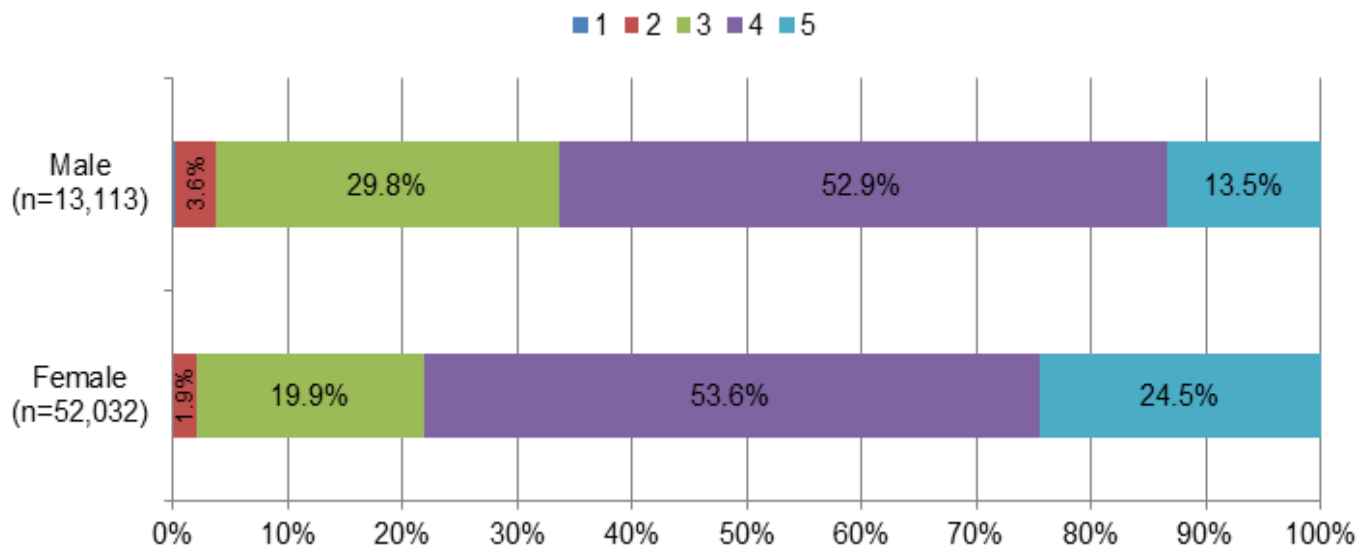
Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Exhibit 26: Qualitative Measures Distributed by Gender, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Exhibit 27: Qualitative Measures Distributed by Gender, 2011-12



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

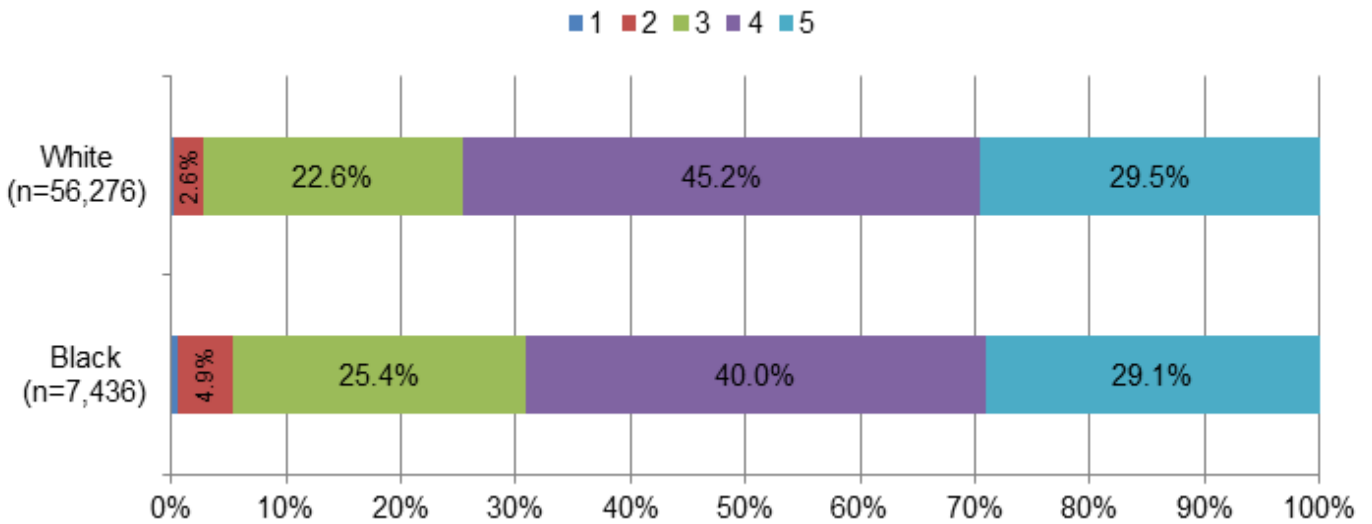
The distribution of qualitative composites by gender illustrates that male teachers, proportionally, receive fewer level 5s than their female counterparts. Both genders receive levels of 1^H and 2 at a similar, low rate. A study of the North Carolina teacher evaluation system found female teachers tended to receive higher performance evaluation (observation) scores than their male counterparts, holding constant many factors, including teacher growth (effect) scores.^{18, I}

The distribution of teachers' qualitative composite scores by race shows very little variation. In 2013, approximately 29 percent of teachers from each group received a level 5. Other score levels are also closely aligned. (See Exhibits 28 and 29.)

^H The percent of teachers receiving a rating of 1 is so small that it is not visible on the graph.

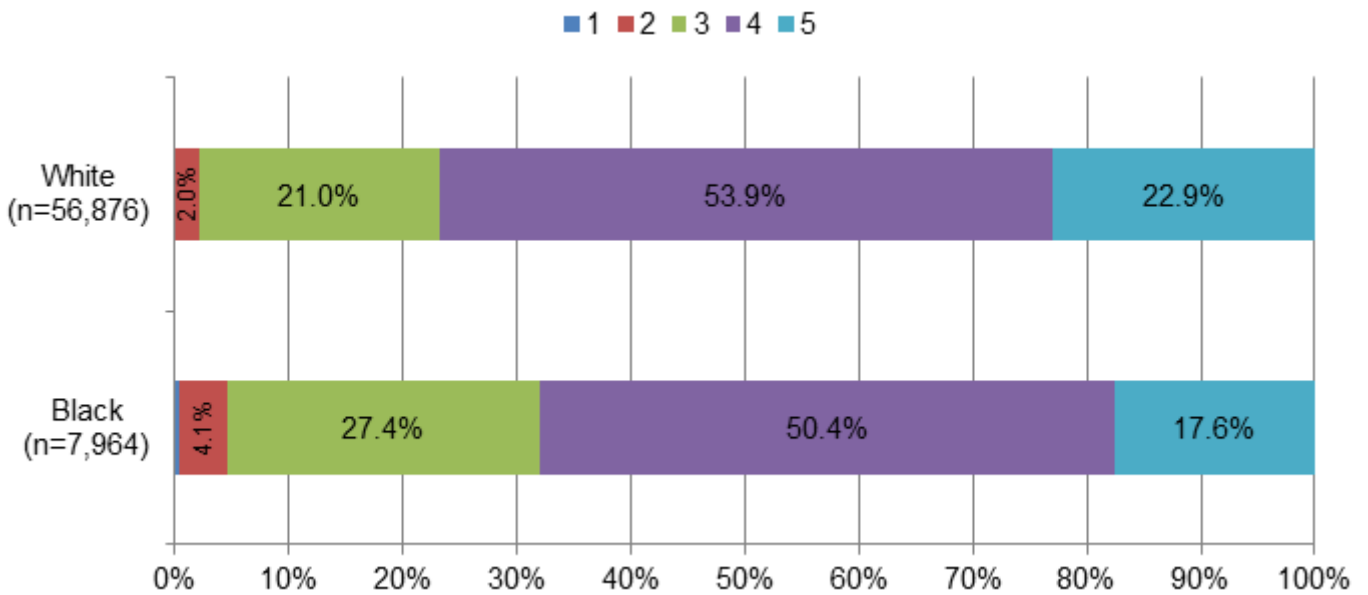
^I The North Carolina study suggests factors that might explain this finding, including: some aspects of teacher performance are not captured in their [the authors] research design model, there is racial or gender bias, or that 'mental short cuts,' like the idea that newer teachers are generally less effective, might impact an administrator's performance evaluation of teachers. See Dayne Batten, Christopher Britt, Jennifer DeNeal, and Lauren Hales, *North Carolina Teacher Evaluations and Teacher Effectiveness: Exploring the Relationship between Value-Added Data and Teacher Evaluations*, Public Schools of North Carolina, State Board of Education, Department of Public Instruction, 2012, p. 7, <http://www.ncpublicschools.org/> (accessed Sept. 8, 2014).

Exhibit 28: Qualitative Measures Distributed by Race, 2012-13



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Exhibit 29: Qualitative Measures Distributed by Race, 2011-12



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

Endnotes

- ¹ Tennessee Department of Education, *TEAMTN website*, <http://team-tn.org/> (accessed Dec. 17, 2014). Tennessee State Board of Education, *Teacher and Principal Evaluation Policy 5.201*, effective July 25, 2014, <http://www.tn.gov/sbe/> (accessed Dec. 17, 2014).
- ² Tennessee Department of Education, "Overview," <http://team-tn.org/> (accessed Dec. 17, 2014).
- ³ Tennessee Department of Education, "Student perception surveys," <http://team-tn.org/> (accessed Dec. 17, 2014).
- ⁴ Tennessee State Board of Education, Chapter 0520-1-1 (a), Local Evaluation of Teachers, Principals and Non-Instructional, Certified Staff, effective Nov. 4, 2011. Previously, in the 2011-12 academic year all teachers had 35% for their growth score metric and 50% overall for the qualitative composite; legislation changed the weights for teachers without individual growth scores for 2012-13.
- ⁵ Tennessee State Board of Education, *Teacher and Principal Evaluation Policy 5.201*, effective July 25, 2014, <http://www.tn.gov/sbe/> (accessed Dec. 17, 2014).
- ⁶ Tennessee Department of Education, *Teacher Evaluation in Tennessee: A Report on Year 2 Implementation*, p. 15, <http://tn.gov/> (accessed Dec. 17, 2014).
- ⁷ Tennessee Department of Education, *Local Evaluation Decisions*, <http://team-tn.org/> (accessed Dec. 1, 2014).
- ⁸ Tennessee Department of Education, *Teacher Evaluation in Tennessee: A Report on Year 2 Implementation*, p. 17, <http://tn.gov/> (accessed Dec. 17, 2014).
- ⁹ Tennessee Department of Education, "Student perception surveys," <http://team-tn.org/> (accessed Dec. 17, 2014).
- ¹⁰ Tennessee State Board of Education, Final Reading Item: II. B., "Teacher and Principal Evaluation Policy 5.201," Aug. 16, 2013. See <http://state.tn.us/sbe/>.
- ¹¹ Tennessee Department of Education, "Teacher Characteristics and Effectiveness," Feb. 2013, <http://tn.gov/> (accessed Dec. 17, 2014).
- ¹² Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.
- ¹³ Tony Pratt, Deputy Assistant Commissioner, Data and Research, Tennessee Department of Education, interview, Sept. 29, 2014.
- ¹⁴ MET Project, *Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains*, The Bill and Melinda Gates Foundation, Jan. 2012, p. 3, <http://files.eric.ed.gov/> (accessed Dec. 18, 2014); Jenny Anderson, "Curious grades for teachers: Nearly all pass," *The New York Times*, March 30, 2013, <http://www.nytimes.com/> (accessed Dec. 18, 2014).
- ¹⁵ Douglas N. Harris and Tim R. Sass, *Teacher Training, Teacher Quality, and Student Achievement*, 2007, National Center for Analysis of Longitudinal Data in Education Research, <http://files.eric.ed.gov/> (accessed Dec. 18, 2014).
- ¹⁶ Dayne Batten, Christopher Britt, Jennifer DeNeal, and Lauren Hales, *North Carolina Teacher Evaluations and Teacher Effectiveness: Exploring the Relationship between Value-Added Data and Teacher Evaluations*, Public Schools of North Carolina, State Board of Education, Department of Public Instruction, 2012, p. 8, <http://www.ncpublicschools.org/> (accessed Sept. 8, 2014).
- ¹⁷ Douglas N. Harris, William K. Ingle, and Stacey A. Rutledge, "How Teacher Evaluation Methods Matter for Accountability: A Comparative Analysis of Teacher Effectiveness Ratings by Principals and Teacher Value-Added Measures," *American Educational Research Journal*, Vol. 51, Issue 1, Jan. 2014, pp. 98-99, <http://www.ncpublicschools.org/> (accessed Sept. 8, 2014).
- ¹⁸ Dayne Batten, Christopher Britt, Jennifer DeNeal, and Lauren Hales, *North Carolina Teacher Evaluations and Teacher Effectiveness: Exploring the Relationship between Value-Added Data and Teacher Evaluations*, Public Schools of North Carolina, State Board of Education, Department of Public Instruction, 2012, p. 8, <http://www.ncpublicschools.org/> (accessed Sept. 8, 2014).

Appendix 1: Legislative Request

This brief responds to the second bullet point in the legislative request. The adjoining report responds to the first bullet point in the legislative request.

HARRY BROOKS
STATE REPRESENTATIVE
19TH LEGISLATIVE DISTRICT

117 WAR MEMORIAL BUILDING
NASHVILLE, TENNESSEE 37243-0119
(615) 741-6879
E-MAIL:
rep.harry.brooks@capitol.tn.gov

DISTRICT ADDRESS:
6600 WASHINGTON PIKE
KNOXVILLE, TENNESSEE 37918
(865) 687-5987



CHAIRMAN HOUSE EDUCATION COMMITTEE

MEMBER OF:
CALENDAR AND RULES COMMITTEE
CIVIL JUSTICE COMMITTEE
EDUCATION SUB COMMITTEE

House of Representatives State of Tennessee

NASHVILLE

March 17, 2014

Comptroller of the Treasury
State Capitol, First Floor
Nashville, TN 37243-9034

Dear Comptroller Wilson,

We have made significant improvements to the education system in Tennessee in recent years. The significant reforms that we have initiated have been well thought out in addition to being discussed in detail prior to implementation. As part of the ongoing improvement efforts, we continue to discuss the changes to our education system and to review them for effectiveness and efficiency.

I would like to request a review by the Offices of Research and Education Accountability of:

- teacher evaluation systems in other states, specifically a selected sample of other states that employ value-added systems for accountability, and
- Tennessee teacher evaluation data, specifically such variables as teacher age, ethnicity, experience, salary, and evaluation scores.

I believe that such a review will help inform the members of the Education Committee and will contribute to the discussion of accountability issues in the General Assembly.

I am available to discuss this request in more detail at your convenience.

Sincerely,

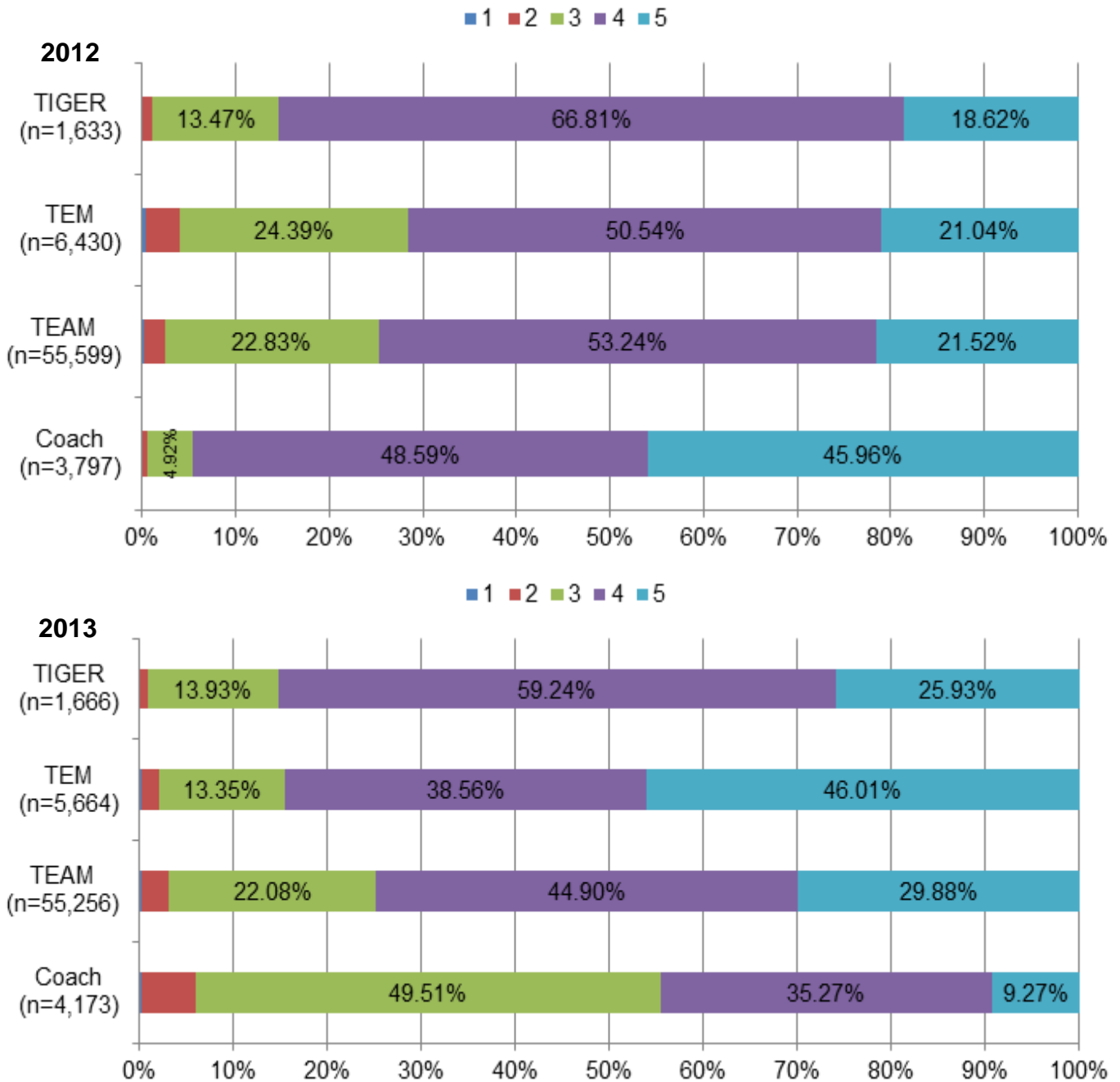
A handwritten signature in cursive script that reads "Harry Brooks".

Harry Brooks
State Representative 19th District

Appendix 2: Evaluation Model Types

Below is a review of observation score distribution by four of Tennessee’s teacher evaluation models: the statewide TEAM model and three alternative evaluation models, including the alternative models used in two of Tennessee’s larger school districts (Hamilton County and Shelby County).^A As stated earlier in the draft, the alternative evaluation models subscribe to different observation procedures and rubrics. (See Exhibit 2.) It is important to note that it is out of the scope of this brief to delve into an analysis of model type and its potential impact on teacher evaluation.

Exhibits 2.1 and 2.2: Qualitative Composite Score Distribution by Teacher Evaluation Model, 2012 and 2013



Source: Tennessee Department of Education, Tennessee Teacher Evaluation Data, 2011-12 and 2012-13, received Aug. 22, 2014.

^A OREA opted to include this analysis for the purpose of illustrating how the different evaluation models fared with regard to qualitative composite scores. The models utilize different observation procedure and rubric, so it was relevant to review the observation score. However, it is out of the scope of this analysis to review the evaluation models against growth scores.

As shown in the two previous charts, the distribution of the qualitative composite of teachers participating in Project COACH changed significantly between the 2011-12 and 2012-13 school years. In 2011-12, approximately 5 percent of teachers received a level 3 for their observation score, while in 2012-13 approximately 50 percent of teachers received a level 3. Additionally, there is a notable variation in the distribution of levels 3, 4, and 5, with very few levels of 1 or 2 across each model within a given year. The differences in the theories of the evaluation models and the feedback and frequency of observations, may impact observation scoring. (See [Exhibit 2.](#))

Appendix 3: Methodology

The data source provided by the Tennessee Department of Education (TDOE) was a compilation of information from several systems and includes data for all individuals evaluated as teachers in the 2011-12 and 2012-13 academic years. Due to a lag on certain metrics and the need to compile information from various sources, the 2013-14 teacher evaluation dataset is unavailable at the time of this analysis.^A

The request sought a review of Tennessee teacher evaluation data and teacher characteristics. The Offices of Research and Education Accountability analyzed the data by comparing teacher evaluation score levels, which range from 1 to 5, relative to teacher characteristic variables (race, gender, level of education, age, years of experience, and salary). TDOE has methods to convert the 1-4 scale scores (Project COACH and TIGER) to 1-5 score levels to ensure that metrics are comparable. Each teacher characteristic variable was reviewed against individual value-added scores and qualitative composite scores, both of which can be directly attributed to an individual teacher.

This review of teacher evaluation data includes variables that are *categorical*, meaning they are not rank ordered (e.g., gender is a categorical variable; a teacher is either male or female); *ordinal*, similar to a categorical variable, but with an order to the variables (e.g., level of education); and *interval*, meaning that the variable is ordered and equally spaced (e.g., age or salary).^B Individual growth scores are a type of ordinal variable; each teacher receives a level ranking based on their individual growth score index, as illustrated in [Exhibit 3.1](#).

Exhibit 3.1: TDOE Index Scores Converted into Levels

Index	<-2	-2 to -1	-1 to 1	1 to 2	>2
Level	1	2	3	4	5

Note: The 1-5 TVAAS levels do not follow exact intervals. Levels 2 and 4 have a 1.0 index range, from -2 to -1 and 1 to 2, respectively, whereas the other levels have a larger index range. For example, level 3 has a 2.0 index range, from -1 to 1.

Source: Tennessee Department of Education, "TVAAS FAQs," <http://team-tn.cloudapp.net/> (accessed Dec. 18, 2014).

Analyses using more advanced statistical techniques, like correlation (an analysis that reflects the variability between the variables in question), were considered. However, teacher evaluation scores and some other teacher characteristics are not interval types of variables, meaning that they are not necessarily rank ordered and equally spaced, and require more rigorous and complex tests of correlation. In addition, the small range of data points, the 1-5 levels for TVAAS and the Qualitative Composite metric, limit the ability to detect variation among characteristics through correlation techniques.^C

^A Tony Pratt, Deputy Assistant Commissioner of Data and Research, Tennessee Department of Education, e-mail, Oct. 1, 2014.

^B Institute for Digital Research and Education UCLA, "What is the difference between categorical, ordinal and interval variables?," <http://www.ats.ucla.edu/> (accessed Dec. 19, 2014).

^C Neil J. Salkind, *Statistics for People who (think they) Hate Statistics*, California: SAGE Publications, 2011, p.80.

A small number of records or partial records were excluded from this analysis. In cases where only certain variables were missing from the data, records were only excluded on those missing variables (e.g., a case is missing salary information, but has all other data; that case would not be included in the analysis of salary, but would be included for the analysis of age). The TEM evaluation model, used in present-day Shelby County Schools, in 2012 was missing both salary and level of education data for all teachers.^D TEM teachers comprised 8.3 percent (n=1,707) of teachers with individual growth scores and 9.5 percent (n=6,430) of teachers with qualitative composite scores in 2012. Across both years, only five analyses had instances with 10 percent or more missing cases:

- Salary (2012) TVAAS and Qualitative Composite
- Age (2012 and 2013) Qualitative Composite Only
- Level of Education (2012) Qualitative Composite Only

Data cleaning was also required to ensure that the information in the dataset was as correct as possible. Changes made include:

- Deleted individual growth score data for 37 cases that had discrepancies in their individual and overall growth data points (i.e., a conflict existed around individual growth score data and was not used in the teachers' evaluation)
- Reconciled issues between years of experience and age by deleting the years of experience variable (41 cases), (e.g., a teacher who had 27 years of experience, but was only 26 years old OR a teacher who was 38 years old and had 20 years of experience)
- Deleted total salary data for teachers whose salaries fell below the minimum allowed for the state (in 2011-12 the minimum salary for a teacher with zero years of experience and a bachelor's degree was \$29,680; in 2012-13 the minimum salary was \$30,420)

In addition, variables were recoded into groups to more easily illustrate findings from the analysis:

- Age was recoded into: Less than 30 years, 31-40 years, 41-50 years, and 51+ years
- Years of experience was recoded into: Less than 4 years, 5-10 years, 11-20 years, and 20+ years
- Salary was recoded into: Less than \$40,000, \$40,001-\$50,000, \$50,001-\$60,000, and \$60,001+
- Race was recoded into: White, Black/African-American, and Other (includes Asian, Hispanic, Pacific Islander, and Alaskan/Native)

^D Tony Pratt, Deputy Assistant Commissioner of Data and Research, Tennessee Department of Education, e-mail, Aug.27, 2014.



OFFICES OF RESEARCH AND EDUCATION ACCOUNTABILITY
Phillip Doss, Director
Suite 1700, James K. Polk Building ▪ 505 Deaderick Street
Nashville, Tennessee 37243 ▪ (615) 401-7866
www.comptroller.tn.gov/orea

