



STATE BOARD of ELEMENTARY and SECONDARY EDUCATION

P.O. Box 94064, Capitol Station, Baton Rouge, LA 70804-9064 · PHONE: 225-342-5840 · FAX: 225-342-5843

March 1, 2013

TO: Senator Conrad Appel, Committee Chairman, and Members of the Senate Committee on Education

Representative Steve Carter, Committee Chairman, and Members of the House Committee on Education

FROM: Charles E. "Chas" Roemer, President 
State Board of Elementary and Secondary Education

Re: Report on the Implementation, Results, and Effectiveness of the Value-Added Assessment Model

Beginning this year, R.S. 17:3883(A)(8) requires the Board of Elementary and Secondary Education (BESE) to annually submit to the Senate Committee on Education and the House Committee on Education a written report, on or before March 1, regarding the implementation, results, and effectiveness of the value-added assessment model.

On behalf of Board, I am pleased to present the attached report that provides an update on the first year of statewide Compass implementation, a description of enhancements made by the Louisiana Department of Education (LDE) and BESE to the Compass system based on educator feedback, and detailed information regarding the value-added model's methodology and results.

CR:HC:naf

Attachment (1)

c: David R. Poynter Legislative Research Library
Jeanne Johnston, Senior Analyst, Senate Committee on Education
Nancy Jolly, Senior Legislative Analyst, House Committee on Education

James Garvey
1st BESE District

Kira Orange Jones
2nd BESE District

Lottie P. Beebe
3rd BESE District

Walter Lee
4th BESE District

Jay Guillot
5th BESE District

Chas Roemer
6th BESE District

Holly Boffy
7th BESE District

Carolyn N. Hill
8th BESE District

Connie Bradford
Member-at-Large

Judith Miranti
Member-at-Large

Stephen Waguespack
Member-at-Large

Heather Cope
Executive Director

John C. White
State Superintendent

**Compass: Report on the Status of Louisiana's Educator Support and Evaluation System
and the Value-Added Assessment Model**

**A Report to the
Senate Education Committee and the House Education Committee
of the Louisiana Legislature**

**From the State Board of Elementary and Secondary Education
*Prepared by the Louisiana Department of Education***

March 1, 2013

Table of Contents

I. Introduction.....	3
II. Processes Supporting Development of Compass and the Value-Added Model.....	4
III. Value-Added Model Technical Process and Findings.....	8
a. Introduction.....	8
b. Combining Student Performance and Other Student Data.....	9
c. Value-Added Analysis.....	10
d. Standards of Effectiveness	11
e. Selected Results	12
<i>Stability of Teacher Results Across Years in Mathematics and English Language Arts.</i>	12
<i>Estimated Average Levels of Achievement.....</i>	12
<i>Preliminary Findings from the 2012-13 Implementation of Compass</i>	13
IV. Conclusion.....	13
V. Appendix	15
<i>Distribution of Student-Teacher Achievement Outcomes for 2011-2012</i>	15

I. Introduction

Compass is Louisiana's educator support and evaluation system. Compass is comprised of tools and supports designed to give teachers meaningful feedback on their classroom practice to help all students achieve rigorous academic goals. The figure below shows how Compass's multiple measures come together to form the basis for teacher and school leader evaluations.

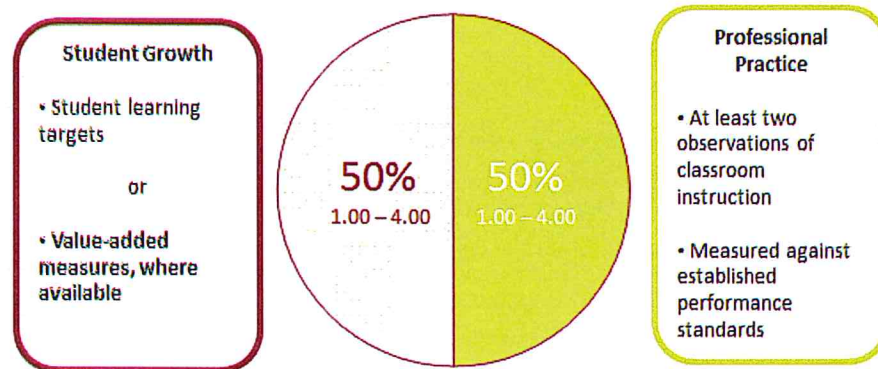


Figure 1. Compass Measures

The value-added assessment model is one of the Compass system's tools. The value-added model generates valuable information about a teacher's impact on student learning. During the 2011-12 school year, the full Compass model was piloted with nine school districts and one charter school, and all districts across the state received value-added data. Compass was implemented statewide at the start of the 2012-13 school year. Preliminary data from this year show that 87 percent of teachers have been rated *Effective: Proficient or Highly Effective*, based upon observations, while only 51 percent of teachers were rated in these categories according to value-added in the previous year. This suggests that there is more work to be done to ensure that teachers across the state are getting the feedback they need to drive gains in student learning.

This report provides an update on this first year of statewide Compass implementation, enhancements the LDE has made to the Compass system based on educator feedback, and information regarding the value-added model's methodology and results. In terms of the value-added model's results, key findings are:

- The value-added model clearly differentiates between teachers who are generally meeting expectations, significantly exceeding expectations, and significantly falling short of expectations;
- The value-added model's results continue to be stable across multiple years; and
- Despite the distinctions between educators' performance surfaced by the value-added model, preliminary data show that observations by local evaluators, (a more traditional form of teacher evaluation,) place most teachers in the top levels of effectiveness.

II. Processes Supporting the Development of Compass and the Value-Added Model

Prior to the launch of Compass statewide, there have been several key points of engagement with stakeholders, which have informed the system's ongoing development. The first was the formation of the Advisory Committee on Educator Evaluation (ACEE). ACEE's composition included diverse representation from across the state, with the majority of the members being practicing teachers. Other committee members included parents, legislators, school board members, Board of Elementary and Secondary Education (BESE) representatives, union representatives, and other educator association representatives. The committee convened its first meeting in September 2010. ACEE members were charged to make recommendations to BESE regarding the value-added model, evaluations for teachers of non-tested grades and subjects, and overall standards of effectiveness for educators. Recommendations regarding these topics were presented to and accepted by BESE in December 2011.

Another key stakeholder engagement point was the Compass pilot. The value-added assessment model was piloted over the course of a three-year period, beginning with the 2008-09 school year. This included the development, testing, and deployment of a secure web portal, through which teachers and educational leaders were able to verify the accuracy of class rosters prior to their use in the value-added analysis and through which they accessed their value-added reports. The portal was developed to address concerns that data quality would be a barrier to accurately estimating teacher contributions to student progress and to create as much transparency as practical into the process of deriving value-added scores. The portal gives teachers the opportunity to know exactly which students are contributing to their results and correct roster errors. It also allows teachers, principals, district superintendents, and district data personnel access to the value-added results. The portal was tested with a small subset of pilot schools and districts in 2008-2009 and 2009-2010. Statewide testing took place during 2010-2011 and the most recent implementation year, 2011-2012. In 2010-2011, 19 volunteer school districts and two charter schools (a total of 328 schools), engaged in a more robust value-added pilot, including field testing of the educator professional development, materials, and training related to the model. In 2011-2012, all schools and districts who requested training were provided guidance. From individual schools to district-wide meetings and educator support groups, over four dozen informational sessions were held, both in-person and virtually. Beginning with the 2013-2014 school year, the value-added portal will be integrated with a web-based data system that support teachers, schools, and districts with managing the data available from all Compass measures, including observations and student learning targets.

The full Compass system was piloted in the 2011-12 school year in nine districts and one charter school. Throughout these pilots, the Department of Education received valuable feedback from practitioners regarding how Compass could be refined to support educators in advancing

student learning. Several significant adjustments were made to the Compass model as a result of this feedback, including the adoption of a modified version of Charlotte Danielson's Framework for Effective Teaching as the model observation tool for teachers and several key enhancements to the Compass data system.

The LDE has continued to collect stakeholder input throughout this first year of statewide implementation. During the first semester of Compass implementation, educators provided additional feedback to the LDE through in-person and virtual town halls, conducted throughout the fall of 2012. Based on this feedback, in January 2013, the LDE proposed additional refinements to Compass, which were approved by BESE and described in three live webinars for educators throughout the course of that month. The key points of feedback gathered during this stage included:

- The original rubric, with eleven performance standards, was too lengthy and cumbersome;
- Teachers need more information on student performance to inform goal-setting;
- Teachers and leaders would benefit from more concrete examples of what instruction looks like at different levels of performance; and
- Reporting clear, complete data and individualized feedback back to educators at the end of the year is critical to understanding performance and identifying ways to continuously improve.

In response to this feedback, the LDE proposed several important refinements to Compass, aimed at increasing clarity and high-quality feedback for teachers, to make the system more effective in guiding educators' professional development. First, we looked to national models to identify strong rubrics that had been researched, refined, and proven elsewhere. The Charlotte Danielson framework provided us with these benefits, as well as the flexibility to narrow our focus on a smaller set of high-impact teacher actions that were clear and evidence-based. The rubric, therefore, was revised to include a set of five performance indicators, and districts were also given the opportunity to propose their own rubrics to meet local needs. Second, the LDE designed new reports for teachers receiving value-added data that show them the estimated expected scores of their students on the relevant state assessment, their prior year's score, and the student-specific factors that will be taken into account in the value-added analysis, early in the school year. These reports will be available for the first time in the fall of 2013. The LDE will also add a new report to the value-added results teachers receive at the end of the 2012-2013 school year, which will provide comparisons of each individual student's expected and actual scores, along with the student-specific factors that were included in the analysis. These reports will allow teachers to analyze their value-added results down to the student level and will provide them with helpful background data on students to inform their goal-setting at the beginning of the year. Figures 2 and 3 provide preliminary mock-ups of these reports.

Student Last Name	Student First Name	Estimated Expected Score	Free/Reduced Price Meal Status	Gifted Status
Baker	Justin	428	None	No
Clinton	Mindy	428	Free	No
Gray	Lacey	485	Reduced	Yes
Jones	Travon	375	None	Yes
Gautreaux	Melissa	428	Free	Yes
Morris	Barbara	389	Reduced	No
Smith	Taylor	399	None	Yes

Note: Students' expected scores are calculated based on their prior achievement, attendance, disability/gifted/Section 504/LEP status, free/reduced lunch status, and discipline record. This report is illustrative and, due to space constraints, includes only some of the factors used to calculate the expected score.

Sort by student status to identify specific needs or achievement trends.

Sort by expected score to identify sub-groups that might need targeted goals or supports.

Figure 2. Beginning of Year Value-Added Report*

Student Last Name	Student First Name	Special Educ. Status	Actual Score	Expected Score	Difference
Baker	Justin	Yes	428	415	+13
Clinton	Ashleen	No	428	406	+22
Gray	Mindy	No	500	390	+110
Jones	Pete	No	375	390	-15
Jackson	Selena	No	428	384	+44
Morris	Brett	Yes	389	389	Met target
Smith	Taylor	Yes	399	371	+28
Value-Added Composite Score:					+28.86
Value-Added Percentile Range:					80th-99th

Highly Effective

Sort by various student factors to identify performance trends.

Basis for the composite score is presented as a function of individual students' actual vs. expected scores.

Composite score, percentile range, and performance level appear together, painting a more complete picture of teacher performance.

Note: Final Compass ratings represent a combination of the value-added score and the teacher's professional practice score.

Figure 3. End of Year Value-Added Report*

* For simulation purposes only; not actual student names or data. Teachers may only access student data for their students. This information was previously provided in paper form, but will now be available electronically at the beginning of the school year for teacher planning purposes.

Third, the LDE committed to building a comprehensive video resource library to catalogue and share instructional videos and related resources from a diverse set of classrooms. The video library will include videos from multiple grade levels, ranging from pre-kindergarten to 12th grade, and multiple effectiveness levels, ranging from *Ineffective* to *Highly Effective*. Videos will be catalogued into a searchable library that allows educators to find relevant videos based on these criteria, or based on the standards they address, the content area they address, or the specific component from the observation rubric that they feature. These video resources will provide individual teachers and evaluators with concrete examples of strong instruction and they will serve as helpful professional development tools for groups of teachers and evaluators who are collaborating to improve practice.

The fourth significant refinement that the LDE proposed to Compass as a result of educator feedback in the first year related to how evaluators provided teachers who were generally meeting expectations according to the value-added model with effective feedback regarding their impact on student learning. While teachers who were significantly exceeding or falling short of expectations according to the model got very definitive feedback, in the form of an *Ineffective* or *Highly Effective* rating, teachers in the mid-range of value-added results received feedback that did not adequately distinguish their performance from that of their peers. Because so many teachers are clustered in this *Effective* range, (between the 20th and 80th percentiles,) the value-added model is unable to differentiate feedback on their performance as precisely as it does for teachers significantly above or below the mean. However, it is these teachers who stand to benefit the most from specific, individualized feedback on their performance, as they are the teachers best positioned to move from ‘good’ to ‘great.’

To address this challenge, the LDE proposed that for teachers in the *Effective* range, evaluators consult multiple sources of student learning data, as well as their value-added results, before assigning a growth score to teachers. Upon doing so, evaluators of teachers in this range will assign a final student growth score of *Effective: Emerging* or *Effective: Proficient*. These growth ratings will be combined with ratings from their observations throughout the year and weighted at 50 percent of their final score. This practice will allow evaluators to consult multiple sources of data in order to give teachers who are generally meeting expectations in terms of advancing student learning more specific, individualized feedback that will help them continue to grow and achieve even greater results with their students.

Moving forward, the LDE will remain committed to gathering feedback from educators on Compass and making refinements that will make it a stronger professional development tool for teachers and leaders. The LDE will also be committed to ensuring transparency in the development and implementation of the system through these legislative reports and end-of-year reports. To this end, the following section summarizes the analytic process and selected aggregated results from the 2011-2012 school year, as well as preliminary results for the 2012-13 school year. The preliminary results from the 2012-2013 school year have been gathered from

available records in the online system that collects Compass data from local school districts. As of February 18, 2013, over 28,000 observations had been logged in the system and over 75,000 student learning goals recorded.

III. Value-Added Model Technical Process and Findings

a. Introduction

This section summarizes the pilot examination of student-teacher achievement outcomes for the 2011-2012 school year that were shared with teachers statewide during the spring of 2012. Outcomes were assessed via a value-added model. The assessment used regression of student data (prior achievement, demographics, and attendance) to estimate typical student achievement, and then compared typical outcomes to actual outcomes.

In the context of this report, *value-added analysis* describes the use of prior achievement history, attendance, demographics, and discipline to estimate typical outcomes for students in a specific content (e.g., mathematics), based on a longitudinal data set derived from all students who took state-mandated tests in grades 3 through 9 in Louisiana. The analysis uses a complex model that reflects the grouping of students within classrooms.

The current model, where feasible, was developed to address concerns raised by researchers, policy makers, and other stakeholders regarding variable selection/inclusion and data quality, as they emerged in the application of value-added models. This included the use of a model process that permitted the inclusion of all students with prior achievement data (described below). The high level of test participation in Louisiana results in a substantially more complete database than is commonly available. The predictor variables were expanded to include non-test variables, such as attendance, disability diagnosis, discipline history, and class composition variables to address peer influences on achievement, as requested by ACEE.

b. Combining Student Performance and Other Student Data

Data were drawn from the standardized test files (*iLEAP, LEAP, Algebra I EOC, and Geometry EOC*) for Spring 2009, 2010, 2011, and 2012; the Louisiana Educational Accountability Data System (LEADS) that links students to teachers; and supplemental student databases. Data analyses for 2008-2009, 2009-2010 and 2010-2011 were also conducted to supplement the current year work and provide a point of comparison. The testing and supplemental databases provided data regarding attendance, enrollment, disability diagnosis, limited English proficiency, free or reduced price lunch status, Section 504 status, and disciplinary infractions. Data regarding teachers were drawn from the teacher demographic database (Personnel Education Profile/PEP). A multistage process was used to create

longitudinal records for students describing achievement, attendance, and demographic factors across years. The student and teacher databases were then linked.

Students' data were linked across years. Table 1 presents the number of records available in each content area.

Table 1. Student and Teacher Records Available Overall and in Each Content Area for 2011-2012

	Overall	English Language Arts	Reading	Mathematics	Science	Social Studies	Algebra I	Geometry
Students	256,287	202,416	164,454	202,666	203,038	200,081	35,424	4,759
Teachers	13,687	5,964	5,307	5,226	4,501	4,864	915	417

Several important decision points are noteworthy. Initial records were limited to students who completed one assessment in grades 4-9 to permit the availability of one-year prior achievement data. The testing program begins in the 3rd grade, so, 4th graders would have their matched 3rd grade achievement data as predictors of 4th grade achievement. In order to be included in the analyses, a student was required to be enrolled in the same school from October 1, 2011 to April 9, 2012. Because the student-teacher-course nexus data are collected only once per year, once a student changes schools within that time period it is not possible to ascribe achievement measured at the end of that period to a particular teacher. Students' attendance and achievement records had to be matched to the LEADS curriculum data to identify which courses the students took and who taught those courses. Additionally, the attendance and course databases were used to confirm that the student was enrolled in the same site.

Indicator variables were created to identify student characteristics. Indicator codes identified students as members of the following special education disability groups: emotional disturbance, specific learning disability, mild mental disability, speech/language impairment, other health impairment, or other special education disability. Additionally, indicator codes were used for limited English proficiency, Section 504 status, gifted status, and free lunch and reduced lunch recipients.

The final data structure contained a number of variables used to estimate typical student achievement outcomes and links students to teachers based on the course. Table 2 displays the variables used in analyses that were included in the databases.

Table 2. Student Level Variables Examined

Variable

Emotional Disturbance
Speech and Language Impairment
Mild Mental Disability
Specific Learning Disability
Other Health Impairment
Special Education - Other
Gifted
Section 504
Free Lunch
Reduced Price Lunch
Student Absences
Suspensions (prior year)
Expulsions (prior year)
Prior Mathematics Test (1-3 years based on path)
Prior Reading Test (1-3 years based on path)
Prior Science Test (1-3 years based on path)
Prior Social Studies Test (1-3 years based on path)
Prior English Language Arts Test (1-3 years based on path)
Squares and Cubes of all prior predictors were also entered

c. Value-Added Analysis

Once student and teacher data was compiled, the assessment of student-teacher achievement outcomes was calculated. Students who had multiple teachers in a content area were weighted in proportion to the number of teachers they had in that subject. For example, if a student had two mathematics teachers, the student would have a 0.5 weight in contributing to each teacher's assessment result. Analysis for each content area was conducted separately. The analysis was conducted in three steps. The first two steps were implemented separately for each promotion path and the final step brought all of the data together to obtain student-teacher achievement outcomes.

The steps taken in the value-added process included running a regression model to identify coefficients for each variable used in the model. These coefficients were then used to determine the student's expected score. The expected score was then subtracted from the

student's actual score. This difference for each student belonging to a teacher was then averaged, controlling for any outliers, to calculate the teacher's final result.

Along with individual value-added scores by content, an overall composite rating was provided for the teacher. To calculate the composite percentile, the number of students a teacher instructs in each content area, along with the teacher's specific content area percentile, was compiled into one database with all teachers statewide, regardless of content. The percentile rankings for each content area were converted into a score. If a teacher only teaches in one content area, that teacher's final composite percentile will not change. However, if a teacher has multiple content areas, the teacher's final composite percentile will reflect a weighted average of how he/she scored in all content areas. This composite percentile ranking will be the final value-added evaluation score that is used to determine the teacher's level of effectiveness.

d. Standards of Effectiveness

As mentioned previously, the ACEE committee was responsible for recommending standards of effectiveness for teacher evaluations. These recommendations were submitted and accepted by BESE in December 2011. Table 4 provides the effectiveness ranges used in the 2011-2012 Compass pilot.

Table 4. Ranges for Standards of Effectiveness for the 2011-2012 Results

Effectiveness Level	Total Score	Composite Percentile
Ineffective	1.00 - 1.45	1-10
Effective: Emerging	1.50 - 2.47	11-49
Effective: Proficient	2.50 - 3.48	50-89
Highly Effective	3.50 - 4.00	90-99

Teachers whose value-added, composite percentile falls within the bottom 10th percentile received an ineffective rating. Teachers in the middle 11th - 89th percentile range received a rating of effective. The top 10 percentile of teachers received a rating of highly effective.

As part of the refinements to Compass recommended by the LDE and approved by BESE in January, 2013, the *Highly Effective* range was expanded to include teachers receiving a composite percentile of 80 or above. This change was made to recognize teachers achieving results with students that are significantly above the mean and therefore distinguished from the results of their colleagues in the same content area. In addition, as described in Section II of this report, evaluators will be empowered to give teachers receiving composite percentile scores between 20 and 80 more specific feedback on their performance, by considering student learning target data in conjunction with value-added results. For teachers in this range, evaluators will

have the ability to review both of these data sources and assign a student growth rating of *Effective: Emerging* or *Effective: Proficient*, based on the evidence available. These changes will go into effect for the 2012-13 school year.

e. Selected Results

Stability of Teacher Results across Years in Mathematics and English Language Arts

In order to examine the degree of stability of teacher outcomes across years, two sets of analyses were conducted. These analyses were conducted with the full set of data across 2010-2011 and 2011-2012.

The first analysis examined the stability of teacher ranks across years. Within each year, teachers were ranked as having results that fell in the set standards of effectiveness ranges. The data were examined for the stability of these rankings across years with verified rosters. The results show moderate stability across years. Teachers who fell in the bottom 10th percentile in 2010-2011 were likely to fall in the bottom 10th percentile of results again or to move up one ranking to the 11th - 49th percentile range (mathematics: 78.6%; ELA: 75.2%). They were unlikely to move to the top of the distribution one year later. Teachers who were in the top 10th percentile in 2010-2011 were most likely to fall in the same range or drop by one range to the 50th - 89th percentile in 2011-2012 (mathematics: 85.9%; ELA: 85.8%). They were unlikely to move to the bottom of the distribution one year later. Another way of examining stability is through the correlation coefficient. Again, the data demonstrate moderate stability across years. The rank stability data suggest that there is a group of teachers who will remain in the top or bottom 10th percentile of teachers over consecutive years, and substantive efforts to either improve the results for their students (bottom 10th percentile) or to retain those teachers (top 10th percentile) may be warranted.

Estimated Average Levels of Achievement

Some have questioned whether the value-added model advantages or penalizes teachers of historically low-performing students or teachers of high-performing students. The data demonstrate a nearly zero correlation between typical achievement and teacher effects for either ELA ($r = -0.013$) or mathematics ($r = -0.004$). In response to these questions, LDE will provide new reports to teachers at the beginning and end of the year, allowing them to compare the scores expected of their students, according to the value-added model, and their actual scores, to better understand how their value-added results are derived. (These reports are described in Section II of this report.) To ensure that teachers of high-achieving students are not disadvantaged compared to their peers in terms of demonstrating growth with students according the value-added model, the model allows additional room for high-achieving students to grow when calculating expected scores for them.

Preliminary Findings from the 2012-13 Implementation of Compass

While value-added results for the 2012-13 school year will not be available until the conclusion of the year, the LDE does have preliminary results from teacher observations that have been conducted and recorded thus far. As illustrated in Figure 13, these preliminary findings show that 87% of teachers have been rated *Effective: Proficient* or *Highly Effective*, based upon observations, while only 51% of teachers were rated in these categories according to value-added in the previous year. This suggests that more work is required to ensure that teachers are getting accurate and meaningful feedback from evaluators to help them produce gains in student learning, particularly according to the rigorous expectations of the Common Core State Standards.

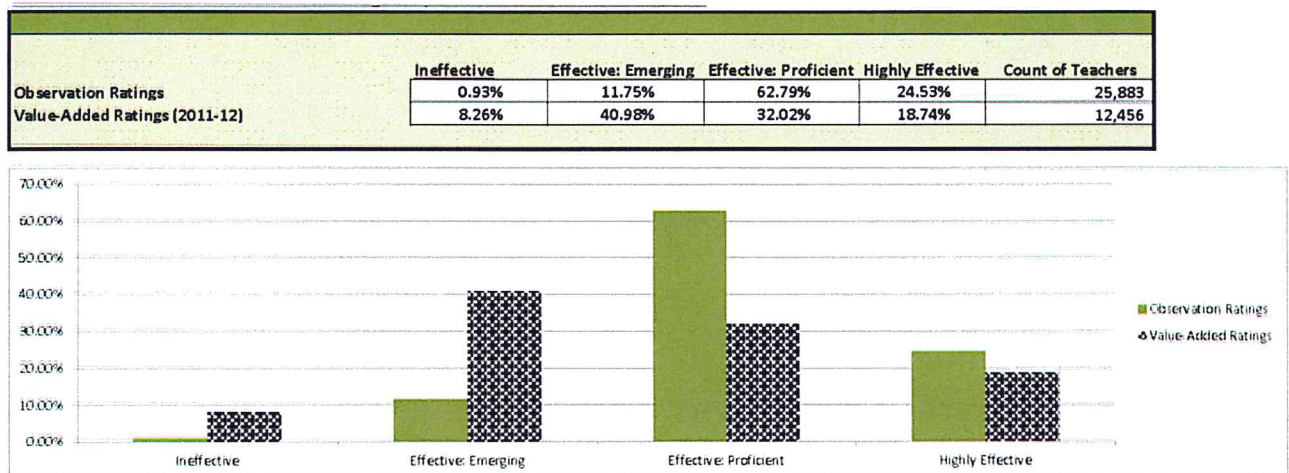


Figure 13. State Distribution of 2012-2013 Observation Ratings and 2011-2012 Value-Added Ratings

IV. Conclusion

The value-added model continues to provide a meaningful and objective reflection of the impact of teacher performance on student achievement, and the tools provided by Compass now offer additional sources of evidence to create a comprehensive picture of educator performance in a given year. This is significant, as it allows evaluators and other instructional leaders to give teachers more meaningful feedback, focused on driving gains in student learning, as Louisiana moves to more rigorous standards.

With the data available following the first year of Compass implementation, policymakers and educators alike will have an unprecedented opportunity to inform their decision-making with

robust information about educators' and students' performance. These data will lead to greater insights for our teachers and leaders about how to refine their craft and even greater results for our students, who will ultimately benefit from this statewide commitment to excellence.

APPENDIX

Selected Results

Distribution of Student-Teacher Achievement Outcomes for 2011-2012

The following figures present the distribution of outcomes across content areas for 2011-2012. The graphs depict the number of teachers (y-axis) with each magnitude of teacher effect (x-axis).

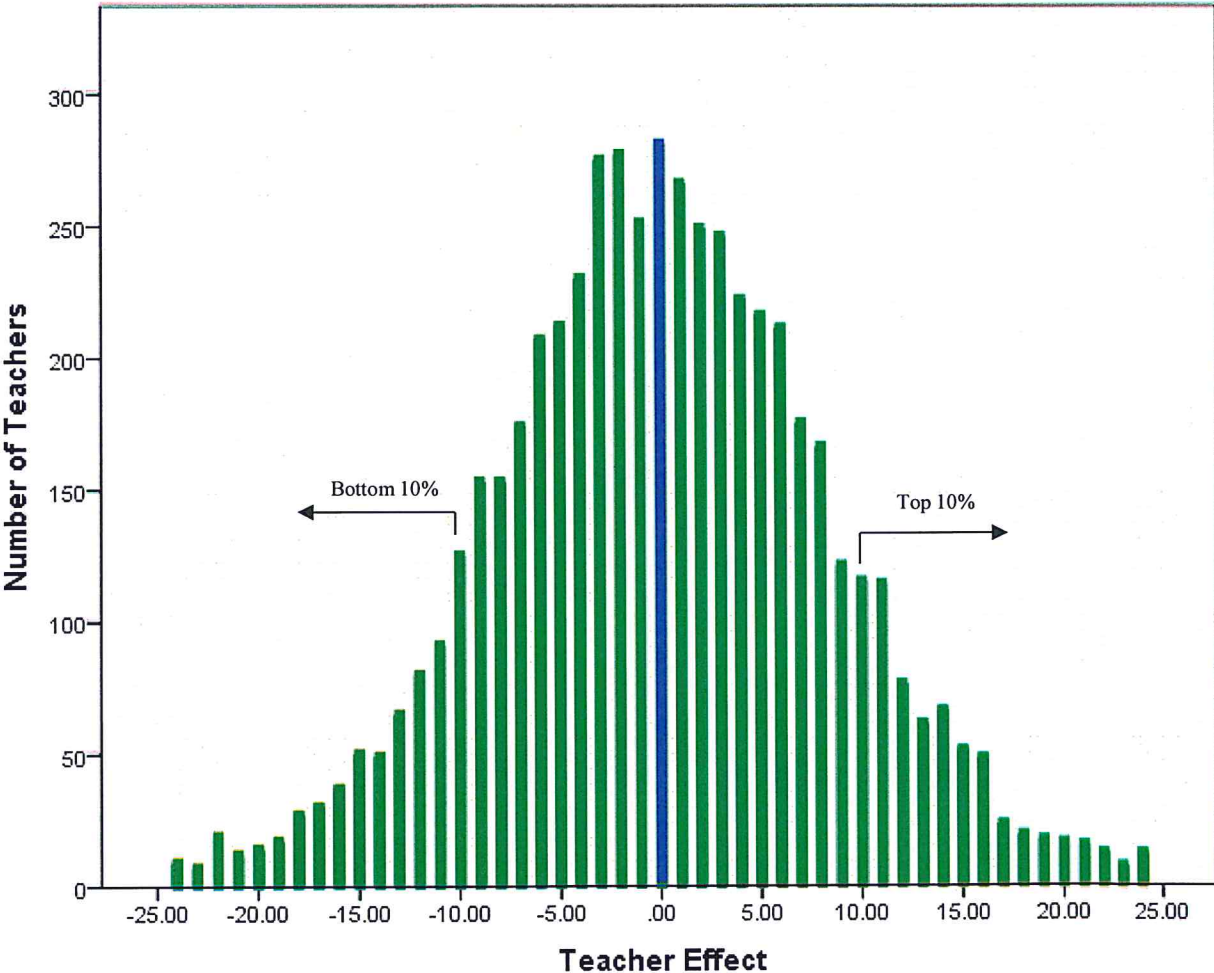


Figure 1. English Language Arts Teacher Effects for 2011-2012

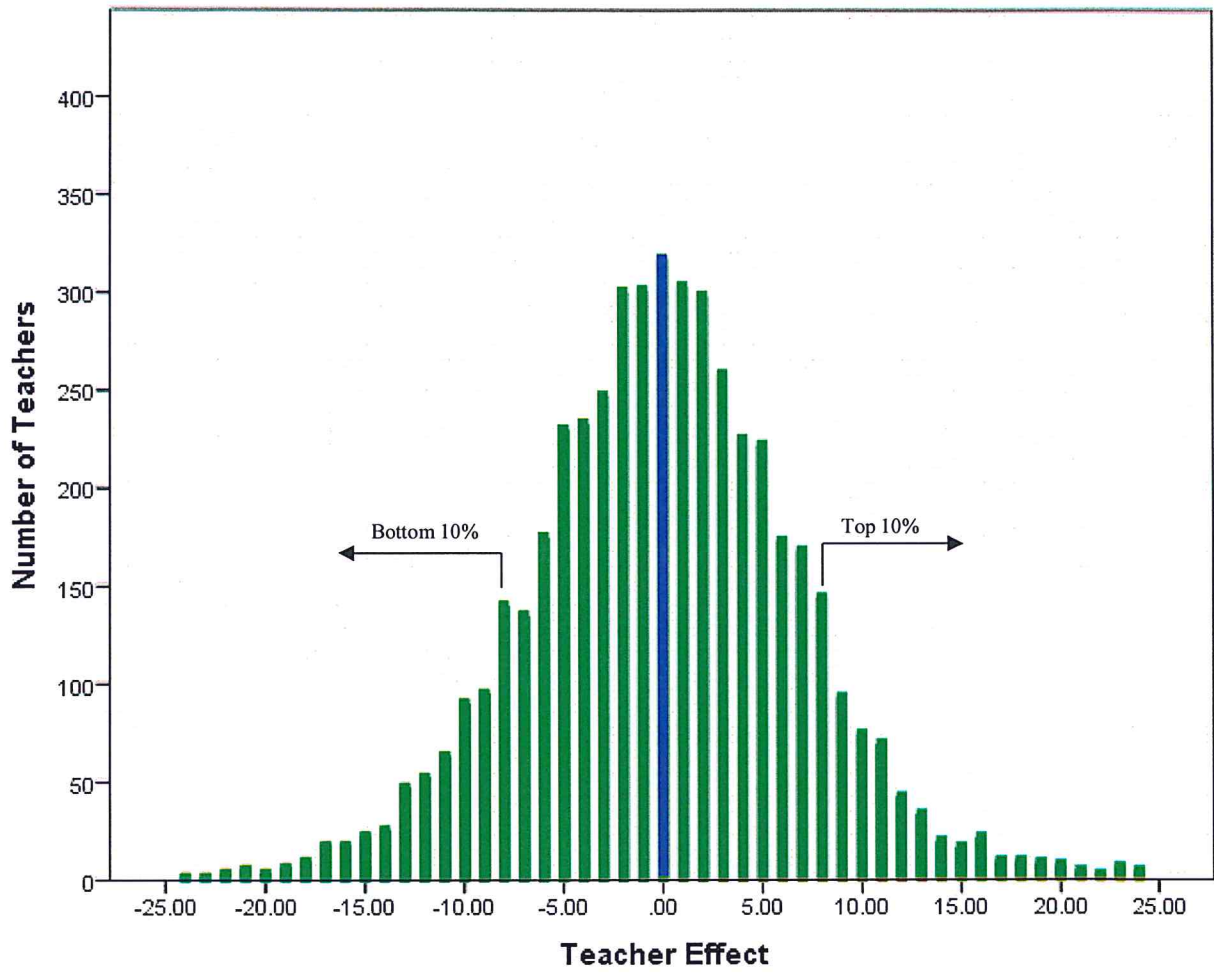


Figure 2. Reading Teacher Effects for 2011-2012

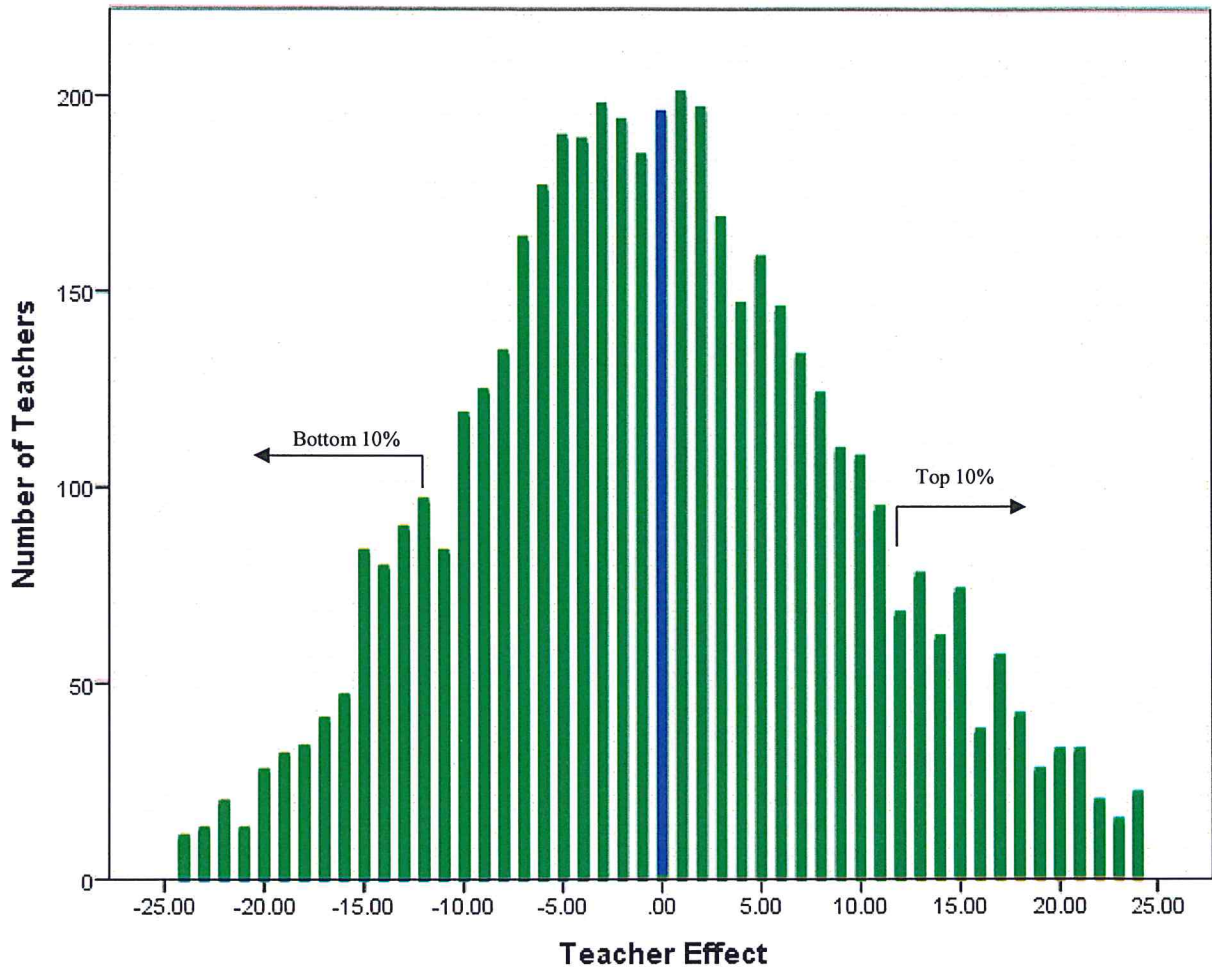


Figure 3. Mathematics Teacher Effects for 2011-2012

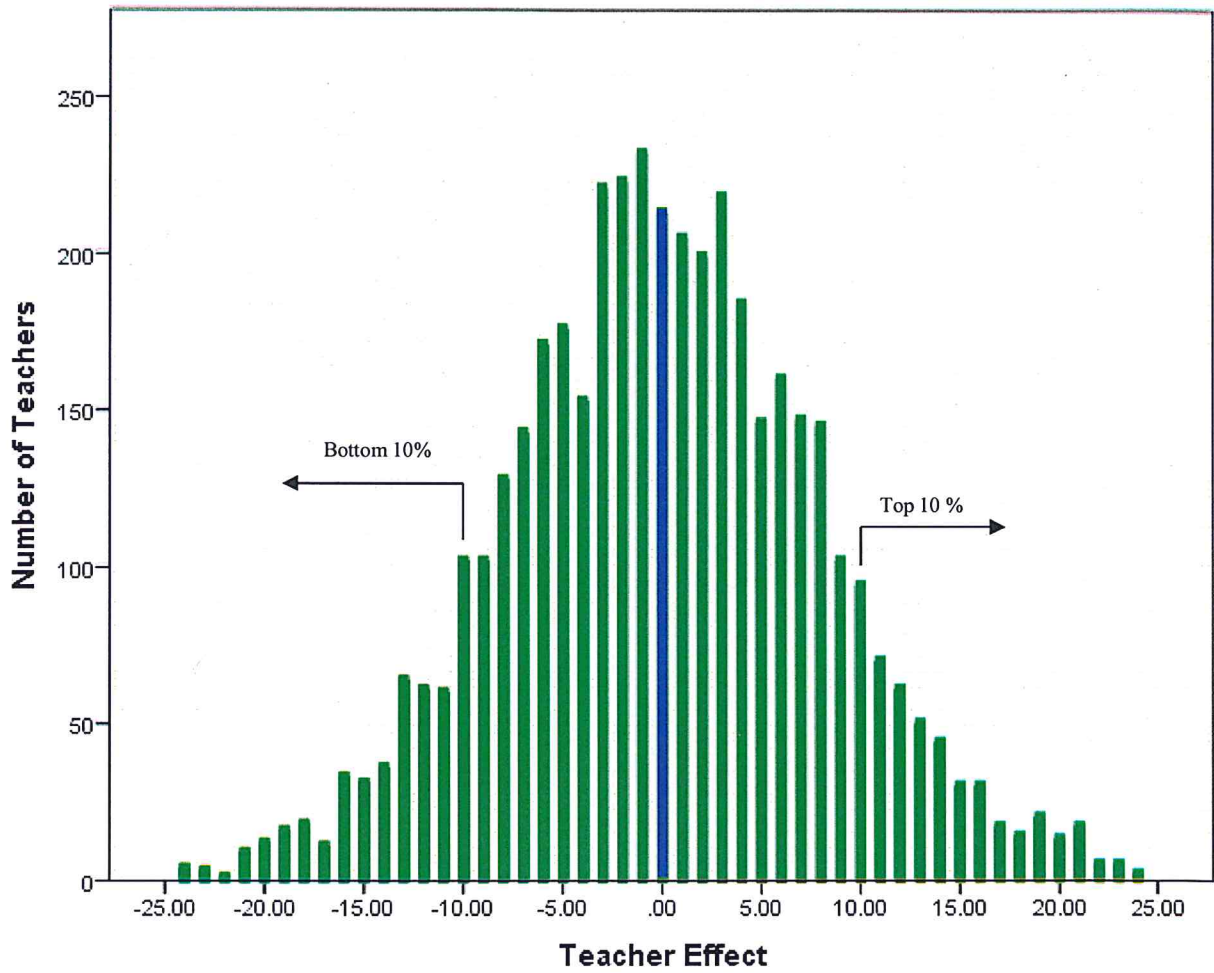


Figure 4. Science Teacher Effects for 2011-2012

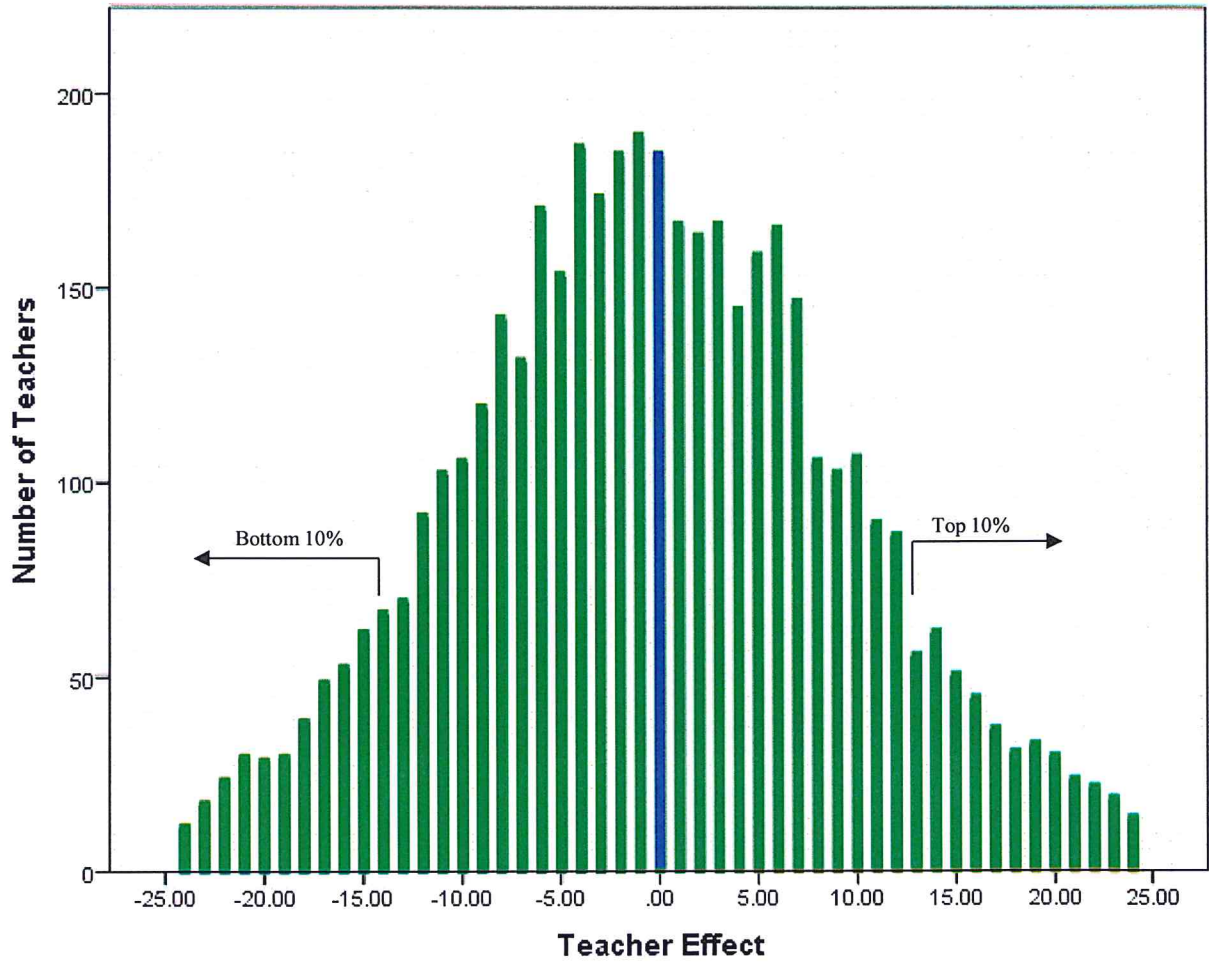


Figure 5. Social Studies Teacher Effects for 2011-2012

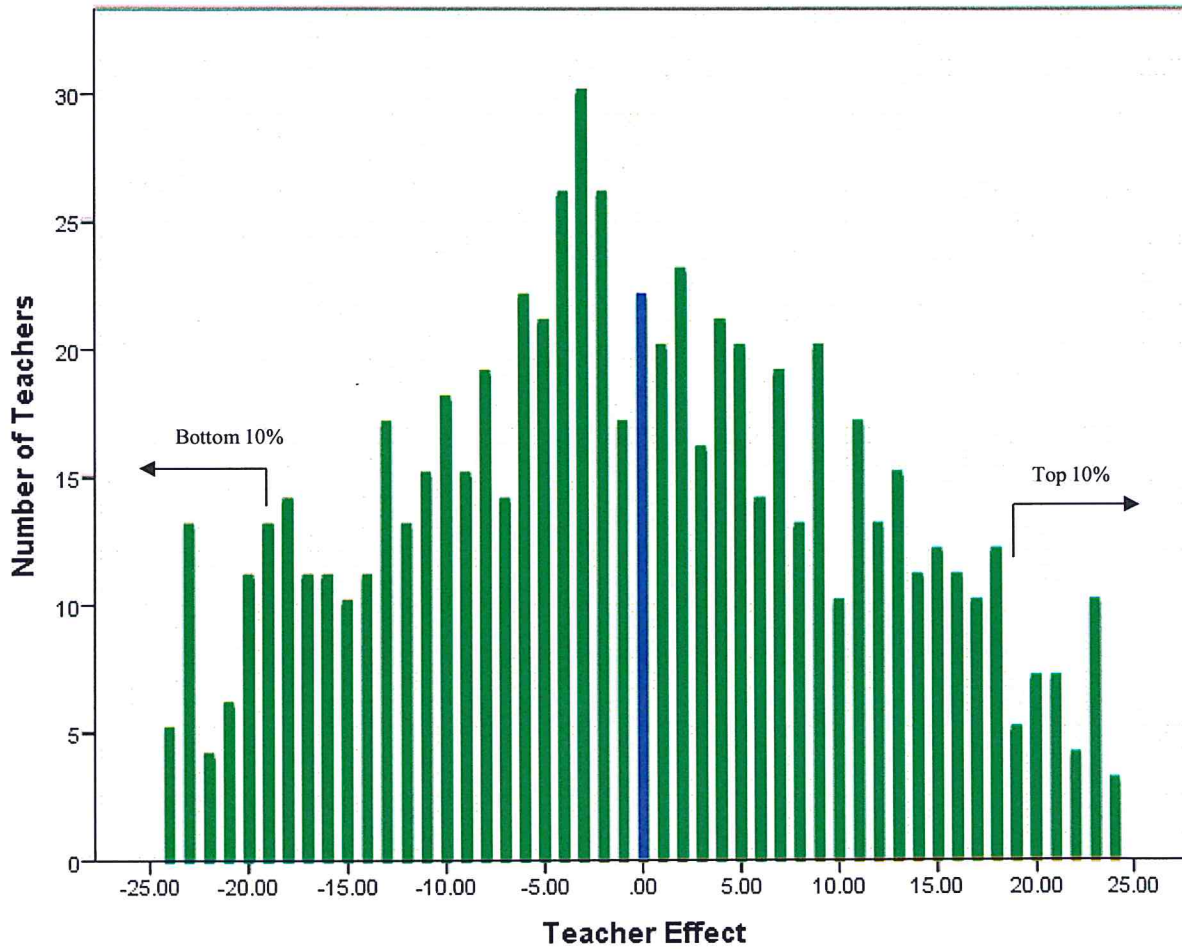


Figure 6. Algebra I Teacher Effects for 2011-2012

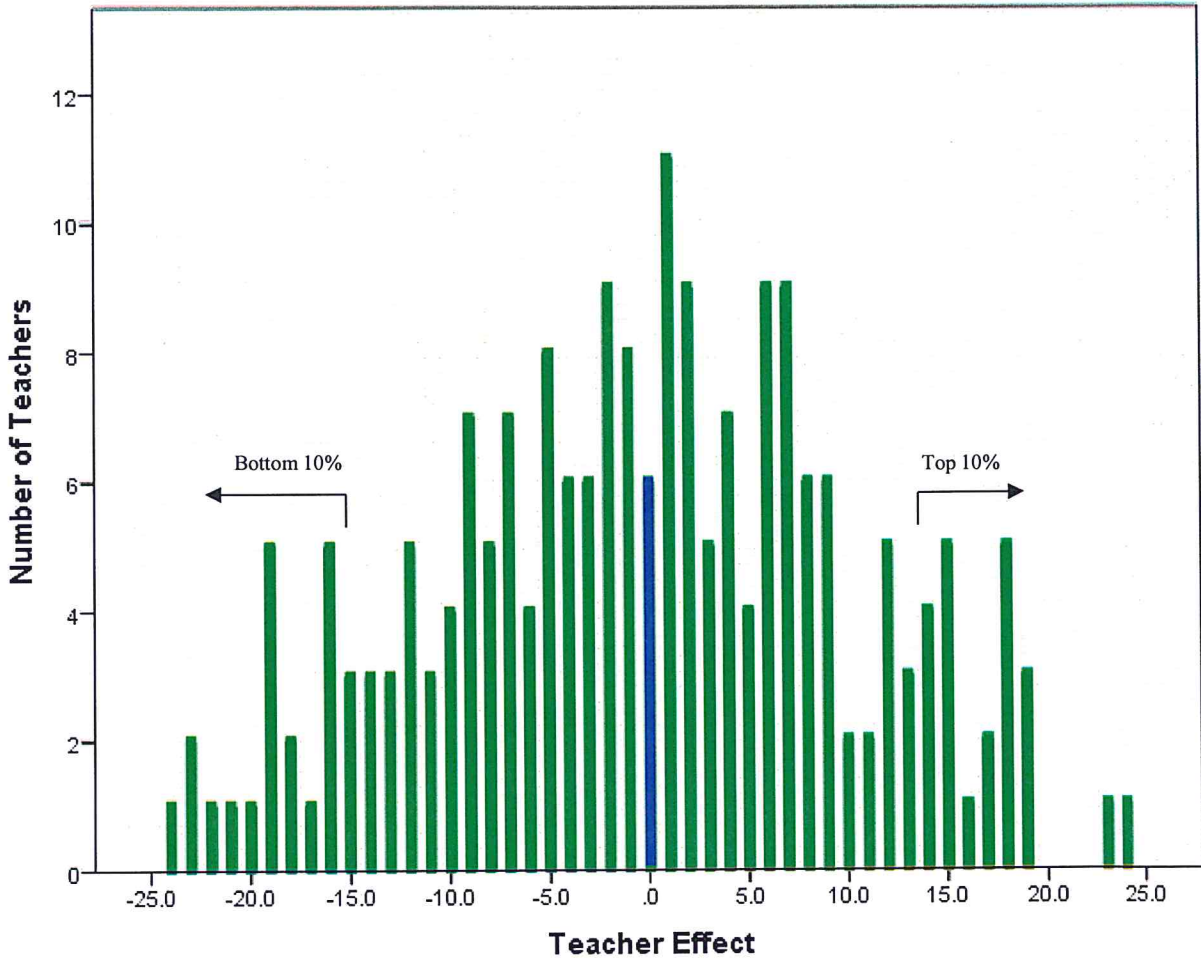


Figure 7. Geometry Teacher Effects for 2011-2012