

Delaware Department of Education

DPAS-II for Administrators (“Year One” Results)

An Addendum to “Continuous Improvement: A Report on ‘Year One’ of the revised DPAS-II Educator Evaluation System”

January 2014

Prepared by:
Teacher and Leader Effectiveness Unit (TLEU)
Delaware Department of Education

Executive Summary

Delaware has invested heavily in revising its educator and administrator evaluation system—the Delaware Performance Appraisal System (DPAS-II)—as an essential component of the state’s efforts to offer all of its students a quality education. Although Delaware has had an evaluation system for many years, it did not have a consistent or sophisticated way to measure each educator and administrator’s impact on student growth. This made it difficult to identify those educators and administrators who were struggling to serve students and who needed extra support. The increased focus on student growth was particularly crucial given the persistent socio-economic and racial achievement gaps among Delaware students and growing evidence that many students are not ready for college and career upon graduation.

The reforms set in place over the last four years, and implemented statewide in the 2012-13 school year, were designed to raise standards for teaching and learning and provide the supports needed for professional growth. The most important and significant change to the DPAS-II system—the revised measure of an educator or administrator’s impact on student achievement (“Component V”)—was designed to give more detailed and objective information about the performance of Delaware students, educators and administrators. A revised DPAS-II system would identify strengths and weaknesses in practice, triangulating this information with student growth data to tailor professional development initiatives and to make decisions regarding promotion, advancement, retention and removal.

This research brief—an addendum to the December 2013 *Continuous Improvement* report on the DPAS-II results for teachers/specialists—provides an analysis of the first year of statewide implementation of the revised DPAS-II system for administrators using evaluation rating data entered in the state’s Evaluation Reporting System (ERS). This brief presents state- and district-level evaluation results and analyzes the relationship between the qualitative and quantitative components of DPAS-II. Similar to the results released in *Continuous Improvement*, the administrator evaluation results show very few differences in administrator ratings on the qualitative evaluation components while student growth ratings show varying performance among administrators. Administrators also receive higher ratings on the student growth measures where they set goals for student growth than on the state-defined metric for student growth using the state assessment. Ultimately, 95 percent of school and district administrators who received a summative evaluation in 2012-13 were rated “effective” or better via the revised DPAS-II system.

The key findings about the administrator evaluation system discussed in the brief include:

There was more variation among administrators in the student improvement component of DPAS-II (Component V) than in the qualitative components

DPAS-II for Administrators 2012-13 Evaluation Results

(Components I-IV, assessing vision and goals, culture of learning, management and professional responsibilities):

- 97 percent of administrators earned “satisfactory” ratings on all four of the qualitative components of the DPAS-II system.
- While student achievement varies across districts, in almost all districts, greater than 90 percent of administrators earned “satisfactory” ratings on Components I through IV.
- On Component V, 8 percent of administrators were rated as “unsatisfactory,” 50 percent were rated “satisfactory” and 42 percent were rated “exceeds.” In comparison, 1 percent of educators were rated “unsatisfactory” on Component V in 2012-13, and 53 percent were rated “exceeds.”
- Overall, 95 percent of administrators were rated “effective” or better and 5 percent of administrators were rated as “ineffective or needs improvement.”

Administrators earned consistently higher scores on the measure of student growth where the targets were set by the administrators and their evaluators than on the state-defined metric for student growth:

- The average score administrators earned on Part A (the state-defined measure of student growth) was 33 out of 50 which was 11 points lower than the average Part B measure of student growth score (44 out of 50) administrators earned.
- There was no correlation between an administrator’s Part A and Part B measures of student growth. Of administrators who opted to have their Component V based on both Part A and B, 71 percent of those with a Part A score of less than 30 out of 50 earned a Part B score of between 40 and 50.
- There was wide variation across Delaware districts in Part B scores. For example, administrators scored an average of 36 out of 50 on Part B in Indian River and an average of 50 out of 50 in Smyrna.

Overview of DPAS-II and “Component V”

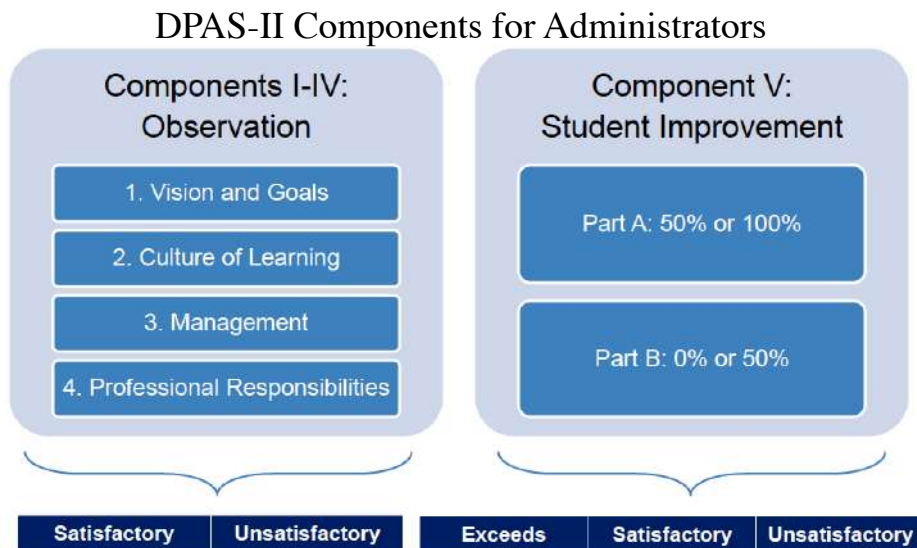
An administrator’s performance is assessed via the DPAS-II system using five components. Components I-IV are grounded in the Interstate School Leaders Licensure Consortium (ISLLC) standards and are based on qualitative assessments of an administrator’s success with the following areas: Vision and Goals, Culture of Learning, Management and Professional Responsibilities. All administrators who are “on cycle” in a given year receive either a “satisfactory” or “unsatisfactory” rating on each of these components. Component V is evaluated yearly (rated as “exceeds,” “satisfactory,” or “unsatisfactory”) based on multiple measures of an administrator’s impact on student improvement. These measures are discussed in further detail below.

Administrators receive a summative rating based on Components I-V every other year with some administrators “on cycle” each year. The summative rating options are: “highly effective,” “effective,” “needs improvement,” or “ineffective.” The newly-created rating of “highly effective” was implemented for the first time in the 2012-2013 school year.

Another important change in 2012-2013 was that an educator or administrator could not earn a summative rating of “effective” if they had an “unsatisfactory” rating in Component V. The opposite also was true. An educator or administrator could not be rated “ineffective” if his or her Component V rating was “satisfactory.”

Component V

The most significant changes in 2012-2013 occurred within Component V (Student Improvement) of the DPAS-II system. In 2012-2013, all educators and administrators were required to show evidence of student growth as a part of their DPAS-II evaluation. There are two parts used to form an administrator’s Component V rating, as seen in the figure below:



DPAS-II for Administrators 2012-13 Evaluation Results

Part A is based upon Delaware Comprehensive Assessment System instructional scale scores for reading and/or mathematics in grades 3 through 10. Administrators select a roster of at least 20 students on which to be evaluated. For each administrator, students from a school, subset of schools, subset of classrooms or the district's schools in their entirety will constitute a "roster" for each "administrator-of-record."¹

Part B is based upon other state-recommended measures, current school success plan measures pertaining to student achievement, or other district priority student achievement measures. These goals for student improvement are decided upon by an administrator and their evaluator.

An administrator and his or her evaluator have the ability to choose whether to be evaluated solely on his or her Part A score, or a combination of Part A and Part B scores where each comprises 50 percent of the Component V overall score (with a maximum of 100 points). Component V is calculated by adding a Part A score to the Part B score if an administrator opted to use both measures. A Part B score will range between 0 and 50 points, and a Part A score is equal to half of the percentage of students on the administrator's roster meeting their growth targets, unless Part A is being taken alone (in which case it comprises 100%). A Component V score of 80 or above is considered "exceeds" between 60 to 79 points is considered "satisfactory" and below 60 points is "unsatisfactory."

¹ For more information on how growth targets for Part A are calculated, please see Appendix.

Distribution of Components I-IV Ratings

During the 2012-13 school year, nearly all administrators (97 percent) who received a summative evaluation were rated “satisfactory” on all four of these components.

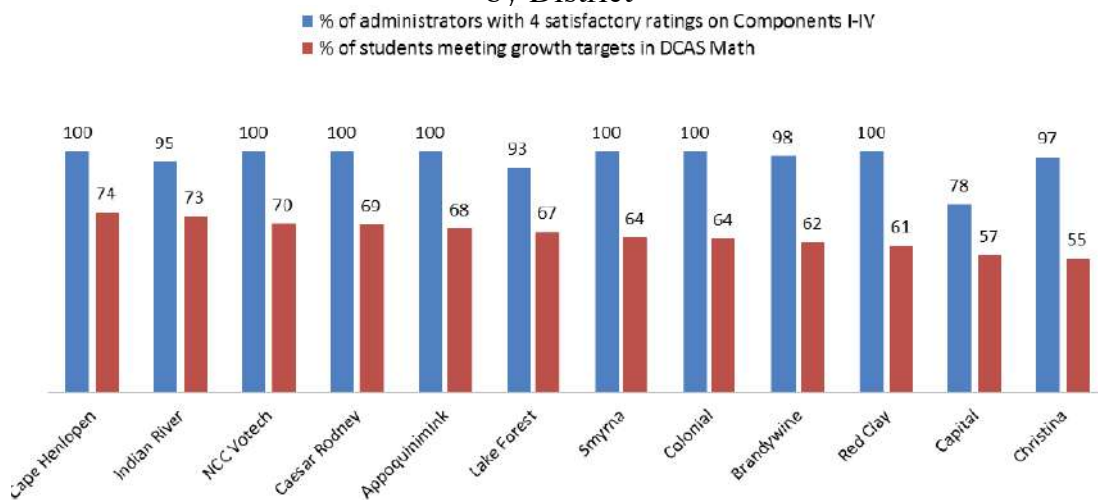
DPAS-II Components 1-4 Distribution for Administrators

| Component I-IV Ratings | Percent |
|---------------------------------------|---------|
| Percent with 2 “Satisfactory” Ratings | 1% |
| Percent with 3 “Satisfactory” Ratings | 2% |
| Percent with 4 “Satisfactory” Ratings | 97% |

Note: Includes 309 administrators with Components 1-4 data entered into ERS.

However, the “satisfactory” ratings many administrators receive are not correlated with students’ success in meeting their growth targets, as shown by the chart below. This chart compares the percentage of administrators in each district who earned “satisfactory” ratings in Components I through IV (blue bars) to the percent of that district’s students that met their DCAS math growth targets in 2012-13 (red bars).

Administrator Component I-IV Ratings vs. Share of Students Meeting Growth targets in Math by District



Notes: Sample includes Delaware districts with 10 or more administrators in a summative evaluation year with Component I-IV ratings entered into ERS. For districts included in this chart, the number of administrators with Component I-IV ratings entered into ERS ranges from 11 in Colonial School District to 45 in Brandywine School District. All data are from state administrative records.

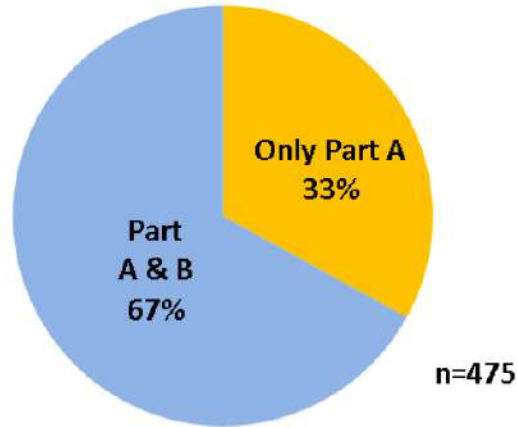
DPAS-II for Administrators 2012-13 Evaluation Results

While the share of students meeting their growth targets varies across the districts (from 74 percent to 55 percent), more than 90 percent of administrators earned “satisfactory” ratings on all four components in each district with the exception of Capital School District (in which case 78 percent of administrators earned four Satisfactory ratings in Components I-IV). These results mirror the teacher evaluation results for the qualitative components (I-IV) of DPAS-II where close to 99 percent of teachers evaluated in each district earned “satisfactory” on each component. Thus, while student performance varies across schools and districts, the qualitative components of administrator and teacher appraisals do not currently present any difference in performance. In the next section we turn to the results of the student improvement component of the DPAS-II system.

Distribution of “Component V” Ratings

The chart below shows the distribution of administrators electing to be evaluated by both Part A and B, or just Part A. Two-thirds of the state’s school and district administrators chose to be evaluated using a combination of Parts A and B, while one-third elected to be evaluated using the state-defined growth targets for their roster of students (Part A) alone.

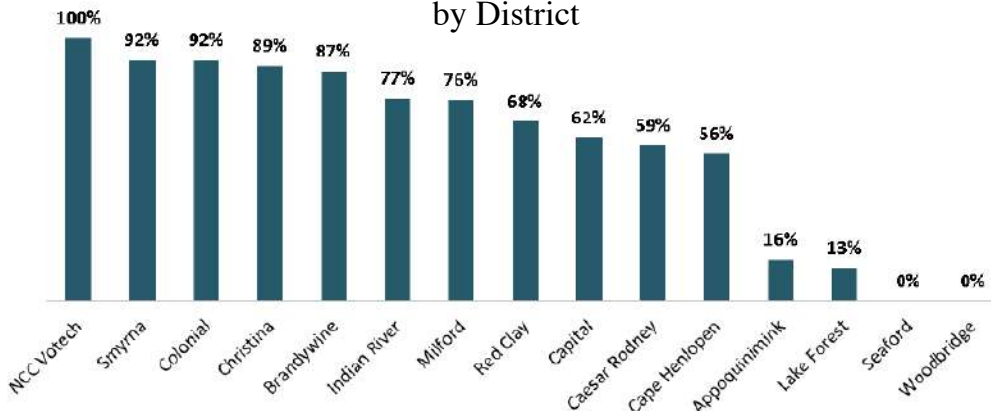
DPAS-II Student Growth Parts Selected by DE Administrators



Notes: Sample includes 475 Delaware district and school administrators with Component V data entered into ERS. All data are from state administrative records.

While the majority of administrators opted to use Part A and B, the approach to Component V varied by district, as seen in the following chart. In certain districts (Seaford and Woodbridge), no administrators chose to be evaluated using any Part B option, while in New Castle County Vo-Tech, all the administrators elected to use a Part A and B combination.

Share of Administrators Selecting Part A & B Option for Component V Rating by District

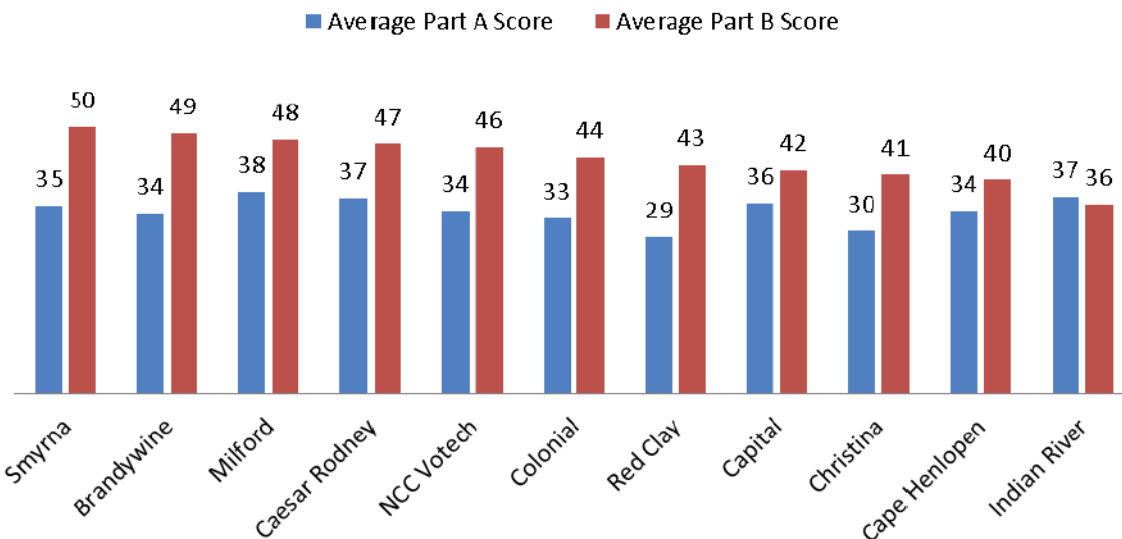


Notes: Sample includes Delaware districts with Component V data for 10 or more administrators entered into ERS. For districts included in this chart, the number of administrators with Component V data entered into ERS ranges from 11 in Woodbridge School District to 82 in Red Clay School District. All data are from state administrative records.

DPAS-II for Administrators 2012-13 Evaluation Results

On average, administrators earned a higher score on their Part A than their Part B measure for Component V. The average Part A score for administrators with both scores was 33 out of 50 while the average Part B score was 44 out of 50. However, there was wide variation across Delaware districts. Administrators earned an average of 36 out of 50 on Part B in Indian River and an average of 50 out of 50 in Smyrna. Indian River School District was the only school district in which the Part B average was lower than the Part A average².

**DPAS-II for Administrators 2012-13 Average Component V Part A & B
Scores (out of 50 points)
by District**



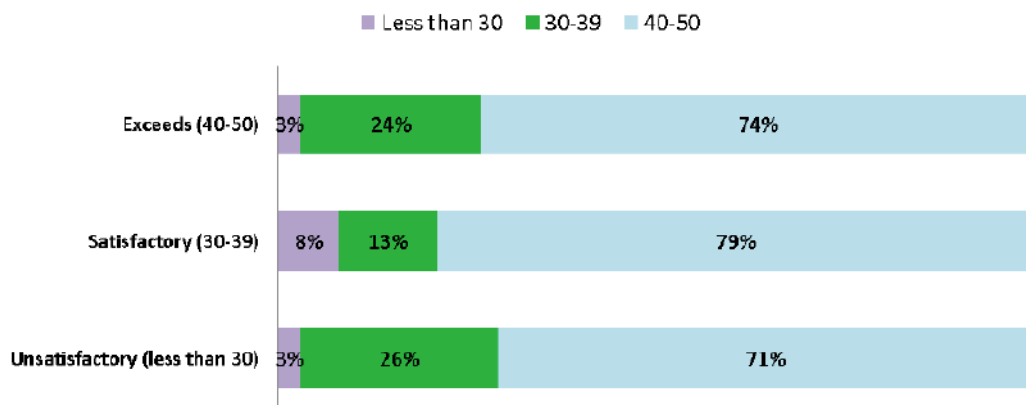
Notes: Sample includes Delaware Districts with Part A and B data for 10 or more administrators entered into ERS. Mean scores are based only on data for administrators with both a Part A and Part B score. For districts included in this chart, the number of administrators with data entered into ERS ranges from 13 in Milford School District to 56 in Red Clay School District. All data are from state administrative records.

Overall, there was no correlation between Part A and B measures of Component V. For administrators scoring at what would be regarded as an “unsatisfactory” level for Part A (less than 60 percent of students met targets for a Part A score lower than 30) the majority earned Part B scores in the range of 40 to 50. The chart below shows a distribution of administrators’ Part B scores grouped by their Part A score. Of all administrators who earned scores lower than 30 out of 50 on that measure, 71 percent earned scores in the range of 40 to 50 on Part B. In contrast, 74 percent of administrators who earned scores in the ranges of 40 to 50 on Part A also earned scores ranging from 40 to 50 on Part B.

² Indian River School District’s Part B for administrators used the DCAS assessment and focused on subgroups.

DPAS-II for Administrators 2012-13 Evaluation Results

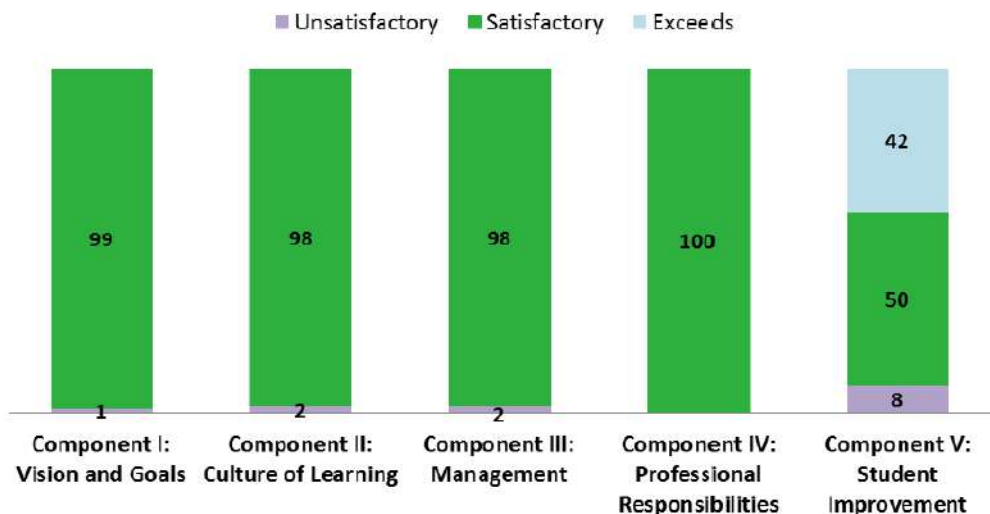
Distribution of Administrator Component V Part B Scores by Part A Scores



Notes: Sample includes administrators with Part A and B scores entered into ERS. 34 Administrators have “exceeds-level” Part A scores, 217 “satisfactory-level,” and 68 “unsatisfactory-level” scores. All data are from Delaware administrative records.

After these Part A and B scores were combined (each weighing a possible 50 points) or Part A was used to determine the entire 100-point Component V scale for administrators, a Component V rating was determined for each school and district administrator. Of all school and district administrators with data entered into Evaluation Reporting System (ERS), 42 percent were rated “exceeds,” 50 percent were rated “satisfactory,” and 8 percent were rated “unsatisfactory.” In contrast, 53 percent of educators were rated “exceeds,” 46 percent “satisfactory,” and 1 percent “unsatisfactory.” For the other components in DPAS-II, at least 98 percent of administrators earned “satisfactory” ratings.

2012-13 DPAS-II for Administrators Component Rating Distribution

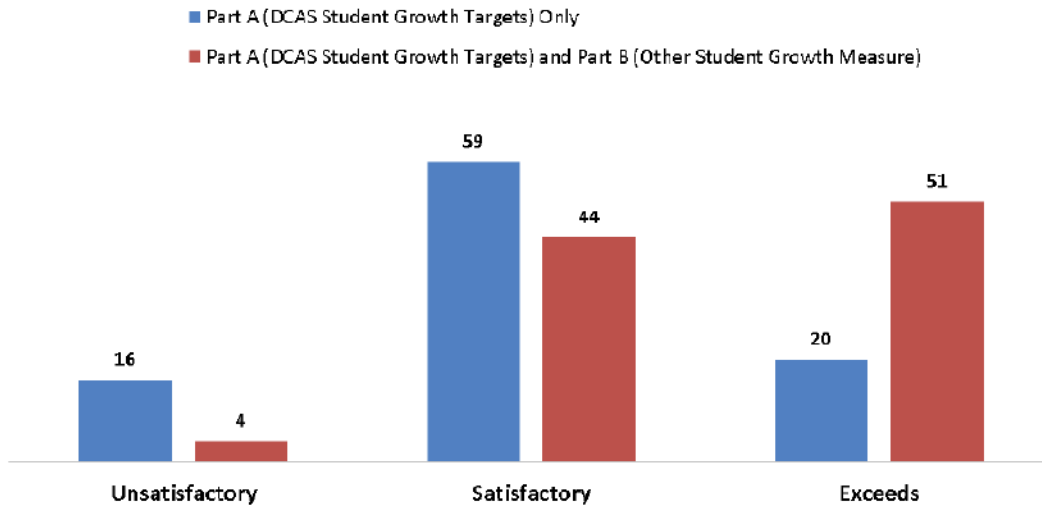


Notes: Only Component V gives an “exceeds” rating option. Sample for Component I-IV distribution includes 318 administrators with Component I-IV data entered into ERS. Sample for Component V distribution includes 484 administrators with Component V data entered into ERS. All data are from Delaware administrative records.

DPAS-II for Administrators 2012-13 Evaluation Results

Since a third of all administrators state-wide chose to be evaluated using both Parts A and B, it is important to see how, if at all, Component V ratings differ among administrators who selected different approaches to Component V. The chart below shows the distribution of Component V ratings for administrators with only Part A or Part A and B scores.

Distribution of 2012-13 Administrator Component V Ratings by Student Growth Measure Option Selected



Notes: Sample includes 148 administrators using only Part A for Component V and 320 administrators using Parts A and B with data entered into ERS. All data are from Delaware administrative records.

Administrators who used Part A and B measures earned higher ratings than those who opted to only use the state-defined measure. Of the administrators who chose just Part A, 20 percent earned “exceeds” ratings, 59 percent earned “satisfactory” and 16 percent earned “unsatisfactory.” For those who were approved to include Parts A and B, 51 percent earned “exceeds,” 44 percent “satisfactory” and 4 percent “unsatisfactory.”

2012-13 Summative Ratings

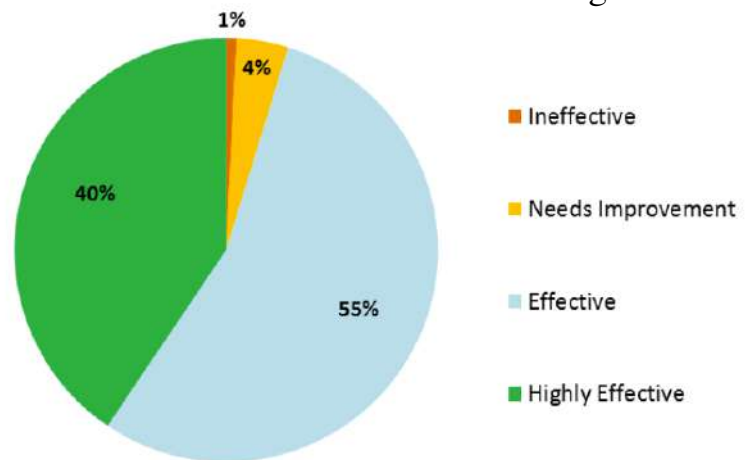
The five components of DPAS-II culminate in a summative rating for administrators every other year. Ratings from the first four components are evaluated alongside an administrator’s Component V score to give an overall rating of “ineffective” “needs improvement” “effective” or “highly effective”. The table below shows how Components I-V are combined to determine the overall summative rating.

Summative Ratings Chart

| Total # of Satisfactory ratings in Components I-IV | Rating in Component V | Summative Rating |
|--|-------------------------|-------------------|
| 3/4 or 4/4 | Exceeds | Highly Effective |
| 3/4 or 4/4 | Satisfactory | Effective |
| 2/4 | Exceeds or Satisfactory | Effective |
| 3/4 or 4/4 | Unsatisfactory | Needs Improvement |
| 0/4 or 1/4 | Exceeds or Satisfactory | Needs Improvement |
| 0/4 or 1/4 or 2/4 | Unsatisfactory | Ineffective |

The chart to the right shows the distribution of overall DPAS-II summative ratings for all administrators considered “on cycle” with data entered into the ERS. Overall, 40 percent of administrators earned the newly-created “highly effective” rating, 50% were rated “effective” and 5% were rated as “ineffective” or “needs improvement.” The distribution of administrators’ summative ratings is comparable to the distribution of ratings for non-administrators in Delaware.; 99% of educators with a summative rating in 2012-13 were rated “effective” or better.

2012-13 DPAS-II for Administrators
Distribution of Summative Ratings



Notes: Sample includes 254 district and school administrators with summative ratings data entered into ERS. All data are from Delaware administrative records.

Appendix

In 2010–2011, the Delaware Department of Education (DDOE) began collecting data to measure student academic growth based on the DCAS. This measure (Part A) makes up one portion of Component V of DPAS-II and is based on the change in performance of students in grades 3 through 10 on DCAS reading and/or mathematics assessments from fall to spring.

The 2012–2013 school year was the first year in which student achievement results using a growth model was fully integrated into evaluations. Administrators select a roster of more than 10 students upon which to be evaluated, however this population can vary based on the role of the administrator as it should be related to his or her responsibilities whenever possible. Ratings for administrators are based on the percentage of their roster’s students’ instructional scores that meet individual student growth targets.

Student growth targets were determined based on the average growth made by “similar” students over the first two years of DCAS (2010–2011 and 2011–2012). “Similar” students are defined as those who had:

- The same fall DCAS instructional score; and
- The same grade and subject; and
- The same Student With Disability (SWD), English Language Learner (ELL), SWD/ELL designation

For the 2011–2012 school year, students had two opportunities in the spring to take the DCAS. The higher of the two scores was utilized in developing the student growth model.

The statistical model that was used to develop expected scores (growth targets) in spring was created by pooling two years of data, i.e., combining all student records for both the 2010–2011 and 2011–2012 school years and creating a regression model for each grade and subject. Using the statistical method of regression analysis, a straight line was drawn that best fits the data. The line represents the average (or statistically expected) spring scores for all students in the state. Using this line, each fall score was associated with a predicted spring growth target score. Students who scored at the highest performance level (PL4) in the fall were expected to remain at this level in the spring. Any student who was at PL4 in the spring was considered to have met his or her growth target for the year.