

## Indiana Department of Education

### Accountability A-F Review: Observed Growth Component Development

#### Purpose

The purpose of this document is to provide details concerning the development of Option D's components. The Department has prepared this information based on 1) data analysis performed for the Accountability System Review Panel, 2) annual A-F Accountability data calculation examination, 3) assistance from external Assessment and Accountability experts, and 4) additional K-12 insight. Information contained within this report is specific to the development of the observed growth metric components and does not reflect details of the remainder of the accountability system. Additional knowledge of the accountability system is assumed by the author.

#### Observed Growth Description

The Department and Accountability System Review Panel have recommended to the State Board of Education the use of Observed Growth (commonly referred to as Option D) in the accountability system. Observed Growth uses a value table with prior year categorical status and current year observed growth score to determine points per student. The points are averaged by subgroup to determine points for the school.

### Observed Growth Component Outline

The Observed Growth metric utilizes various data components to results in a final score per school. Those components are as follows:

- **Point Values**
  - Point values to be awarded based on a student's prior year assessment category status and current year observed growth score.
  - Utilize principles expressed through the Panel.
- **Prior Year Status**
  - A student's prior year assessment status broken into subcategories.
  - Utilizes cut scores to break current statuses into additional subcategories.
- **Observed Growth Target Ranges**
  - Range of observed growth scores associated with negative movement, static movement or positive movement.
  - Point values are associated with each prior year status and growth target range combination.
- **Current Year Observed Growth Score**
  - A student's current observed growth score.
  - Utilizes Indiana Growth Model calculations with criterion, baseline-referenced analysis.

### Process for Setting Components

For each Observed Growth component, specific standards were established. The Department, in conjunction with CECL, completed this work utilizing input from the Panel and assistance from external Assessment and Accountability experts.

The following sections detail the process for establishing each component:

#### **Point Values**

The culmination of the Observed Growth component is the assignment of points per student. The Panel performed an exercise in which the members ranked each outcome on the categorical status improvement table from -2 to 2, with -2 being the least valued outcome and 2 being the most valued outcome. The Panel worked in groups to complete the rankings. The results were then averaged to determine the final values of each result. Each outcome value was translated into points. Various point allocations were analyzed based on different expectations of improvement. Through data review, the Panel elected the assignment of points

more similar to Categorical Status Improvement Option B, in which improving one or more category status or maintaining proficiency receives full points.

In transitioning to Observed Growth Option D, the values from the values table were considered for three primary outcomes: static, improved one category, or declined one category. The points associated with positive movement, static movement and negative movement were assigned under two strategies:

- Point values are adjusted per prior year status category to reflect growth expectations. For this option, point values reflect the diagonal of the categorical status improvement table.
- Point values are constant per proficiency category Did Not Pass, Pass and Pass Plus. Points reflect the growth expectations defined by the Panel. For this option, point values reflect the diagonal of the categorical status improvement table, averaged per proficiency category.

The point values have been tested through iterative data runs which examined impact and school grade distributions. Each iteration considers the original values of the Panel.

### **Prior Year Status**

Indiana currently has three proficiency statuses: Did Not Pass, Pass and Pass Plus. The three statuses are broad and do not supply precision detail concerning student status within each band. Experts recommended that the three primary categories be divided into additional status categories. Cut scores must be established as thresholds between each status category. Experts advised that the definition of subcategories cuts would be an iterative process. Stakeholders determined the following regards:

- Ensure students within the Did Not Pass Status can show incremental progress.
- Create at least 7 status categories

The Department first established 8 preliminary status categories. Each proficiency status point band was split evenly into the defined number of subcategories. The preliminary bands were defined as follows:

- **Pass Plus** Divided into three status categories. Cut scores were established for each grade level, breaking the span of scale score points equally.
- **Pass** Divided into two status categories. Cut scores were established for each grade level breaking the span of scale score points equally.
- **Did Not Pass** Divided into three status categories. Cut scores were established for each

grade level, breaking the span of scale score points equally.

Data was analyzed to determine if student distributions reflected the expected outcome of more refined groups. The results showed the scale could not support three status categories within the Pass Plus status. Additional review indicated the status categories did not provide reasonable distribution among many additional categories. Greater than 95% of students were clustered within the inner most four categories. In effect, only one additional status category was established with this iteration.

**Student Distribution Across Status Categories**

<b>Prior Year Status</b>	<b>Percent Students in Band</b>
DNP1	0.0%
PP2	1.6%
PP1	16.8%
P2	31.3%
P1	31.4%
DNP3	18.2%
DNP2	0.6%

The Department prepared the next iteration of the cut scores based upon feedback from growth experts as well as Panel members. The premise for the second iteration was to provide a more even distribution of the student population. The Department used the prior four years of ISTEP+ data at each grade level to determine cut scores that create status categories as follows:

- **Pass Plus** Divided into two status categories. Cut scores were established for each grade level breaking the prior year student distribution equally.
- **Pass** Divided into three status categories. Cut scores were established for each grade level breaking the prior year student distribution equally.
- **Did Not Pass** Divided into three status categories. Cut scores were established for each grade level breaking the prior year student distribution equally.

Data was analyzed to determine if student distributions reflected the expected outcome of more refined groups. Utilizing this method of cut score setting, students were more evenly distributed across all eight category statuses.

**Student Distribution Across Status Categories with Revised Cut Scores**

Status Improvement Table: Frequency: Math 2013								
Previous Year Level	Current Year Level							
	Did Not Pass-1	Did Not Pass-2	Did Not Pass-3	Pass-1	Pass-2	Pass-3	Pass Plus-1	Pass Plus-2
Pass Plus-2	0.00%	0.00%	0.00%	0.01%	0.11%	0.89%	3.35%	7.62%
Pass Plus -1	0.00%	0.01%	0.01%	0.23%	1.32%	4.62%	6.72%	3.96%
Pass-3	0.00%	0.03%	0.12%	1.45%	4.55%	7.15%	4.87%	1.15%
Pass-2	0.02%	0.21%	0.62%	4.39%	6.50%	4.68%	1.57%	0.18%
Pass-1	0.14%	1.21%	2.28%	7.33%	4.59%	1.75%	0.36%	0.02%
Did Not Pass-3	0.24%	1.32%	1.56%	2.68%	0.91%	0.23%	0.04%	0.00%
Did Not Pass-2	0.79%	2.20%	1.51%	1.59%	0.33%	0.08%	0.01%	0.00%
Did Not Pass-1	0.94%	0.96%	0.33%	0.20%	0.04%	0.01%	0.00%	0.00%
<b>TOTAL:</b>	<b>2.13%</b>	<b>5.93%</b>	<b>6.44%</b>	<b>17.87%</b>	<b>18.35%</b>	<b>19.41%</b>	<b>16.92%</b>	<b>12.94%</b>

Status Improvement Table: Frequency: ELA 2013								
Previous Year Level	Current Year Level							
	Did Not Pass-1	Did Not Pass-2	Did Not Pass-3	Pass-1	Pass-2	Pass-3	Pass Plus-1	Pass Plus-2
Pass Plus-2	0.00%	0.00%	0.00%	0.04%	0.28%	1.55%	2.61%	4.12%
Pass Plus -1	0.00%	0.01%	0.03%	0.31%	1.37%	4.05%	3.55%	2.55%
Pass-3	0.01%	0.07%	0.23%	2.09%	5.43%	7.93%	3.83%	1.58%
Pass-2	0.03%	0.38%	1.09%	5.74%	7.13%	4.78%	1.20%	0.26%
Pass-1	0.24%	1.99%	3.42%	8.62%	4.63%	1.55%	0.22%	0.04%
Did Not Pass-3	0.32%	1.92%	2.03%	2.57%	0.68%	0.15%	0.01%	0.00%
Did Not Pass-2	1.06%	2.79%	1.58%	1.16%	0.21%	0.04%	0.00%	0.00%
Did Not Pass-1	1.25%	0.91%	0.22%	0.10%	0.02%	0.00%	0.00%	0.00%
<b>TOTAL:</b>	<b>2.91%</b>	<b>8.08%</b>	<b>8.59%</b>	<b>20.63%</b>	<b>19.76%</b>	<b>20.05%</b>	<b>11.42%</b>	<b>8.55%</b>

The Department, CEI and industry experts defined the evaluation to prove the validity of the cut scores set for each area. The examination explored distribution analysis, student growth percentile information and minimum size of scale score bands.

As noted in the above table, this iteration of cut scores provides a reasonable distribution of the student population in the all eight status categories. In addition, the cut scores provide for increased movement across category status lines, with 60% of students either increasing one or more categories or decreasing one or more categories.

The cut scores were also reviewed through student growth analysis. For each status category, the median and mean student growth percentiles were examined for those students that did

not transition a status category. The median student growth percentile of the static students shows an incremental increase across each category status. This observation provides evidence that the bands shows a difference in growth levels as well as higher median growth within the high performing bands than the lower performing bands. Mean and median growth required to move up one category within the next year also was reviewed for consistent intervals. This review also provided evidence that the breaks in the status categories provide evidence of viable cut scores.

**Mean Student Growth in Each Category Status**

2013				
ELA_PRYR_IMP_STAT	Mean Growth	MATH_PRYR_IMP_STAT	Mean Growth	Overall
PP2	63	PP2	65	64
PP1	59	PP1	54	56.5
P3	54	P3	51	52.5
P2	52	P2	48	50
P1	48	P1	43	45.5
DNP3	44	DNP3	37	40.5
DNP2	38	DNP2	32	35
DNP1	29	DNP1	27	28

2012				
ELA_PRYR_IMP_STAT	Mean Growth	MATH_PRYR_IMP_STAT	Mean Growth	Overall
PP2	68	PP2	66	67
PP1	59	PP1	55	57
P3	53	P3	52	52.5
P2	52	P2	48	50
P1	48	P1	44	46
DNP3	44	DNP3	40	42
DNP2	38	DNP2	36	37
DNP1	30	DNP1	25	27.5

Finally, scale score point spans in each status category for each grade level were examined by Department Assessment staff to confirm sufficient scale score points existed within each band. The minimum threshold for point bands was set to 20 points. This examination confirmed each band contained sufficient points.



## Observed Growth Target Ranges

The Observed Growth target ranges and movement classifications are based on the examination of the Categorical Status Improvement model and the observed pattern of incremental mean growth across category statuses. The Observed Growth Target Ranges are a key component for establishing points with the values table. The target ranges create a threshold for which points are to be assigned. In review of student distribution charts for the Categorical Status Improvement, it was found that between 85% and 90% of the students fell within one of three outcomes: static, improved one category, or declined one category. This diagonal of the Categorical Status Improvement table became the focus to the Observed Growth option. Three movement classifications were created to categorize the type of category movement present within each band.

The Option D modified values table is an expansion of the previous Option C, in which students received a full point for either actual positive category movement or observed growth, showing a student was on target achieve a positive category movement within the next year. In transitioning to Observed Growth Option D, the actual outcome of category movement was incorporated into the value table through analysis of the Observed Growth. The mean Observed Growth for each movement classification was analyzed to establish target ranges that represent actual category movement or trajectory-to-category movement. Two strategies were used to create target growth ranges:

- Target ranges are constant within each proficiency category Did Not Pass, Pass and Pass Plus and point values are adjusted per prior year status category to reflect growth expectations. For this option, the target ranges were established based on the median growth observed in each proficiency category. Positive movement targets are based on the median observed growth for students experiencing positive growth through categorical status improvement. Negative movement targets are based on the median observed growth for students experiencing negative growth through categorical status improvement. The target ranges were validated through analysis to actual category change outcomes. Students who changed a category status through Categorical Status Improvement were compared to the Observed Growth movement classification. It was found that between 85% and 90% of those students showing movement through Categorical Status Improvement showed the same movement positive or negative movement through the Observed Growth target ranges.
- Target ranges are specific to each status category and point values are constant for each proficiency category Did Not Pass, Pass and Pass Plus. For this option, the target ranges were established based of the median growth observed in each proficiency category for students not changing statuses through Categorical Status Improvement. Positive

movement targets are based on the median observed growth for students experiencing positive growth through categorical status improvement offset to reflect the median growth of the students not changing categories. Negative movement targets are based on the median observed growth for students experiencing negative growth through categorical status improvement offset to reflect the median growth of the students not changing categories. The target ranges were validated through analysis to actual category change outcomes. Students who changed a category status through Categorical Status Improvement were compared to the Observed Growth movement classification. The target ranges were validated through analysis to actual category change outcomes. Students who changed a category status through Categorical Status Improvement were compared to the Observed Growth movement classification. It was found that between 90% and 95% of those students showing movement through Categorical Status Improvement showed the same movement positive or negative movement through the Observed Growth target ranges.

The Panel elected to utilize target ranges specific to each prior year category status.

### **Current Year Observed Growth Score**

The current year observed growth metric is a component of the Indiana Growth Model. Observed growth is a criterion baseline-referenced growth calculation. The use of baseline referencing creates an anchored growth metric which meets the growth requirements established in IC 20-31-8-5.4(a). Multiple years of data are used to establish matrices or reference tables for growth. These tables are used for each student, using only the student's own data to determine the student's observed growth score. Key elements of Observed Growth include:

- Score is assessed based on student's own data.
- Baseline matrices are static for a period of time defined by stakeholders (i.e. evaluated every 5 years).
- Baselines are not normed annually as done for norm-referenced SGP analyses.
- Observed Growth is a criterion metric of growth.
- It is possible for all students to achieve the highest levels of Observed Growth.

The Department would continue to calculate growth data and provide information to schools through secure reports and public reports. Because the underlying matrices are defined and static in the baseline process, this information can be exposed to stakeholders to allow schools

to calculate the observed growth for students or validate the calculations provided by the Department.

### **Summary**

Through statistical analysis, the Department has concluded that the process for establishing each data component is valid and defensible. This conclusion is supported by additional evidence provided by external growth and accountability experts.