

Aggregating Measures for Accountability Determinations

**Background Information for the
Arizona School Accountability
Components & System RFI**

Ryan Reyna

Key Issues in Aggregation

1. Purpose of System

2. Type of Aggregation

1. Method for Calculating Measures

2. “Weights”

3. Display of determination

Purpose of Accountability System

- Accountability systems may serve multiple purposes, including
 - Performance information for parents, students, policymakers and the public
 - Designations for reward, support and consequence
 - Continuous improvement of teaching and learning in classroom
- Purposes are not mutually exclusive; **however** the main value a state places on its accountability system will lead it to make very different choices in the way it calculates and aggregates performance

Types of Accountability Aggregation

1. Index
2. Matrix
3. Goal-based
4. Dashboard



Type of Aggregation – Index

Numerical Combination of Performance Across Measures

- Benefits:
 - Simple for parents and the public to understand
 - Can maximize differentiation between schools and create clarity between the rating thresholds
- Limitations:
 - Can minimize transparency of performance on individual measures
 - May be difficult to weight appropriately
- Key considerations:
 - Policy “weights” may not match numerical weights in calculation
 - Performance thresholds can be set for individual metrics and/or overall, which can lead to unintended consequences
 - Performance thresholds can be normative (i.e., top 5% of schools) or criterion-based (i.e., >90% performance)
- Examples: DE, KY, FL, WV, CA CORE

Index Example – High School

Area/Measures	Weight	Points
<i>Academic Achievement</i>	25%	50
Proficiency ELA	10%	20
Proficiency Math	10%	20
Proficiency Science	5%	10
<i>Growth</i>	40%	80
Growth in ELA	15%	30
Growth in Math	15%	30
Progress in EL Proficiency	10%	20
<i>On Track to CCR</i>	20%	40
On Track to College and Career Ready in 9 th Grade	5%	10
4-year Cohort Graduation Rate	10%	20
6-year Cohort Graduation Rate	5%	10
<i>College and Career Readiness</i>	15%	30
College and Career Preparation	10%	20
College and Career Transitions	5%	10
Total	100%	200

Type of Aggregation – Matrix

Balance of Performance in Two Domains

- **Benefits:**
 - Provides parents and public with more transparent information about both current performance and improvement
 - Can better reflect the policy values of the accountability system
- **Limitations:**
 - Can be more difficult to explain to stakeholders as few examples currently exist
 - May be harder to establish federal school classifications
- **Key considerations:**
 - Must decide which indicators fit into which domain or whether to use improvement as its own domain
 - Establishing cut lines requires political agreement regarding accountability system values (i.e., importance of status vs. growth)
 - Proposed regulations requiring each measure to have a performance designation of at least 3 levels can complicate the communication of a matrix
- **Examples:** California (potential), also found in many educator evaluation systems

Matrix Example 1 – Indicators

Status

- ELA
- Math
- Science
- Chronic Absence
- 4-year Grad Rate
- On-track to CCR
- College and Career Preparation

Progress

- ELA growth
- Math growth
- Progress in EL Proficiency
- Extended-year Grad Rate
- College and Career Transitions

Status	5	C	B	B	A	A
	4	C	B	B	A	A
	3	D	C	B	B	A
	2	D	D	C	B	B
	1	F	D	D	C	C
			1	2	3	4
		Progress				

Note: Categories are for illustrative purposes. They may shift according to SEA/stakeholder values and could change over time.

Matrix Example 2 – Performance Over Time

Status – Current Year

- ELA
- Math
- Science
- Chronic absence
- Growth ELA
- Growth Math
- Progress in EL Proficiency
- 4- and Extended year Grad Rate
- College and Career Preparation
- College and Career Transition

Status	5	C	B	A	A	A
	4	C	B	B	A	A
	3	D	C	B	B	B
	2	D	D	C	C	B
	1	F	D	D	C	C
			1	2	3	4
		Progress				

Progress – Year to Year Improvement

- All or subset of above measures

Note: Categories are for illustrative purposes. They may shift according to SEA/stakeholder values and could change over time.

Type of Aggregation – Goal-based

Performance Against Specific Benchmarks

- Benefits:
 - Simple for parents and the public to understand
 - Has historical precedence
 - Can take advantage of new predictive analytic models
- Limitations:
 - Can narrow focus on improvement to students near benchmark
 - May discourage schools far from the benchmark
- Key considerations:
 - Negative historical connotations may discourage next-generation models
 - Proposed regulations require at least 3 performance levels, which increases the number of decisions about the “appropriate” level of performance, and whether/how that differs for different schools or student populations
- Ex: pre-waiver NCLB, TN (mix of Goal and Index), CT (mix of Goal and Index),

Goal-based Example – Theoretical

Measure	Performance	Goal	Status
Proficiency ELA	63%	60%	Yes
Proficiency Math	49%	60%	No
Proficiency Science	71%	70%	Yes
Progress in ELP	<N size	60%	<N size
Growth ELA	Above Avg	Average	Yes
Growth Math	Above Avg	Average	Yes
On Track to CCR in 9th	75%	70%	Yes
4-year Grad	81%	83%	No
6-year Grad	86%	85%	Yes
College and Career Readiness	41%	35%	Yes
College and Career Transitions	22%	25%	No

Note: Determination based on either meeting the goal on X number of indicators or creating a minimum “performance gate” on specific indicators to reach a certain letter grade (i.e., only schools that meet Proficiency goals are eligible for A letter grade). Goals can be set statewide and/or based on peer performance or statistical projections

Goal-based Example – Connecticut

- **Chronic Absenteeism**
 - Full points awarded if the chronic absenteeism rate is 5% or lower
 - No points awarded if rate is 30% or greater
 - Chronic absenteeism rates between 30% and 5% will be awarded proportional points
- **Preparation for Postsecondary and Career Readiness**
 - Ultimate target is 75%. Points will be prorated based on the percentage of the ultimate target achieved.
- **On Track in 9th Grade**
 - Ultimate target is 94%. Points will be prorated based on the percentage of the ultimate target achieved.

Type of Aggregation – Dashboard

Data Presentation without Specific Performance Ratings

- Pro:
 - Can maximize transparency of performance on individual measures and minimize performance threshold decisions
 - Allows the stakeholder to determine its own values about the data
- Con:
 - More difficult for the public to interpret overall performance across schools
 - More difficult for educators/administrators in low performing schools to understand why a specific school was identified in a federal improvement category
 - Effective communication with dashboards takes considerable design work
- Key considerations:
 - As currently proposed, regulations would not allow this model
 - Another aggregation approach “behind the scenes” would be used to identify comprehensive and targeted support and intervention schools
 - Can include comparison data (to similar schools, district, state) to provide additional performance context
- Example: Illinois

Dashboard Example – Illinois

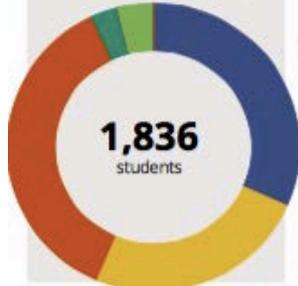
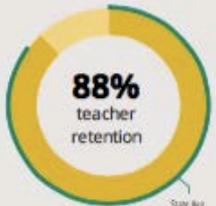
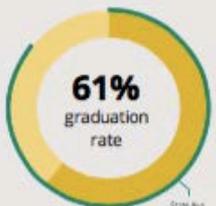
JEFFERSON HIGH SCHOOL

4145 SAMUELSON RD ROCKFORD, IL 61109 3249 (815) 814-9536

Grades: 9-12
District: ROCKFORD SD 205

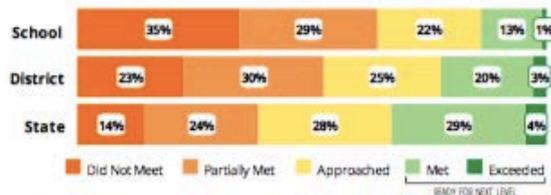
Principal: Don Rundal
Superintendent: Dr. Jhen Janet

FAST FACTS



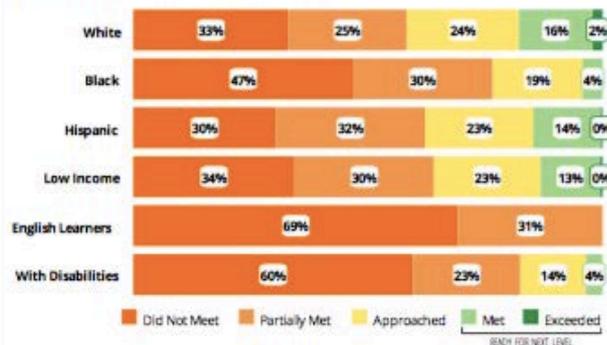
Academic Success

All Illinois students in grades 3-8 take the PARCC assessment each year. High school students take the PARCC in specific Math or English Language Arts (ELA) courses.



Success by Student Group

This display shows PARCC performance levels for each student group. No data is shown for groups with fewer than 10 students.

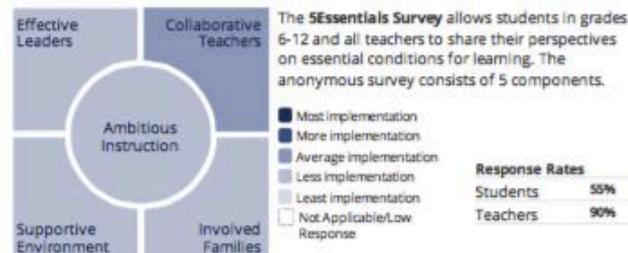


Student Characteristics

White	32%	Low Income	87%
Black	25%	English Learners	3%
Hispanic	37%	With Disabilities	16%
Asian	3%	Homeless	7%
American Indian	0%		
Two or More Races	4%		
Pacific Islander	0%		

School Environment

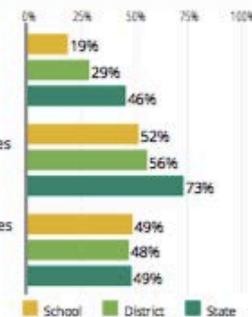
The **5Essentials Survey** allows students in grades 6-12 and all teachers to share their perspectives on essential conditions for learning. The anonymous survey consists of 5 components.



- Less Effective Leaders**
Do principals and teachers implement a shared vision for success?
- Average Collaborative Teachers**
Do teachers collaborate to promote professional growth?
- Less Ambitious Instruction**
Are the classes challenging and engaging?
- Less Supportive Environment**
Is the school safe, demanding, and supportive?
- Less Involved Families**
Does the entire staff build strong external relationships?

College Readiness

Ready for College Coursework
Students who meet or exceed ACT college readiness benchmarks



Postsecondary Enrollment

Students who enroll at colleges and universities

Postsecondary Remediation

(lower is better)
Students enrolled in Illinois community colleges who require remedial coursework

Identifying the Bottom 5%

- System **must** identify bottom 5%

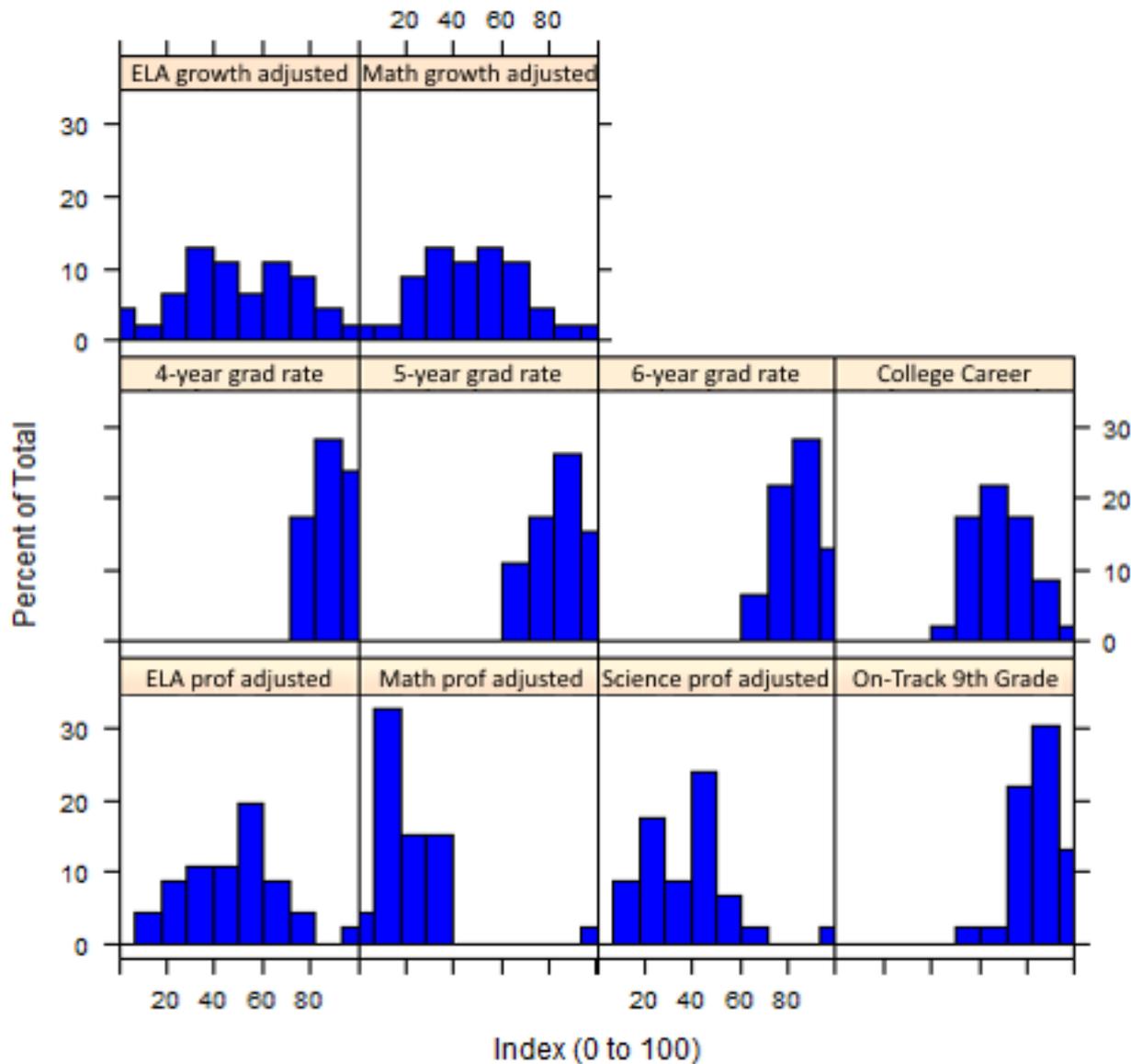
5% of what?

- Index score makes this relatively easy
 - May be affected by statistical weights
- Matrix interpreted to mean 5% in the “bottom left”
 - May require resetting “bottom left” annually
- Goal-oriented depends on relative ambition of goals
 - Certain types of schools may have “hard” goals and end up in bottom 5% (and vice versa)
- Dashboard will need to use one of the above to identify 5% behind the scenes

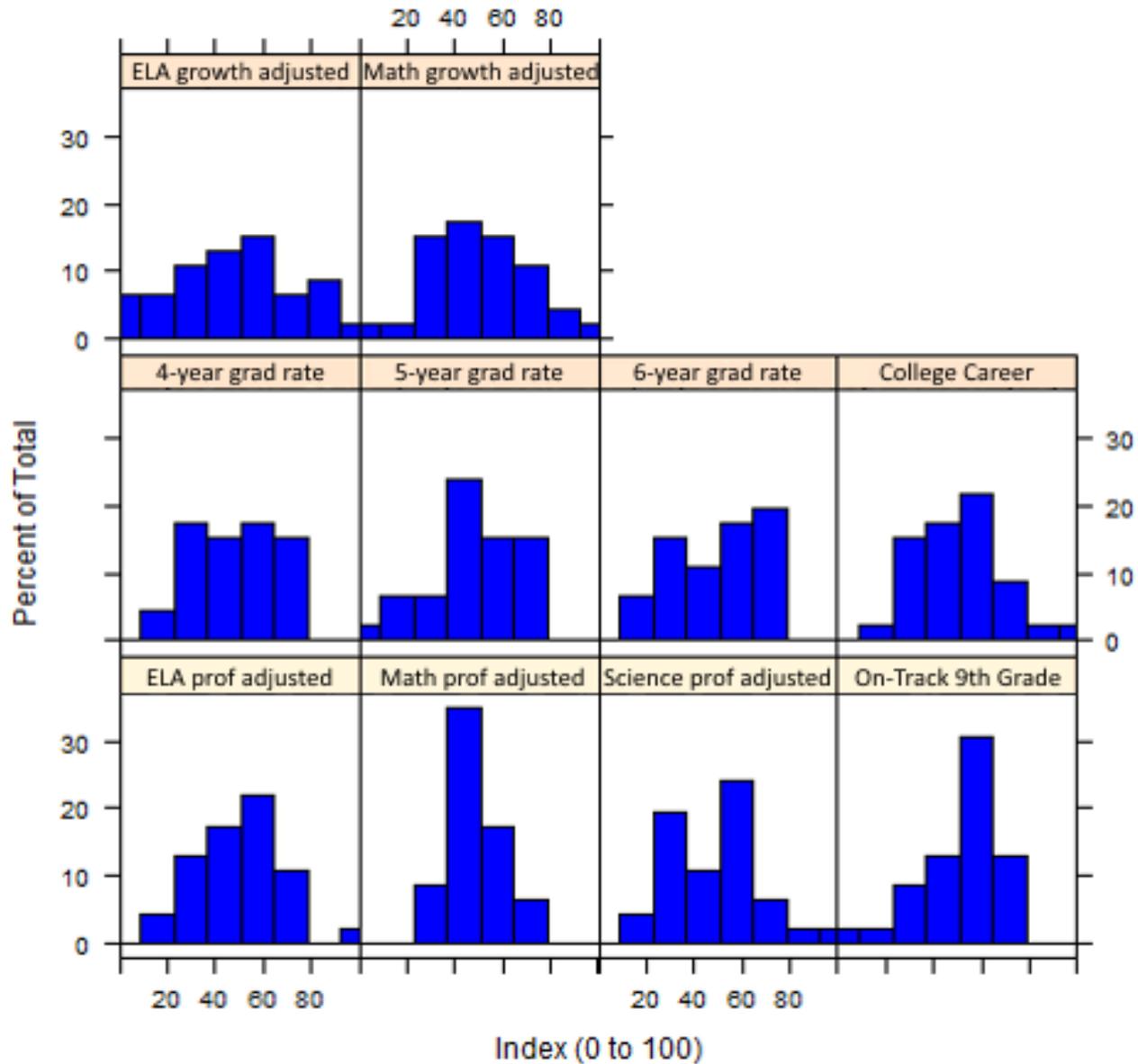
Weights

- Policy weights represent system values
- Numerical weights may not end up representing policy weights
- Caution: well-intentioned specifications may end up with wildly mis-weighted components
- Seek technical assistance (internal or external) on statistical weights regardless of system type

Non-standardized HS Measures



Standardized HS Measures



Communication of Designations

1. Numbers

- Traditional: 0 – 100, 1 – 5
- Nontraditional: 0 – 150, 1 – 4, GPA

2. Words

- State determined language (ex. below expectations, meet expectations)
- Federal categories (ex. comprehensive support, reward)

3. Letter grades

- A – F

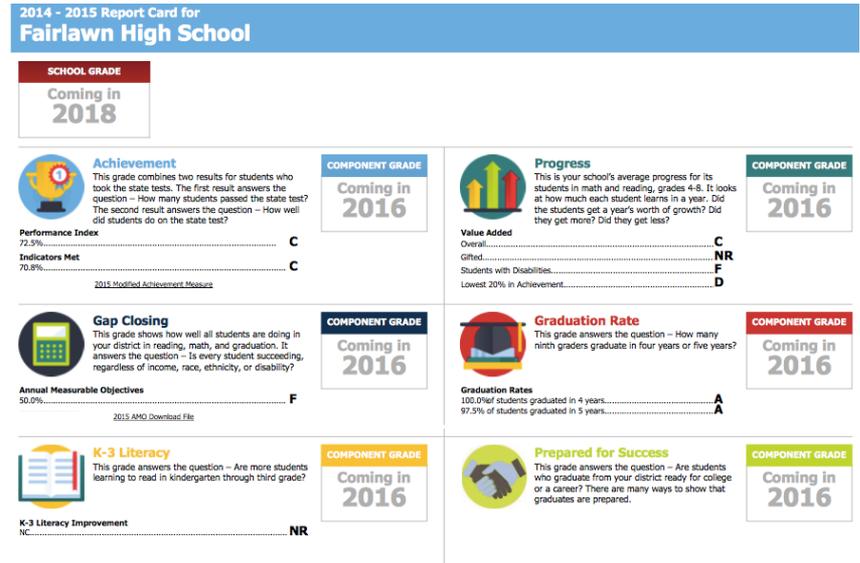
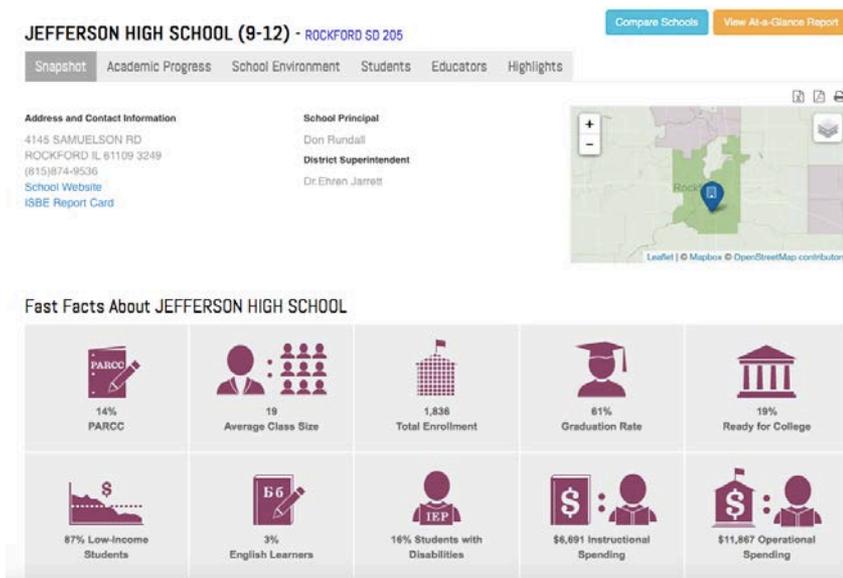
4. Symbols

- 5 stars

5. Colors

- Red, Yellow, Green

Reporting Does Not Stop at Overall Designations



- ⌘ Need to provide classifications for each set of measures
- ⌘ Additional context information is both required and advisable
- ⌘ Examples of interactive dashboards:
 - Illinois School Report Card
 - Ohio School report Card
 - LearnDC

Make Accountability Data Actionable for Educators

District Information | **Academic Dashboard** | **Dashboard Usage** | **Early Learning Dashboard**

District Overview | Attendance and Discipline | Assessments | Grades and Credits | Advanced Academics | College and Career Readiness

METRIC CATEGORY	PERFORMANCE SUMMARY	METRIC STATUS
Attendance and Discipline Students' attendance and discipline patterns	7 of 8 metrics met goal	
<ul style="list-style-type: none"> Attendance Daily and class period attendance Discipline Discipline incidents and actions 	2 of 3 metrics met goal	
Assessments State and local examinations and assessments	4 metrics met goal	
<ul style="list-style-type: none"> State Standardized Assessments Performance and progress on state standardized tests Early Childhood Assessments Performance and progress on Early Learner Survey Early Childhood Assessment Participation Participation of K-2 students on Early Learner Survey Local Assessment Performance and progress on local benchmark assessments 	6 metrics met goal	
Grades and Credits Students' progression in coursework	4 metrics met goal	
<ul style="list-style-type: none"> Course Grades: Primary Performance and progress in subject areas Course Grades: Secondary Performance and progress in coursework Credits Student progress toward graduation 	8 of 9 metrics met goal	
Advanced Academics Advanced coursework opportunity and performance	6 of 10 metrics met goal	
<ul style="list-style-type: none"> Advanced Coursework Student opportunity and performance in advanced coursework 	6 of 10 metrics met goal	
College and Career Readiness Students' ability to succeed in higher education and the workforce	7 of 9 metrics met goal	
<ul style="list-style-type: none"> College Entrance Exams Student performance on college entrance exams 	7 of 9 metrics met goal	

CREDITS	STATUS	VALUE	SCHOOL GOAL	△	DETAILS
Credit Accumulation (Through Year Round 2016)					
% of students meeting required # of credits for Recommended Graduation Plan as of the end of the prior school year					
10th Grade	99.1%	349 of 352	80.0%	19.1%	More
11th Grade	96.8%	339 of 350	80.0%	16.8%	More
12th Grade	96.4%	327 of 339	90.0%	6.4%	More

On-Track to Graduate (Through Year Round 2016)					
% of students who have earned at least 1 credit in each of the four core subject areas at the end of the prior school year					
10th Grade	91.4%	322 of 352	80.0%	11.4%	More
11th Grade	86.0%	301 of 350	80.0%	6.0%	More
12th Grade	77.5%	263 of 339	90.0%	-12.5%	More

ADVANCED COURSE POTENTIAL AND CAPACITY	STATUS	VALUE	TREND	STATE AVG.	△	DETAILS
Advanced Course Potential						
% of students that have potential and are enrolled in an advance placement course						
English	36.6%	94 of 264		52.0%	-16.4%	More
Math and Computer Science	25.8%	39 of 151		58.0%	-32.2%	More
Sciences	28.3%	41 of 145		51.0%	-22.7%	More

Advanced Course Potential - Sciences - Student List

↓ CUSTOMIZE VIEW

STUDENT ^	GRADE LEVEL	ADVANCED COURSE POTENTIAL	ATTENDANCE / DISCIPLINE		ASSESSMENTS		
			Last 4 Wks Attendance	School Violations	SAT Composite	PSAT Composite	ELA SBA
# Students Meeting Goal			98 of 104	104 of 104	14 of 17	85 of 100	16 of 1
ARMSTRONG, CEDRIC	10th	Yes	95.0% ▲	0 ◀▶			
BEALS, TRACEY	10th	Yes	100.0% ▲	0 ◀▶		142	

Questions?

For additional questions or information, contact:

Ryan Reyna

rreyna@edstrategy.org