



EmpowerK12 ESSA Accountability Design Proposal

EMPOWERK12 OVERVIEW

EmpowerK12 was born in 2013 as a data support organization for DC charter schools. We now work in multiple states and have a broader scope that includes edudata advocacy and professional development, data warehousing and report creation for schools, and builder of district data collaboratives (see [here](#) and [here](#).)

Josh Boots, EmpowerK12 Executive Director, played a key role as a charter data leader during the DC Public School Board's development of the nationally-recognized [Performance Management Framework](#) (PMF), a school quality index system with multiple performance indicators. While we feel the PMF is a terrific example of a data-focused next generation accountability system, incorporating important academic and non-academic indicators, our proposal for an ESSA-generation accountability system envisions an advanced accountability model truly driven by data that matters.

OUR DESIGN OBJECTIVES

Initial NCLB state accountability plans and those approved by ED through the waiver process often opted for simplistic measures in an effort to increase transparency at the expense of accuracy. As our society grows more accustomed to big data, trust in advanced analytics' ability to garner the best results has increased. EmpowerK12 proposes an educational accountability system supported by advanced statistical modelling in order to improve the match between school needs and state/LEA differentiated accountability.

The two main objectives for our ESSA accountability design include:

- A focus on high expectations for student growth in math and reading for all levels of achievers and all demographics, including analysis of growth gaps; and
- Using a variety of data points, we propose a comprehensive statistical model that estimates the probability a school can achieve greatness in the future.

OUR PROPOSED ACCOUNTABILITY SYSTEM

First, What is a Great School?

In order to fully understand our recommended accountability design, we begin by defining what we believe constitutes a great school. Great schools demonstrate the following characteristics (in descending order of importance):

- ▮ High math and reading growth for students across the achievement spectrum;
- ▮ Above national average percentages of students meeting or exceeding expectations in all subject areas tested, no matter their students' backgrounds;
- ▮ Little to no gap in achievement or growth in math and reading across all demographics, including special education and English language learners; and
- ▮ Foster safe learning environments with few behavioral incidents.

Our accountability model design includes a school-level composite index score encompassing these four elements and a probability of future greatness based on recent index score outcomes and other non-academic factors which contribute to success.

Indicators of Academic Achievement

Nationally, the next generation math and reading assessments are raising the proficiency bar. In fact, the bar is so much higher that many educators are moving away from the word proficiency altogether. Large percentages of schools have less than 20% of students meeting the new, higher benchmarks on the PARCC and Smarter Balanced assessments, and significant proportions of the student population perform well below benchmark.

If accountability systems merely focus on students meeting college readiness benchmarks, we worry schools looking for quick wins may utilize their resources to move average students across the proficiency line to the potential detriment of lower performing students. Therefore, our proposal recommends half credit for students approaching expectations. While it is likely unreasonable to expect 100% of students to be college ready right now, we think it is entirely appropriate to expect all students to be at least approaching the college readiness standard.

The EmpowerK12 design proposes separate indices for achievement and achievement gaps aggregated into an overall school index score. Below is an example of achievement index scores based on actual 2015 PARCC statewide results for the entire District of Columbia. The percentages below the index values show how much that result contributes to the overall school index score.

Overall Achievement Index			Achievement Gap Index							
Math	ELA	Other	Math Gaps				ELA Gaps			
All	All	All	FARMS	Race/Ethn. [^]	Special Ed*	ELL*	FARMS	Race/Ethn. [^]	Special Ed*	ELL*
35	36	N/A	43	40	31	80	39	37	27	61
8%	8%	4%**	1.25%	1.25%	1.25%	1.25%	1.25%	1.25%	1.25%	1.25%
35.5			44.6							

* Students who have exited special ed and ELL services remain in that subgroup calculation for two years

[^] Race/Ethnicity gap score is average gap score for ethnicity (Hispanic vs. Non-Hispanic) and race (highest performing race vs. all others, typically white vs. non-white)

** The weighting of science and social studies assessments would be based on number of valid tests compared to math/ELA, 4% is an example where half the students take another test

The math and ELA achievement index scores are the sum of the percent of students achieving Level 4+ on PARCC and one-half the percent of students achieving Level 3 Approaching Expectations. The Achievement Gap Index is the percent gap between math and ELA achievement index scores of subgroups. For example, in math, the achievement index score for non-ELL's is 50 and for ELL's is 40 with the resulting gap index score being $40 \div 50 \times 100$ or 80

Schools with an overall achievement index or achievement gap index score more than two standard deviations below the state school baseline-year mean would be automatically flagged for accountability purposes. Any school not already flagged but with any individual achievement gap index score more than two standard deviations below state baseline-year mean (e.g. a special education math gap score in the bottom two percent of all schools) would be marked for targeted assistance.

Indicators of Academic Growth

Truly amazing schools impact student growth across all levels of prior achievement. Several states and independent school rating systems have chosen growth models which include median growth percentiles (MGP). In general, we think MGP is decent measure. However, we worry about the reliability of growth percentiles at the tails of achievement as well as their tendency to be annually-normed measures which contribute to a competitive instead of collaborative environment amongst schools. Annually-recalculated growth percentiles mean results remain the same even if overall student growth improved over the prior year.

Our proposal for an academic growth index score will analyze the percent of students who exceed the 75th percentile of baseline growth of students who performed in the same decile. By focusing on decile bands of performance, we hope to reduce some of the issues related to floor and ceiling effects. Schools are credited for improving growth over time by setting a baseline growth standard that lasts at least three years.

So how does this proposal work in practice? Let's examine a sample student and how they would count towards the school's academic growth score:

Sarah is a 4th grader who earned a PARCC score of 678 in math as a 3rd grader. When the state calculated baseline year data, a 678 places Sarah in the second decile, and in the baseline dataset, the top 25% of second decile students gained 12 scale score points in 4th grade. Therefore, if Sarah earns a 690 or higher on her 4th grade math test, she will positively count towards her school's academic growth index score.

For growth, we calculate both overall growth by subject as well as growth gaps similar to the achievement gap index analysis. Overall, growth counts twice as much as achievement in the final school index score. The next table is an example of growth indices. The Overall Growth Index section indicates the percent of students with growth above the 75th percentile by subject,

and the higher Growth Gap Index score, the less of a growth gap between subgroups (a 100 here illustrates no gap.)

Overall Growth Index		Growth Gap Index							
Math	ELA	Math Gaps				ELA Gaps			
All	All	FARMS	Race/Ethn.^	Special Ed*	ELL*	FARMS	Race/Ethn.^	Special Ed*	ELL*
42	31	78	67	45	97	88	86	23	100
20%	20%	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%
36.5		73.0							

* Students who have exited special ed and ELL services remain in that subgroup calculation for two years

^ Race/Ethnicity gap score is average gap score for ethnicity (Hispanic vs. Non-Hispanic) and race (highest performing race vs. all others, typically white vs. non-white)

Similar to achievement, schools with an overall growth index score two standard deviations below the state mean would be flagged for accountability, and schools with an individual growth gap two deviations below will receive targeted assistance for that subgroup.

Indicators of Progress towards English Language Proficiency

We include a measure for English language proficiency progress because we must, but it is weighted low because most English acquisition assessments we know are not great for measuring growth. Our preference for this metric is a growth measure closely aligned to the math and ELA methodology, meaning a calculation where students positively count when they attain proficiency or exceed growth at the 75th percentile.

Indicators of Student Success or School Quality

This new requirement of state accountability systems is sure to open up a whole world of creative metrics. We expect many meaningful metric ideas will be presented to state boards, but it is likely only a small fraction of them can be implemented with the level of validity and reliability required for a high stakes system. In DC, where school choice is a big deal across the whole state, the state board is expected to strongly consider re-enrollment rates as an indicator of school quality. In larger states, student surveys, parent surveys, and teacher observations may be considered, but we fear their affordability and whether results will pass the high stakes readiness test.

Our guess is states will likely default to a measure based on data already collected from school districts, say behavioral incident and suspension data as an indicator of school climate, for example. Schools report in-school suspensions, out-of-school suspensions, and expulsions through the Civil Rights Data Collection as well as discipline incidents in EDFacts. You could create some fancy statistic that weighs incidents, expulsions, and instructional days missed due to suspension in a larger behavioral index score. However, for our design, we choose to focus on

the percent of students who served one or more days for out-of-school suspension during the school year. Why? Well, we are not convinced that going fancier adds any tangible value to the metric.

To calculate the school climate quality index score based on suspension rate, we start by finding the 10th and 90th percentile of school rates for the baseline year and calculate the index score:

$$\frac{90th\ percentile - \alpha}{(90th - 10th\ percentile)} \text{ where } \alpha = \text{school suspension rate}$$

A school with a suspension rate of 3.9% in a state where the 10th percentile is 0.7% and 90th percentile is 14.1% would have a behavior index score of 76.1.

Cumulative Summary School Grades

Our school accountability index design includes the following elements and relative weights in determining the overall index score:

- ▮ Student growth in math and reading (40%)
- ▮ Gaps in levels of growth amongst subgroups (20%)
- ▮ Achievement in math, reading, and science/social studies (20%)
- ▮ Subgroup achievement gaps in math and reading (10%)
- ▮ Suspension rate as indicator of school climate quality (5%)
- ▮ English language proficiency progress (5%)

Below is a sample overall scorecard based on the results from prior sections.

Overall School Index					
Growth	Growth Gap	Achievement	Achievement Gap	Eng Lang Prof Growth	Behavior
36.5	73.0	35.5	44.6	59.2	76.1
40%	20%	20%	10%	5%	5%
47.5					

Remember when we discussed the elements of what makes a “great school”? Well, here is what a great school would look like translated into an index score:

GREAT School Index Scorecard					
Growth	Growth Gap	Achievement	Achievement Gap	Eng Lang Prof Growth	Behavior
50	90	76	80	80	95
40%	20%	20%	10%	5%	5%
70.0					

This prototypical great school has half of their students demonstrating growth in the top quartile of all students. Their achievement scores are double the statewide average, and their gaps are small. The great school only had a few behavioral issues that required out-of-school suspensions, and English language learners are demonstrating appropriate growth.

Using Index Scores and Non-academic Data to Project Future School Performance

The innovative component of our accountability design is not our use of a comprehensive index score (well, maybe a little bit “innovative” because of an emphasis on growth and growth gaps) rather that we propose to incorporate a whole slew of additional school-level data to project the probability a school can be a great school within the next three years. The most obvious factors in our advanced statistical model would be current composite index score, prior year index scores, and changes to index subscores. We also propose, using historical non-academic data to develop the model, the following variables (as examples, not an exhaustive list) be analyzed for significance and included if they help reduce the inherent entropy in identifying great schools:

- Average years of staff experience;
- Rate of staff turnover;
- School leadership changes and tenure;
- Student re-enrollment;
- Per pupil spending rates;
- Number of instructional days on the school calendar;
- Presence of extended school day academic activities;
- Reported teacher quality indicators; and
- Student population demographics.

Each school would receive a composite probability that estimates the likelihood it will meet great school criteria within the next three years, and the state can use the model’s data to drive decisions about eligibility for comprehensive support and improvement services.

Schools with high probabilities of being great schools, which are likely already at or near 70 index points, would be considered “reward schools” and eligible for dissemination grants. Schools with lower probabilities, perhaps those with less than a 25% chance of future greatness, may be eligible for certain intervention opportunities, and depending on the greatest factors contributing to their lower probability could be asked to submit action plans targeting those specific factors. In some states and jurisdictions, schools with extremely low probabilities may be those to enter the school turnaround or revocation process.

EmpowerK12 believes this “big data” approach to projecting future school success will result in improved use of school improvement funds by targeting schools that truly need the additional support. States can also better differentiate among low performing schools, requiring districts to submit more aggressive action plans for schools with low growth and little current potential for improvement.