5-2017

# Allocation of Federal Title IA Funds Under ESSA: Issues and Recommendations

Amy Johnson
*University of Southern Maine*

Janet C. Fairman Ph.D.
*University of Southern Maine*

# Allocation of Federal Title IA Funds Under ESSA: Issues and Recommendations

Amy Johnson
University of Southern Maine

Janet Fairman
University of Maine

Maine Education Policy Research Institute

May 2017

Maine Education Policy Research Institute
McLellan House, 140 School Street
Gorham, ME 04038
207.780.5044 or 1.888.800.5044

Maine Education Policy Research Institute
5766 Shibles Hall, Room 314
Orono, ME 04469-5766
207.581.2475

Published by the Maine Education Policy Research Institute in the Center for Education Policy, Applied Research, and Evaluation (CEPARE) in the School of Education and Human Development, University of Southern Maine, and the Maine Education Policy Research Institute in the School of Education and Human Development at the University of Maine at Orono.

The Maine Education Policy Research Institute (MEPRI), is jointly funded by the Maine State Legislature and the University of Maine System. This institute was established to conduct studies on Maine education policy and the Maine public education system for the Maine Legislature.



This study was funded by the Maine State Legislature, and the University of Maine System.

# Allocation of Federal Title IA Funds Under ESSA: Issues and Recommendations

Amy Johnson
*amyj@maine.edu*

Janet Fairman
*Janet.fairman@maine.edu*

## Table of Contents

# Allocation of Federal Title IA Funds Under ESSA: Issues and Recommendations

Amy Johnson
*amyj@maine.edu*

Janet Fairman
*Janet.fairman@maine.edu*

## Executive Summary

With the passage of the Every Student Succeeds Act in 2015, each state that receives federal funding has the opportunity and responsibility to develop systems for holding schools accountable for student learning outcomes. One mechanism is annual public reports for every school district, which will provide an opportunity for parents, school board members, and the general public to compare their schools against state averages on a variety of data measures. Additional accountability rests with a state-developed set of learning goals for all students, and an accompanying system of evaluating whether schools are meeting those goals.

Working within the timeline of the federal requirements, Maine was challenged to develop a robust system of learning goals and interim progress measures because the state recently adopted new standardized assessments. The initial state plan was developed with only one year of assessment data upon which to base expectations for school outcomes. In addition, the initial plan is limited to using data measures that are currently available and comparable for each individual student in the state; data that are only available to the state in aggregate, or are not reported from school districts to the state Department of Education, could not be considered because they cannot be analyzed by student subgroups. Thus the state has a plan for continuing to review and revise its accountability measures in the coming years as more data and research become available. This report highlights some issues the Department may wish to consider in those future analyses and discussions.

In reviewing historical assessment data (based on prior state tests) and analyzing current school enrollments, MEPRI researchers identified several challenges for building an accountability system for Maine:

- Small schools may have less reliable data, or even missing data, that make statistical analyses more challenging. Maine has many small schools. They may fluctuate widely in their annual performance, making it more difficult to understand and

monitor their students' academic outcomes. This is especially problematic for evaluating performance of student subgroups.

- A number of elementary schools do not include any grades that participate in annual state assessments (e.g. schools with a grades K-2 configuration). These cannot be included in measures based on student proficiency rates or student growth. An additional number of elementary schools only have one tested grade (e.g. K-3 schools) and thus cannot calculate student growth scores in the current system.
- Maine has maintained its past practice of basing evaluating student achievement on the percent of students in a school that score at the "proficient" level or higher on state exams. This approach works well for small schools, but there are trade-offs for choosing this measure instead of other options, such as using average scaled student test scores.
- There is considerable measurement overlap between the student proficiency rate and student growth measures as defined in historical data. This means that schools with low proficiency rates also tended to have low growth rates, and the two measures were not an optimal combination for differentiating between schools. This is a factor to consider when refining the measures for achievement and growth in the future.

The report concludes with suggestions for options to test out when more years of current assessment data are available, in addition to other suggestions from the ESSA Advisory Group that are already planned for future consideration.

## Background

### Purpose & Methods

In January of 2017, the Maine Education Policy Research Institute provided a report to the Maine Legislature's Joint Standing Committee on Education and Cultural Affairs summarizing the issues and implications of the recent federal Every Student Succeeds Act (ESSA) for Maine policy. This Title IA Allocation report is intended as a companion document to delve more deeply into questions related to federal funding for states and districts under the Improving the Academic Achievement of the Disadvantaged formula grant program.

Maine Department of Education leaders and other stakeholders have engaged in ongoing planning and discussions to finalize an initial consolidated application for federal funds, including Title IA funds, as required by ESSA. Maine's application was submitted in April 2017. At the time of writing, it remains under review for completeness by U.S. Department of Education staff, and is not yet formally finalized. Thus this report serves a purpose of providing current information to policymakers while also identifying questions that may be revisited in coming years as Maine's approach to school improvement and accountability is studied and improved.

General background information provided in the report was compiled from federal statute, other available documentation, and Maine's consolidated application for federal funds. A list of data questions was developed in collaboration with leadership at the Maine Department of Education, and analyses were conducted on student enrollment and achievement data provided from Department staff or from public reports. Analytic methods are described in more detail within each data topic.

### Title I Basics

According to the U.S. Department of Education, "Title I, Part A (Title I) of the Elementary and Secondary Education Act (ESEA), as amended [by the Every Student Succeeds Act of 2015] provides financial assistance to local educational agencies (LEAs) and schools with high numbers or high percentages of children from low-income families

to help ensure that all children meet challenging state academic standards."[1] These federal funds are apportioned to states based on the most recent available census data estimates of the population of children living in poverty. They are in the category of "formula" grants rather than "discretionary" grants because they are allocated to all states and territories based on established objective criteria rather than by competition.

In the FY 2017 federal budget passed on May 1, 2017, Congress appropriated a total of $15.5B for Title IA allocations.[2] This was $550M more than the $14.9B amount provided in FY2016, and $100M more than the amount requested for FY2017 in the original administrative budget request. This 4% increase from FY2016 indicates that Maine should expect an allocation that is similar to its FY2016 amount of $50.1M, given that the criteria for allocating amounts to states remains similar to those used in the past.

*State allocation*

Maine's total state allocation will be computed based on the varying criteria set for four different Title IA fund categories: basic, concentration, targeted, and Education Finance Incentive Grants (EFIG). Basic and concentration grants are proportional to the *number* of students living in poverty, while targeted grants are proportional to the *rate* of poverty. Thus some of the funds flow in greater quantity to large districts with greater numbers of students, while others are focused more specifically on low-income communities. EFIG funds are allocated to favor states that have equitable funding formulas and provide high levels of education funding relative to per capita income. From year to year, states must meet a "maintenance of effort" provision in order to receive their full allocation; this requires spending at least 90% as much as the prior year, either on a total or per-student basis. State level allocation estimates for FY 2017 were not yet available at the time of report publication.

*Allocations to Districts (LEAs)*

Once Maine's total share is determined, the funds are divided amongst school districts and the state Department of Education. By law, at least 90% of the funds must be

[1] From https://www2.ed.gov/programs/titleiparta/index.html
[2]https://rules.house.gov/sites/republicans.rules.house.gov/files/115/OMNI/DIVISION%20H-%20LABORHHS%20SOM%20OCR%20FY17.pdf

directed to school districts (called Local Educational Agencies, or LEAs). District funding is determined by federally established formulas for each of the four categories listed above (basic, concentration, targeted, and Education Finance Incentive Grants), which are detailed in Sections 1124, 1124A, 1125, and 1125A of the ESEA (as amended by ESSA).

Each state must set aside 7% of its total allocation for school improvement activities to help struggling schools. Most of these funds (at least 95%) must be directed to school districts identified as needing additional support. This category of funds, including methods for identifying struggling schools, is discussed in detail in following sections of this report.

The remaining 3% of the total state award can be used either for a new option under ESSA known as Direct Student Services, or can be added to the 90% funneled to schools by formula. Maine has elected to pursue the latter option.

*Allocations to Schools*

LEAs must apply a combination of federal requirements and local priorities to determine how to divide the funds among their schools. Section 1113 [20 USC 6313] of ESEA details the criteria districts must use in determining which of its member schools may receive Title IA funding. In general, only schools with poverty rates above the district average poverty rate are eligible for funds, and schools with the highest percentages of low-income students must receive priority for funds before those with lower poverty rates. There are specific provisions for expanded options for funds in schools with 40% or more of students identified as low-income (detailed in ESEA section 1114 [20 USC 6314]), which are eligible for schoolwide programs. Schools that have lower poverty rates or that do not choose to pursue schoolwide services can use targeted assistance funds to implement more narrowly focused programs, as dictated by ESEA Section 1115 [20 USC 6315].

## Title IA Accountability Considerations

The federal law that provides for Title IA funding for public schools contains two mechanisms for holding schools accountable for student learning. The first is a requirement that states have a system for identifying schools that are struggling to achieve state-established target levels for student outcomes. The second is a mandatory annual report card that must be released publicly about each grade in each district that receives Title I funding.

**Identification of Low-Performing Schools**

In the past, the federal ESEA law dictated what the identification and support systems looked like for underperforming schools. The criteria and methods for measuring and comparing school performance were based largely on standardized test scores, and states had limited options for providing interventions to help schools improve. Under the new ESSA legislation, states have been provided with more flexibility in both how to identify struggling schools and how to provide them with meaningful and effective supports. This presents an opportunity for Maine to build an accountability system that is suitable to its unique context and is aligned to state priorities and ongoing initiatives.

Evaluation systems can serve a variety of different purposes. For example, a school rating system could be intended to identify best-practice schools to serve as exemplars or to locate schools that may benefit from a certain focused grant program. The rating system that states are required to develop under Title IA (Section 1111(c), Statewide Accountability System) must, at a *minimum*, serve the purpose of identifying schools that are eligible to receive funding and/or state support from the 7% set-aside for school improvement activities. The system must identify 1) "*comprehensive support and improvement*" schools as the lowest-performing 5% of all schools and high schools with graduation rates below 67%, and 2) "*targeted support*" schools as those with a consistently underperforming subgroup(s) of students. Maine's proposed system for measuring attainment goals for all students and subgroups of students is summarized in Table 1.

Working within the timeline of the federal requirements, Maine was challenged to develop a robust system of learning goals and interim progress measures that would allow identification of comprehensive and targeted support schools because the state recently adopted new standardized assessments. The initial state plan was developed based on historical (prior NECAP assessment) data since only one year of the current assessment data were available. It was not the preferred approach to base expectations for school outcomes on legacy data. In addition, the initial plan was limited to using data measures that are currently available and comparable for each individual student in the state; data that are only available to the state in aggregate, or are not reported from school districts to the state Department of Education, could not be considered because they cannot be

analyzed by student subgroups. The Department of Education initially requested to be given until June 30, 2017 to develop its measurement criteria and weights for the system so that it could test the model using two years of current Maine assessment data. However, that request was not accepted by the U.S. Department of Education (USDE). The Maine Department of Education application had to be finalized based on the data available prior to 2016-17. Thus the state has an initial plan currently under USDE review, and intends to review and revise its accountability measures in the coming years as more data and research become available.

**Table 1: Maine Proposed Accountability System Indicators**

| ESEA Requirement: Section 1111(c)(4)(B) | Maine Measure | Weight for Elementary Schools | Weight for High Schools |
|---|---|---|---|
| (i) An indicator of student academic achievement (I) as measured by proficiency on annual assessments | Proficiency rate as measured on annual statewide assessments in English language arts and mathematics, using eMPowerME (Grades 3–8) and SAT (Grade 11). | 42% * | 40% |
| (ii) For public elementary schools, (I) a measure of student growth, if determined appropriate by the State, and | Progress as measured on the annual statewide assessments in English language arts, mathematics, using eMPowerME for Grades 4–8 | 38% * | -- |
| (iii) For public high schools, (I) the four-year adjusted cohort graduation rate, and, at the State's discretion, (II) the extended-year adjusted cohort graduation rate | Adjusted cohort graduation rates (four-year rate, as well as five- and six-year rates) | -- | 40% |
| (iv) Progress in achieving English proficiency for English Learners (ELs) within a state-determined timeline | English Learner Progress (ACCESS for ELLs 2.0 assessment) | 10% | 10% |
| (v) At least one additional statewide measure of school quality or student success | K-12: Consistent attendance | 10% | 10% |

**See also Appendix A for additional detail about measure weights*

Each state has considerable discretion to tailor its parameters. Maine can choose to include additional criteria or to use the above measures in such a way that would also

allow other functions beyond identification of comprehensive and targeted support schools, such as recognition of improving or high-performing schools.

In addition to revising the above model based on additional years of testing data, the Maine Department of Education will also continue to investigate other data measures that were recommended by the ESSA Advisory Group. These include but are not limited to:

- College and career readiness data points for high schools such as participation in advanced coursework, CTE, and/or assessments such as Accuplacer or ASVAB
- Inclusion of science assessment data
- Non-academic measures of school climate such as disciplinary actions
- Student feedback obtained by survey to measure engagement, school climate, and/or other social or emotional issues.

These measures are not currently available to the Department of Education, and if adopted into the accountability system would need to be implemented statewide in a way that would allow disaggregation by student subgroups. Thus these are longer-term considerations that are not likely to able to be fully implemented before the first comprehensive improvement schools are identified for AY 2018-19.

Discussion of the components of Maine's proposed accountability system, and options for consideration in future iterations, is the major focus of the remaining sections of this report.

**Annual report cards**

The required annual reports depicting each district as well as the state's overall performance are often overlooked as accountability mechanisms. Yet the inclusion of multiple mandatory data elements in these report cards will ensure that there is public transparency about multiple factors impacting public schools' contexts as well as their students' performance. As reported in the companion ESSA report (MEPRI, 2017), "The components to be included in state and district report cards are:

- Details of the state accountability system, including goals, indicators, weights of indicators, and schools identified for Comprehensive Support and Improvement and Targeted Support and Improvement.
- Disaggregated results on all accountability indicators, such as achievement on reading/language arts, mathematics and science on state assessments and graduation

rates. (This includes subgroups including homeless students, students in foster care, and students with a parent who is a member of the Armed Forces.)
- Disaggregated assessment participation rates.
- Information on the acquisition of English language proficiency.
- The state's threshold for the minimum number of students (n-size) necessary to be included in subgroup reporting.
- Disaggregated results on the indicators that the state and its districts are already reporting to the Civil Rights Data Collection, including, but not limited to: access to advanced coursework, such as Advanced Placement (AP), International Baccalaureate (IB), and dual enrollment; exclusionary discipline rates; and chronic absenteeism.
- The professional qualifications of educators including the number and percentage of inexperienced teachers, principals, and other school leaders as well as teachers with emergency or provisional credentials and teachers who are not in a subject or field for which they are certified.
- State, local and federal per pupil expenditures, including actual personnel and non-personnel expenditures.
- The number and percentage of students with the most significant cognitive disabilities taking the alternate assessment.
- At the state level, results of the National Assessment of Educational Progress (NAEP), as compared with national averages (state report card only).
- Where available, for each high school, the disaggregated rates of students who graduate from high school and enroll in higher education.
- Other information as determined by the state or district

Additional details about the mandatory data elements in these report cards are described in ESEA Section 1111(h)(1) and 111(h)(2).

## Criteria for Evaluating School Performance

When developing accountability systems, states must consider some basic principles. The following foundational criteria were provided by the Maine Department of Education to Maine's ESSA Advisory Group, which was assembled to provide stakeholder input into the state-developed plan required by ESSA, at the outset of its work in fall 2016:

- Data measures must be valid, reliable, and research-based.
- Indicators should be easily understood.
- Indicators should be difficult to manipulate or corrupt.

It was also noted by Department leadership that the combination of criteria used to identify schools for comprehensive or targeted supports are not necessarily the same as those needed to further identify the "root causes" of school underperformance and thus the

types of supports that would be most appropriate and effective. In other words, the accountability metrics do not need to answer all the questions one might want to know about school performance; other data sources, including the annual report cards, can serve that purpose. Rather, the identification system needs to be able to discern which schools are struggling to achieve adequate results for students, so that additional analysis and assistance can be provided.

**Data Validity and Reliability**

To ensure that an accountability system can accurately identify schools that are in need of additional resources, it must be based on data that are both valid and reliable. The principle of validity means that a data indicator captures what it purports to measure. Reliability means that a measure produces predictable and stable results without large variation or error.

To illustrate these principles one can consider the data measure of attendance. It is routine practice for schools to capture student attendance at the start of each school day. Schools typically have late arrivals each day, for various students and various reasons. Some schools have a regular practice of updating their records to distinguish between absent students and tardy students. Others do not have a robust update process, and their absentee data may include students who were merely ten minutes late. In secondary schools, it is increasingly common for schools to capture attendance for each class period rather than once per day during homeroom. Each of these methods for measuring attendance could produce different rates. Depending on how the data are to be used and interpreted, the different methods would have varying validity to measure the construct of "attendance". To illustrate the idea of reliability, one can compare the practice of a school that measures attendance daily to a hypothetical school where teachers are asked to complete attendance rosters each Friday based on their memory of who had attended each day that week. The latter environment would likely have problems with reliability, as different teachers would have different systems for remembering, and some would be more accurate than others. When data are to be used for accountability purposes—with resulting high stakes decisions involving school funding and possible interventions—it is critical that they meet both tests of validity and reliability.

**Data Challenges of Small Schools**

An issue that greatly impacts Maine's ability to develop a robust accountability system is the prevalence of small schools. When a school is represented by a small number of students in any grade (or subgroup), its data is more likely to fluctuate. This can arise from the influence of a single student on average scores, and also from data variation from year to year. The general rule of thumb when comparing groups using statistical analysis is to strive for 20 to 25 individuals to represent each group, so that the influence of outliers can be mitigated and less likely to unduly affect a group's average.

When small numbers of students are represented in data, there is an additional challenge of protecting student privacy. It may be possible to guess the identity of a student if small numbers of students are aggregated into public data reports. Thus it is standard practice to establish a minimum "N-size" for the number of students (data points) that must be available for data to be presented in a public report. Different guidelines are used for different types of data; the most commonly used minimum N-sizes are 5 or 10. Maine has elected to use a minimum N of 10 in all data reports related to accountability. This means that if there are fewer than 10 students comprising a subgroup for any data measure for either the school improvement identification system or the annual reports, the data point will be suppressed (i.e. replaced with a * to indicate inadequate N size).

The proposed system for rating schools combines assessment results for all tested students in a school into a single proficiency rate. This greatly reduces issues related to data suppression as students from multiple grades can be combined together in reporting and calculations. However, in prior reporting requirements schools and districts provided combined student performance levels on state exams by grade level (i.e. for each different state examination). If grade-level reporting is maintained on annual public reports, schools that have fewer than 10 test-takers in a grade will have challenges because that data would be suppressed. This would be the case for a substantial proportion of Maine schools. Table 2 provides an overview of the varying grade configurations in Maine schools based on Fall 2016 attending enrollment data. The "Number with <10 in a tested grade" column is an indication of schools that would potentially be lacking *grade-level* data for at least one tested grade if that were to be required reporting.

In addition, Maine also has 37 elementary schools that do not include an annual statewide test because of their grade configurations (i.e. do not include any grades in the 3-8 range or grade 11), and thus would be unable to be rated on the academic performance or student growth indicators (combined 80% of the accountability metric). An additional 25 include only one tested grade, and thus cannot evaluate student growth using annual summative state assessments (38% of the accountability metric for elementary and middle schools). The "Number unable to calculate growth" column in Table 2 depicts the schools that would not be able to calculate student growth because their configurations do not include two tested grades.

**Table 2. Maine School Challenges with N-Size and Number of Tested Grades**

| Number of Tested Grades | Grade Configurations | Number of Schools | Number with <10 Students in a Tested Grade | Number Unable to Calculate Growth |
|---|---|---|---|---|
| **Elementary & K-8** | | | | |
| ZERO | K-2, K-1, K only | 37 | 37 | 37 |
| ONE | K-3, 1-3, 2-3 | 25 | 3 | 25 |
| TWO | K-4, 1-4, 2-4, 3-4, 4-5 | 38 | 0 | 0 |
| THREE | K-5, 1-5, 2-5, 3-5 | 116 | 13 | 0 |
| FOUR or FIVE | K-6, 3-6, K-7 | 61 | 8 | 0 |
| SIX | K-8, 1-8, 3-8 | 86 | 38 | 0 |
| ***Subtotal*** | | 363 | 99 (27%) | 62 (17%) |
| **Middle** | | | | |
| TWO | 6-7, 7-8 | 9 | 0 | 0 |
| THREE | 6-8 | 53 | 0 | 0 |
| FOUR or more | 5-8, 3-8, 4-8 | 27 | 0 | 0 |
| | | | | |
| **K-12** | K-12 | 10 | 9 | 0 |
| **Secondary** | | | | |
| ONE | 9-12 | 104 | 0 | N/A |
| > ONE | 5-12, 6-12, 7-12, 8-12 | 22 | 1 | N/A |
| **TOTAL** | -- | 587 | 108 (18%) | 62 (10%) |

* *Excludes four pre-K only schools*.

An additional 71 schools (12%) have between 10 and 20 students in at least one tested grade. Combined with the 108 schools with fewer than 10 in a grade, the 178 total schools with at least one grade with fewer than 20 students may experience challenges

with growth model calculations because of possible instability in scores when dealing with small numbers of cases.

The effect of small schools on data reliability is even more pronounced when looking at subgroups. While demographic data for the current school year were not yet available for this report, prior years give an indication of the data challenges that arise with small numbers. In historical data from the 2014 testing year, comparatively few schools had 10 or more students in some of the required subgroup categories. Table 3 shows the number of schools with 10 or more students in various student subgroups, out of 433 total schools administering state tests that year.

**Table 3. Student Subgroup Counts, 2014 Historical Data**

| Student Subgroup | Number of Schools with 10 or more Students in Subgroup (Total N=433) |
|---|---|
| White | 403 |
| Black | 20 |
| Hispanic | 10 |
| Asian | 10 |
| Native American | 6 |
| Multi-racial | 5 |
| Special Education | 267 |
| Economically Disadvantaged | 385 |
| Limited English Proficient | 27 |
| Migrant Education | 0 |

This points to potential difficulties in identifying underperformance of student subgroups other than the economically disadvantaged and special education students. Notably, the small numbers of schools with 10 or more English Learners (Limited English Proficient) means that the large majority of Maine schools would not have sufficient data to be evaluated on their performance, a metric that is weighted 10% in the accountability model.

Notably, there are two factors that can adversely impact N sizes and result in increased data suppression. First, the number of students tested in a given year is not identical to school enrollment in tested grades. Due to student absenteeism and parental choice to opt their children out of testing, exam participation rates are rarely 100%. Participation rate is an important factor that varies significantly among schools and is not

entirely within school control. Also, students who do participate in testing are only included in a school's data calculations if they are present for at least half of the school year. This means that schools with high student mobility rates may have more problems with data suppression and small N sizes in their rating calculations.

Lastly, the federal ESEA policy allows districts to exempt small high schools (defined as those enrolling fewer than 100 students) from implementing comprehensive support activities that would otherwise be required, per ESEA Sec. 1111 (d)(1)(C)(ii). In 2016-17, Maine has 21 high schools with fewer than 100 students enrolled in grades 9-12, out of 126 total schools that include grades 9-12 (including eleven K-12 schools plus other configurations other than strictly grades 9-12). This means that up to 17% of Maine secondary schools could possibly be exempt from school improvement accountability, depending on whether the statute can be interpreted to apply to all schools that incorporate grades 9 through 12.

**Comparing Schools with Different Grade Levels**

As described above, Maine's initial accountability plan for identifying comprehensive and targeted support schools uses a method of calculating proficiency rates for all tested students in a school. As Table 2 showed, this means that some schools would have proficiency rates calculated based on only one tested grade level, while others could have as many as six tested grades included in their overall rate. This solves a very real obstacle of having to exclude grades with fewer than 10 tested students if students were compared only to students in similar grades. However, the system's validity is built on the assumption that proficiency rates are equivalent at all grade levels. This has not been verified to be the case.

In prior years of Maine testing data, proficiency rates tended to decrease between grades 8 and 11. To illustrate, Table 4 provides the overall statewide proficiency levels in English Language Arts and mathematics for each tested grade in AY 2013-14. For comparison, data for AY 2015-16 on the first administration of the new eMPowerME assessment (grades 3-8) and the recently revised SAT exam (grade 11) are also provided. The new assessment data do not appear to have the same trend for English Language Arts, since more 11$^{th}$ grade students than lower grade students were rated as proficient.

However, data for English Language Arts for AY 2014 and both years of mathematics data show marked decreases in proficiency from grades 3-8 to grade 11.

**Table 4. Maine Proficiency Levels by Grade**

| | English Language Arts Proficiency rate | | Mathematics Proficiency Rate | |
|---|---|---|---|---|
| | AY 2014 | AY 2016 | AY 2014 | AY 2016 |
| Grade 3 | 68% | 48% | 60% | 49% |
| Grade 4 | 66% | 52% | 63% | 40% |
| Grade 5 | 71% | 52% | 63% | 36% |
| Grade 6 | 70% | 46% | 61% | 33% |
| Grade 7 | 69% | 48% | 59% | 40% |
| Grade 8 | 71% | 48% | 56% | 35% |
| Grade 11 | 48% | 60% | 49% | 35% |

Once a second year of data are available using the new assessments, it will be important to review proficiency levels by grade level to assess their comparability. If there is a demonstrable pattern of difference in proficiency rates at different grade levels, it may lead to questions about the validity of combining proficiency rates across all grade levels. For example, if 11th grade proficiency levels are generally lower than those in grades 3-5, then high schools will generally have lower proficiency rates than elementary schools. Comparing high school and elementary schools directly would lead to high schools being more likely than elementary schools to fall in the bottom 5% of schools as measured by student performance. Additional study with a second year of data will reveal whether this is the case. If it is determined that there is reason to question to comparability of proficiency rates across grade levels, then the state could pursue alternative methods of ranking schools for comprehensive and targeted support identification. For example, schools could be grouped with others having similar grade spans, and the lowest performing 5% in each group could be identified as needing comprehensive supports. Methods would need to be developed to treat schools that embody multiple grade spans, such as K-8 and K-12 schools.

## Metrics for Capturing Student Performance

*Proficiency rate vs. average scale score*

Under prior accountability policies, states were required to report student assessment results as the percent of students who scored at the proficient level or higher on state standardized tests. Maine has elected to retain this general reporting approach in its initial plan, using the percent of students that are proficient on statewide tests as its indicator of academic achievement.

The flexibility now afforded to states would allow Maine to use average test scores to evaluate student performance, rather than percent proficient. A group of prominent education researchers filed a statement supporting this practice during the public comment opportunity for draft ESSA regulations (Polikoff, 2016).[3] The following excerpt from that statement succinctly summarizes the reasons to question the use of proficiency rates: "Reporting performance in terms of the percentage above proficient is problematic in several important ways. Percent proficient:

1. Incentivizes schools to focus only on students around the proficiency cutoff rather than all students in a school (Booher-Jennings, 2005; Neal & Schanzenbach, 2010). This can divert resources from students who are at lower or higher points in the achievement distribution, some of whom may need as much or more support than students just around the proficiency cut score (Schwartz, Hamilton, Stecher, & Steele, 2011). This has been shown to influence which students in a state benefit (i.e., experience gains in their academic achievement) from accountability regulations (Neal & Schanzenbach, 2010).
2. Encourages teachers to focus on bringing students to a minimum level of proficiency rather than continuing to advance student learning to higher levels of performance beyond proficiency.
3. Is not a reliable measure of school performance. For example, percent proficient is an inappropriate measure of progress over time because changes in proficiency rates are unstable and measured with error (Ho, 2008; Linn, 2003). The percent proficient is also dependent upon the state-determined cut score for proficiency on annual assessments (Ho, 2008), which varies from state to state and over time. Percent proficient further depends on details of the testing program that shouldn't matter, such as the composition of the items on the state test or the type of method used to set performance standards. These problems are compounded in small schools or in subgroups that are small in size.
4. Is a very poor measure of performance gaps between subgroups, because percent

[3] Full text accessed 5/1/17 at https://morganpolikoff.com/2016/07/12/a-letter-to-the-u-s-department-of-education/

proficient will be affected by how a proficiency cut score on the state assessments is chosen (Ho, 2008; Holland, 2002). Indeed, prior research suggests that using percent proficient can even *reverse the sign* of changes in achievement gaps over time relative to if a more accurate method is used (Linn, 2007).

5. Penalizes schools that serve larger proportions of low-achieving students (Kober & Riddle, 2012) as schools are not given credit for improvements in performance other than the move to proficiency from not-proficient."

The letter further states, "The use of mean scores places the focus on improving the academic achievement of *all* students within a school and not just those whose performance is around the state proficiency cut score (Center for Education Policy, 2011). Such a practice also increases the amount of variation in school performance measures each year, providing for improved differentiation between schools that may have otherwise similar proficiency rates. In fact Ho (2008) argues if a single rating is going to be used for reporting on performance, it should be a measure of the average performance because such measures incorporate the value of every score (student) into the calculation and the average can be used for more advanced analyses. The measurement of gaps between key demographic groups of students, a key goal of the ESSA law, is dramatically improved with the use of average scores rather than the proportion of proficient students (Holland, 2002; Linn, 2007)." (Polikoff, 2016)

In addition to the above, the use of average scale score can be a benefit to states like Maine with numerous small schools, as individual student performance is easier to guess from proficiency rates (where one student may represent as much as 10% of the overall rate). Thus average scores make it easier to protect student confidentiality. When future changes are considered based on additional years of available assessment data using the eMPowerME test, we recommend that consideration be given to switching to average scores. The potential benefits of using average scores should first be verified based on analysis of actual Maine data.

*Growth vs. Proficiency*

Maine's system for measuring school performance, like that of many other states, will incorporate a measure of growth in student learning. The growth measure will be used in evaluating elementary and middle schools (using assessments in grades 3-8) and is 38%

of the overall school rating. (High schools will be evaluated based on graduation rates in place of student progress).

Details about the growth measure to be initially used were not available at the time of report development, as the Maine Department of Education had to finalize that definition in April 2017 rather than during June 2017 as they has initially planned. Therefore, MEPRI analysis of this concept was based on the definition of student growth used in the legacy accountability system from in AY 2011-12 and AY 2012-13. In that metric, student growth was based on changes in individual student proficiency levels; each student received "progress points" based on how their state assessment score compared to their score the prior year. Generally, students who did not score as proficient but who improved from the prior year earned points, and students who scored as proficient also earned points. Lower performing students that did not improve did not score points. An average was calculated based on the total points divided by the number of tested students in the school.

The student progress or growth measure is intended to capture different information than the student achievement measure. However, both the achievement and progress measures are to be calculated from the same state assessment data. This raises a question about whether the measures are providing meaningful differentiation between schools, or merely providing two different measures that are highly interrelated. To investigate this question, MEPRI researchers analyzed data from school ratings calculated in AY2012 and AY2013 that incorporated both an academic achievement measure and a student progress measure.

The Pearson's correlation between the reading proficiency rate and the reading growth score in Maine elementary schools was 0.68 in 2011-12 and 0.69 in 2012-13. The square of the correlation is about 0.46, indicating that 46% of the variance in these two measures is shared—i.e. the measures are capturing some of the same information about student performance. This is a moderately high correlation, indicating meaningful overlap between the two measures. In mathematics the picture is similar, with a correlation of 0.69 in math proficiency and math growth scores in AY2012 and 0.77 in AY 2013.

In the accountability system, the 5% of schools that perform lowest overall are flagged as needing comprehensive improvement supports. Table 5 provides another illustration of the overlap between schools identified as being in the bottom 5% on each

measure. Slightly more than half of the schools that were in the bottom 5% of elementary schools based on proficiency rates were also in the bottom 5% based on low student growth. About two thirds of low-proficiency rate schools could be found in the bottom 10% of schools based on growth. In practice, the use of both growth scores and proficiency rates was helpful to some schools, but not most.

**Table 5. Similarity in Rankings Between Proficiency and Growth Measures**

| | # Schools in bottom 5% based on proficiency | # of lowest proficiency schools also in bottom **5%** based on Growth | # of lowest proficiency schools also in bottom **10%** based on Growth |
|---|---|---|---|
| Reading, AY 2012 | 21 | 12 (57%) | 17 (81%) |
| Reading, AY 2013 | 20 | 8 (40%) | 12 (60%) |
| Mathematics, AY 2012 | 21 | 11 (52%) | 15 (71%) |
| Mathematics, AY 2013 | 20 | 11 (55%) | 14 (70%) |

To reiterate, the above analysis was based on a prior growth measure that may or may not match the metric that is ultimately proposed by the Maine Department of Education. However, the analysis does illustrate the question that may need to be asked when reviewing the effectiveness of the measure that is proposed.

The overlap between growth and achievement measures is not necessarily problematic for the integrity of the accountability system. The fact that there is some convergence in the lowest performing schools on these two student learning measures can be seen as validation that the schools identified are truly those most in need of state assistance. However, it is desirable to explore multiple methods of measuring student growth to identify those that have less overlap with the student proficiency measure, and thus may produce a more robust depiction of schools in need of support.

## Conclusions & Recommendations

This report describes several challenges for Maine policymakers in building an effective system for identifying schools to receive state assistance via comprehensive support or targeted support designations. As additional assessment data becomes available, the Maine Department of Education has the opportunity to conduct additional analyses and

consult with the ESSA Advisory Group to consider improvements to the state plan. The following recommendations are intended to guide these discussions.

*General*

While the school rating system garners the most attention, the potential impact of annual report cards should not be underestimated. The rating system for accountability and school improvement will ultimately identify about 10%-20% of Maine schools as needing either comprehensive supports or targeted supports. Yet 100% of schools will be included in annual report cards. The report cards also provide more data points for each school and thus can present more nuanced information about a broad range of school inputs, outputs, and student outcomes. The state should invest commensurate energy into developing a report card system that is easily accessible and provides clear data visuals to the public. This may also be a venue for requiring additional data beyond what is required in federal statute to address criteria that were of interest to stakeholders but unable to be included in the accountability rating system for any reason.

*Small Student Groups*

There are several options for mitigating the potential problems with unreliable or unavailable data arising from small student groups:

- Review standard deviations (or other measures of variance) when evaluating whether small schools are to be designated as low-performing. If a small school has large variance, consider using multi-year averages or removal of outliers to achieve more stable measures.
- Revisit the prior administrative practice of creating a "super subgroup" of students from multiple underrepresented racial or ethnic groups to increase the number of schools that have data for at least 10 students.

*Measures of Student Learning*

The Metrics for Capturing Student Performance section of this report raises options for future consideration, now that federal policy has become more flexible in allowing states to design the systems that work best for their schools and students. With that in

mind, we propose these future analyses to explore other measures that have the potential to improve upon Maine's proposed methods.

- When additional assessment data are available, investigate whether using average student scores would enhance understanding of student academic performance without producing any unforeseen data challenges.
- Consider alternative methods of calculating student learning growth that may improve the meaningful differentiation of schools. Examples include measures based on changes in student score percentiles or scale scores, which may or may not include value-added methods to consider student factors such as poverty level, special education status, or English proficiency.
- To address the concern that the proficiency rate and growth measures have substantial overlap, consider use of NWEA scores for the student growth measure instead of state summative assessments. Used by a large proportion of Maine districts, and available for additional grade levels, the NWEA exam may potentially provide for greater differentiation of schools. The NWEA has a wider variance in scores because with the computer-adaptive test students can be assessed on material that is below or above their actual grade level. This may be philosophically more aligned to the expectations for teachers in a proficiency-based learning system, and thus better able to capture growth in student learning that is beyond the content in the grade level. This would necessitate substantial changes in state policy to require additional assessments.

In summary, Maine has already done substantial high-quality work to develop the framework for a strong accountability system. But given the stakes involved for schools and students, the work of building a high-quality model must be ongoing.

## References


Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal*, *42*(1), 231–268.

Center on Education Policy. (2011, May 3). An open letter from the Center on Education Policy to the SMARTER Balanced Assessment Consortium and the Partnership for Assessment of Readiness for College and Career. Retrieved from http://cep-dc.org/displayDocument.cfm?DocumentID=359

Ho, A. D. (2008). The problem with "proficiency": Limitations of statistics and policy under No Child Left Behind. *Educational Researcher*, *37*(6), 351–360.

Holland, P. W. (2002). Two measures of change in the gaps between the CDFs of test-score distributions. *Journal of Educational Behavioral Statistics*, *27*(1), 3–17.

Kober, N., & Riddle, W. (2012). *Accountability issues to watch under NCLB waivers*. Washington, DC: Center on Education Policy.

Linn, R. L. (2003). Accountability: Responsibility and reasonable expectations. *Educational Researcher*, *32*(7), 3–13.

Linn, R. L. (2007). *Educational accountability systems*. Paper presented at the The CRESST Conference: The Future of Test-Based Educational Accountability.

Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher*, *31*(6), 3–16.

Maine Education Policy Research Institute (2017). The Every Student Succeeds Act: A Summary of Federal Policy and Implications for Maine.

Neal, D., & Schanzenbach, D. W. (2010). Left behind by design: Proficiency counts and test-based accountability. *Review of Economics and Statistics*, *92*, 263–283.

Ng, H. L., & Koretz, D. (2015). Sensitivity of school-performance ratings to scaling decisions. *Applied Measurement in Education*, *28*(4), 330–349.

Polikoff, M. S., McEachin, A., Wrabel, S. L., & Duque, M. (2014). The waive of the future? School accountability in the waiver era. *Educational Researcher*, *43*(1), 45–54. http://doi.org/10.3102/0013189X13517137

Polikoff, M.S. et. al. (2016). A Letter to the U.S. Department of Education. Accessed at https://morganpolikoff.com/2016/07/12/a-letter-to-the-u-s-department-of-education/

Schwartz, H. L., Hamilton, L. S., Stecher, B. M., & Steele, J. L. (2011). *Expanded measures of school performance*. Santa Monica, CA: The RAND Corporation.

**Appendix A: Report Addendum**

Maine's consolidated application for federal funds was deemed complete by the US Department of Education and released as final after this report had been sent for printing. Upon review, it became apparent that the description in Table 2 of Maine's proposed accountability measure weights was incomplete.

Maine intends to use varied weights within the academic progress measure, which comprises 40% of the overall rating for elementary schools. (High schools are evaluated on graduation rates rather than academic progress). As described in the consolidated application:

> "Maine's Academic Progress measure is computed based on a blended approach that incorporates both measures of academic proficiency and growth for the school. This approach was recommended by Maine's Accountability Advisory Working Group and is based on the Student Learning Index (SLI) presented by AdvancED at the ESSA Symposium in September 2016. Under this approach, schools are divided into quartiles based on their proficiency in the content area (i.e., <sub> = ELA or mathematics). The quartile to which a school is assigned determines the weighting scheme for the proficiency and growth measures that are used in the Academic Progress calculation" (Maine Consolidated Application, p. 32).

The final matrix of measure weights for each quartile of proficiency rates is not yet final, and will be determined after a second year of current assessment data are available. In example weights provided as an illustration, schools in the lowest quartile of student proficiency rates would be rated with a greater emphasis on student growth. They might have 25% of their academic progress based on proficiency rates and 75% based on student growth. Schools in the highest proficiency quartile would be treated in the opposite with a greater weight for proficiency, such as 75% proficiency rate and 25% growth.