

STATE ESSA PLANS: Uneven Progress toward Better Assessment and Accountability

II. FINDINGS, DISCUSSION AND ANALYSIS OF FIRST 16 APPROVED STATE ESSA PLANS

FairTest reviewed the 16 state ESSA plans submitted in spring 2017 and approved by the US Department of Education (DoE) as of the end of December 2017. On one topic, the requirement to test 95% of all students, FairTest reviewed three additional states – Colorado, New York and Utah – whose initial plans were rejected by DoE. We also reviewed New Hampshire’s now-approved proposal because it is the only state that does not require the use of standardized tests but instead builds on its pilot program to largely replace such tests with performance assessments. New Hampshire had received a waiver from No Child Left Behind for this pilot program under previous DoE Secretary Arne Duncan. (See *Assessment Matters: Constructing Model State Systems to Replace Testing Overkill*, at <http://www.fairtest.org/assessment-matters-constructing-model-state-system>.) All state plans can be found at <https://www2.ed.gov/admins/lead/account/stateplan17/statesubmission.html>. That page lists approved as well as unapproved submissions that are updated when approved.

FairTest reviewed five topic areas: the School Quality/Student Success (SQSS) indicators identified by each state, the weighting given all the indicators, how many levels of schools are identified by the state, the core approach to improving low-ranked schools (whether a state retains the punitive NCLB approach or changes to a more supportive plan), and how a state responds to the federal requirement that each school test at least 95% of all its students and of each students in a subgroup (race/ethnicity groups, low-income, disability, English learner). All are summarized in the chart below. The analysis also looked at funding adequacy, since the combination of poverty, race, disability and school funding are powerful predictors of student success in school, even though ESSA requires little in this regard.

School Quality/Student Success (SQSS). ESSA does not specify what SQSS indicators should be, but it does list some options: student engagement; educator engagement; student access to and completion of advanced coursework;

postsecondary readiness; or school climate and safety. Of these, only post-secondary readiness is commonly used, and states have for the most part selected a narrow range of primarily test-based indicators on top of readiness and chronic absenteeism.

Among 16 approved states, 13 use chronic absenteeism and 12 include college/career readiness or a similar indicator.¹ These are far and away the most common SQSS indicators. Seven of 16 states include science test scores. IL and MI include access to arts, while CT, MI and VT include physical education or fitness. DE and LA use social studies scores, and DC specifically notes AP/IB participation and success (though other states may include that in college readiness). IL includes access to a librarian or media specialist.

School climate was included in just three states (IL, NM, ND). ND also includes student engagement. Only OR said it was considering incorporation of exclusionary discipline data (e.g., suspension and expulsion rates), while LA will use that data in determining a school's level but not in determining subsequent accountability actions. DC listed inclusion of students with disabilities in general education and "access and opportunities," NM included an "opportunity to learn" survey, and LA will incorporate "interests and opportunities." Many of these terms are not described in any detail. Some states are either reporting but not weighting additional indicators, or say they are developing more indicators for future use if they are valid, reliable and differentiate among schools.

There are serious limitations with the indicators emphasized by most states. For example, college/career readiness is most often measured by scores on various tests – SAT or ACT, Work Keys, International Baccalaureate, Advanced Placement, etc. But SAT and ACT are not strong predictors of college success compared with, for example, student grades (See "Study: High school grades best predictor of college success — not SAT/ACT scores," https://www.washingtonpost.com/news/answer-sheet/wp/2014/02/21/a-telling-study-about-act-sat-scores/?utm_term=.5528b778dba0). Access to AP and IB courses are strongly correlated with community wealth and student SES. A few states include college enrollment or intend to track actual student success in college (though this too is heavily influenced by family wealth/income). Science and social studies are measured by test scores. Chronic absenteeism is important to track and address, but it too likely correlates closely with family socio-economic status (SES), meaning once again under-resourced schools with many low-income students (often students of color) will be flagged. These schools frequently need additional assistance, but states should address that regardless of ESSA indicators. One

¹ This is similar to what Education Commission of the States found in reviewing all plans, approved or not. ECS said 36 of 51 will use absenteeism and 35 will use readiness; available at <https://www.ecs.org/50-state-comparison-states-school-accountability-systems/>.

state, IL, said it wanted few SQSS indicators because it lacked resources to support schools in these areas. A refusal to adequately fund public education (and IL has one of the weakest state funding programs) here leads to a continuing focus on reading and math test scores.

Few states have strong, non-test indicators of opportunity to learn. All students deserve access to a full range of academics, including the arts and advanced work options. They need nurses, librarians, media specialists and counselors. Wrap-around services (support in obtaining housing, medical and dental care, counseling, etc.), often provided by “Community Schools,” are a growing, successful practice. (See “Community Schools: An Evidence-Based Strategy for Equitable School Improvement,” https://drive.google.com/file/d/0BzYGC6_i31OhV0VfOU51eGtTcHc/view.) In short, states should ensure the education of the whole child. (See *The Campaign for the Education of the Whole Child*, http://www.citizensforpublicschools.org/wp-content/uploads/2009/08/whole_child_report.pdf.)

School climate surveys are sufficiently well-established that some states (IL and ND) include them while others are considering adding them. (See “Making the Most of the Every Student Succeeds Act (ESSA) – Helping States Focus on School Equity, Quality and Climate,” <http://nepc.colorado.edu/publication/ESSA>.) Unfortunately, most do not. Surveys can include student or parent perceptions of access to a rich curriculum, quality of teaching, fairness in such things as discipline and course access, student safety, and how welcoming the school is to parents. Every state must report suspension/expulsion data under separate federal policies, but only LA and OR will use this data. These are important indicators of school quality that often affect student progress, including whether they are engaged or disengaged, or at risk for leaving or being pushed out of school. Similarly, some states have included opportunity to learn, so there is no question that DoE accepts it as an indicator. At a minimum, states should test out such factors to develop additional reliable indicators. ESSA says indicators must differentiate among schools; done well, climate and discipline data are likely to do so. If an indicator showed almost all schools were doing well or poorly, that information should be shared with the public – and if poorly, then all schools should improve.

Weighting of indicators. States must assign weights to each indicator, usually expressed as a percent of their weighting formula with the total equaling 100%. ESSA requires states to include:

- 1) Achievement (state test or assessment scores) in ELA and Math (often labeled “proficiency”);

- 2) Progress by English Language Learners (ELL) in schools with enough such students, measured by an English language proficiency assessment;
- 3) At least for elementary and middle schools, student growth or another indicator of school progress, which in most states means score gains on the state test (e.g., Student Growth Percentiles or Value Added Measurement);
- 4) For high schools, the four-year graduation rate and, at state choice, extended year rates; and
- 5) At least one other indicator of school quality or student success (SQSS). ESSA requires states to give the four core indicators “significantly” greater weighting, but does not specify what that means.

Each indicator must be able to produce data by “subgroups” and to “meaningfully differentiate” among schools. States often use complex formulas to establish these weights and translate them into school ratings.

In most states, achievement and growth/progress carry the most importance, although in a few states graduation rates or SQSS are more heavily weighted (see the table). Weighting also varies between elementary-middle (EM) and high schools (HS). Ten of the 16 states factor growth more heavily than achievement, and two weigh them equally in EM. (ESSA does not require high schools to include growth. Because most states only test one high school grade, many do not include growth as a factor.)

While both types of scores are derived from the same limited tests, growth indicators at the school level can to some extent break the extremely tight correlation between achievement scores and family socio-economic status. That is, low-scoring schools may show relatively fast score gains. Thus, increased emphasis on growth can reduce the number of low SES schools being identified as lowest-performing. This perhaps mattered more under No Child Left Behind (NCLB), when schools faced escalating punitive sanctions, than it will under ESSA, at least in states that have moved away from punitive actions toward a focus on assistance. *Overall, this is a positive step in recognizing that when it comes to test scores, progress is more important than status. However, because of the underlying limitations and dangers, heavy reliance on standardized tests remains ESSA’s primary flaw.* (See “How Standardized Testing Damages Education,” <http://fairtest.org/how-standardized-testing-damages-education-pdf>; and *Failing Our Children*, <http://www.fairtest.org/node/1778>.)

States vary in whether they use 4-year or more than 4-year graduation rates. ESSA says the 4-year rate must count for more, so states have formulas for weighting 4-, 5-, 6- and even 7-year rates. *FairTest supports multiple-year graduation rates: Schools that keep students enrolled an extra year or two so they stay on to graduate should receive credit.* This is especially important for schools that serve transfer students who have fared poorly in other high schools or those who are

returning dropouts. Fortunately, only two of the 16 states, DC and LA, rely solely on four-year rates, and DC includes an unelaborated “alternate graduation metric.” In general, the weight given graduation rates ranges widely, indicating lack of agreement about its importance. Four states weight it at 40% or higher, 10 at 20% or lower, with 2 of those at less than 10%. CT and OR use a non-percentage-based weighting formula (as is also the case for their other indicators).

English language learners are included in accountability formulas if a school has sufficient numbers. Thus, states will increase weights for other factors if ELL enrollment is lower than the minimum required “N-size” for counting them. Most states are weighting ELL in the 10% range; DC and IL are at only 5%, and NJ is at 20%. This factor is based on English-proficiency standardized tests. Specific inclusion of ELL progress is important, but there are many concerns about the tests, including that they are too long. In addition, some ELL teachers think it would make more sense to integrate ELA/reading into ELL tests, especially for those in earlier stages of learning English, instead of testing these students twice.

For weighting SQSS, three states (DC, NV and ND) weight it at 35% or higher, but only NV does so for both EM and HS; nine at between 20% and 34%; and six at 19% or lower. (Again, CT and OR use weighting formulas that do not reveal this information. Because states often weight SQSS differently for elementary-middle and high schools, the total exceeds 16). FairTest thinks states should have multiple, diverse indicators that come close to 50% of the total. Having few SQSS indicators or assigning low weights to SQSS usually means even more emphasis on test results. *States generally have too few indicators, too many are test-based, too few of them get at school quality in a rich or deep way, and SQSS weightings are far too low. In addition to moving quickly to include more indicators, states should increase their weighting to 40% or more.*

School ratings/levels. ESSA requires states to identify schools requiring Comprehensive Support and Improvement (CSI) and schools requiring Targeted Support and Improvement (TSI). A state can, at its discretion, create sub-categories of CSI or TSI (a few do) and can divide all other schools into various levels. Of the 16 states, 6 require only three levels (including MA, which intends to change its approved plan to shift from 6 to 3), 3 have 4 levels, and 6 use 5 levels, most of which rank schools with an A-F letter grade. CT has 5 reporting levels but only 3 intervention levels.

There is no good rationale for having more levels than the three required by the federal government (CSI, TSI and all others). The primary result of extra levels will be to perpetuate unhealthy competition among schools and districts aimed at boosting standardized test scores, the primary factor in most states’ weighting formulas. Thus, the more levels, the more pressure to over-emphasize tested subjects, narrow the curriculum, and teach to the test. As discussed below, the primary causes of low school “performance” are community/family poverty, racial isolation/segregation, and a lack of

school resources. It is certainly important to strengthen curriculum, instruction and assessment, among other school factors, to improve the daily experience and learning outcomes of students. However, labeling schools as failing largely because of factors beyond their control, as is most clearly done with “F” and “D” labels, blames the victims, including students, teachers and other school staff, as well as the impoverished communities states fail to adequately serve.

Even indicators that can be helpful can be misused in the levels-setting process. For example, Louisiana laudably includes overuse of punitive discipline as an indicator. However, punitive discipline is more prevalent in schools with large numbers of youth of color, so this can add another negative outcome to the typically lower test scores and graduation rates at such institutions. (See “Discipline Policies, Successful Schools, and Racial Justice,” <https://www.civilrightsproject.ucla.edu/research/k-12-education/school-discipline/discipline-policies-successful-schools-and-racial-justice?searchterm=discipline>.) States should use such indicators as a level to improve schools, not undermine them.

Some states use a fourth or fifth top level with the idea of rewarding the “best” schools, at least symbolically. In almost all cases, this means recognizing schools for the wealth and often whiteness of their community because most of the indicators closely track socio-economic status. If a state wants to identify schools with high-quality practices, it should look at instruction and other behaviors associated with all-around strong schools.

Support/intervention and funding adequacy. ESSA requires states to establish plans for providing support to schools in CSI or TSI status. Districts must prepare school improvement plans that follow ESSA and state requirements. State plans describe various forms of support. If a school does not make sufficient progress (usually after three years, occasionally two or four years), ESSA requires states to intervene in a stronger way, including by providing additional support. Most states say they will do so but provide little specificity.

Some states maintain No Child Left Behind-era forms of intervention: firing staff, privatizing school management, or closing (and perhaps reopening) the school. Six of the 16 states call for such interventions: CT, LA, MA, NV, NM and TN. Some do so from the very start of when a school is identified, more of them take this step at the start of the second cycle when a school has not sufficiently improved. But three or four years is not enough time to solve deep-rooted problems, especially those heavily determined by poverty. States should not jump to these harsh measures after too short a time. That said, if after six years of support, including provision of adequate funding, a school does not show meaningful progress, it *would* be time for much stronger actions.

ESSA requires states to provide funding information, including per-student expenditures, down to the district and school levels. This can help reveal important inequities. However, ESSA does not require that this information be used to inform needs assessment and improvement plans. Only three of 16 states -- Delaware, Illinois and Oregon -- include funding adequacy or equity in their needs assessments. When it comes to ensuring adequacy for CSI/TSI schools, states plans are generally not concrete. They may require districts to address the issue, leaving often impoverished locales to solve the problem on their own.

ESSA mandates that states set aside 7% of their Title I funding to support improvement. This will help but often will not be sufficient. Moreover, there is a risk that a short-term infusion of additional money produces improvements that lead to termination of the additional funding and a recurrence of the problems. *The lack of inclusion of funding adequacy in needs assessments and improvement plans is one way for states to continue ignoring this problem.*

States aim to provide a variety of supports to schools. Under the law, districts design improvements for a CSI or TSI school based on the local needs assessment and any state requirements, such as having to work with approved external partners.

States can distribute ESSA Title I improvement funds via formula grants or competitive grants or a mix. Among the 16 states, two use formula grants, five use competitive grants, and three use a hybrid of the two. The remaining six are not clear in their ESSA plans. Distributing funds on a formula basis is better; it is inequitable and wrong to deny students needed resources simply because a district does not write a good proposal. States should help districts prepare high-quality plans and implement them well.²

Lastly, “improvement” efforts under NCLB and Duncan-era waiver requirements focused on firings, closings and privatization. These were ineffective, with only spotty successes in raising test scores, never mind improving anything else. *As a result of NCLB’s failure, states and districts have to address anew the question of how to help schools.* Many die-hard defenders of NCLB’s failed, punitive approach have charged states are not specific enough in what they will do. Indeed, states use vague language such as ‘if a school does not improve sufficiently after three years, the state will engage in stronger interventions.’ But given the lack of proven strategies, and despite ESSA requirements to use

² A look at this issue is “ESSA Leverage Points: 50-State Report on Promising Practices for Using Evidence to Improve Student Outcomes,” available at http://blogs.edweek.org/edweek/campaign-k-12/RFA%20ESSA%2050%20State%20Report_final.pdf. Unfortunately, it supports competitive awarding of funds on the basis of district plans rather than by formula.

“evidence-based” interventions, this is a rational response. Whether all the efforts will produce much genuine improvement remains to be seen, but repeating NCLB failed approaches is not a solution.

95% participation. ESSA mandates that 95% of all eligible students and of students in each subgroup take federally mandated state exams. DoE expects states to factor participation into state school ratings. This means determining school average test scores by having a denominator representing at least 95% of eligible students.³ *The consequence of this mandate is to lower test scores for schools with more than 5% non-participants (most of whom will be opt outs, because other reasons such as illness are rare).* This can lead to placing a school at a lower level, even to identification as CSI or TSI, which could trigger interventions. As discussed above, 14 of the 16 states will impose additional sanctions on schools in which participation does not reach 95%. Because lower-income districts are more likely to have schools in CSI or TSI, this policy is more likely to inhibit opting out on such districts, particularly in more punitive states, as people seek to avoid labels and sanctions. On the other hand, wealthier districts could end up labeled low-scoring and be eligible for federal funds they do not need, as New York State Allies for Public Education has pointed out.

LA and NJ do not specify any actions other than a zero score for non-participants. NJ and NY will maintain two lists of schools, one for the federal government plus one for state reference with a denominator that includes only the test takers. It is not clear how the separate lists will be used, but it seems the state list will be used to ensure schools are not subject to interventions based on low participation.

CO and UT also sought to protect families and schools, but their initial plans were rejected by DoE. CO now proposes to maintain two lists but will identify schools using the federal list, and if any schools end up in CSI or TSI, they “will receive differentiated support and intervention compared to those that are identified based on the actual performance of tested students.” UT essentially ignored the whole question in its application; the DoE has asked UT to respond to the issue.

Opt out opponents point to language in ESSA that suggests states need to take action in addition to the 95% denominator mandate. However, LA and NJ have not done so, and their plans were approved. Hopefully, requiring a district to develop a plan to tell parents their children should take the tests will amount to little or nothing. That may well depend on the strength of the local opt-out movement.

ESSA essentially demands that states penalize schools for the actions of parents and students, and does so even though the law explicitly says states can allow students to opt out, a provision DoE has completely ignored. Currently, 10 states

³ In the chart, the assumption is a state uses the 95%-in-denominator requirement, though not all states explicitly state they do.

have laws allowing some or all students to refuse to take the test. Of these, two have approved ESSA plans. Both require districts to develop plans to lower the opt-out rate, but neither proposes any other penalties.

Lastly, states have often continued to propose “ambitious” goals that tend to perpetuate the completely unrealistic NCLB mandate that 100% of all students will score proficient or higher on state tests. Fortunately, NCLB sanctions are no longer attached to these goals. However, unrealistic goals can distort other policies, often causing damage to low-income schools, their students and teachers.⁴

⁴ See, for example, D. Koretz, *The Testing Charade*, 2017, University of Chicago Press; F. Hess, “How Cheap Talk Fuels Bad School Accountability,” Education Week; http://blogs.edweek.org/edweek/rick_hess_straight_up/2017/12/how_cheap_talk_fuels_bad_school_accountability.html.