Shifts in Assessment and Accountability Policy Under the Every Student Succeeds Act

March 23, 2016

Robert Stonehill, Ph.D., Managing Director
David English, Technical Assistance Consultant
Objectives

• Describe the Every Student Succeeds Act (ESSA) implementation timeline
• Describe the negotiated rulemaking process under the ESSA
• Describe ESSA effects on statewide assessment and accountability policy, in context of shifts from No Child Left Behind (NCLB) and Elementary and Secondary Education Act (ESEA) flexibility
## Every Student Succeeds Act Timeline

<table>
<thead>
<tr>
<th>Date</th>
<th>Milestone</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 10, 2015</td>
<td>ESSA was passed.</td>
</tr>
<tr>
<td>March 21, 2016; early and late April 2016</td>
<td>Negotiated rulemaking sessions</td>
</tr>
<tr>
<td>July 1, 2016</td>
<td>ESSA generally applies to formula programs</td>
</tr>
<tr>
<td>August 1, 2016</td>
<td>ESEA flexibility waivers become null and void (but continue support for priority and focus schools)</td>
</tr>
<tr>
<td>Fall 2016</td>
<td>Earliest possible date of regulations</td>
</tr>
<tr>
<td>October 1, 2016</td>
<td>ESSA applies generally to competitive federal grants given out after this date</td>
</tr>
<tr>
<td>Prior to SY2017–18</td>
<td>State plan review will occur, including peer review (March calls between ED and states will include state plan timeline)</td>
</tr>
<tr>
<td>No later than 120 days after submission</td>
<td>State plan approval (includes right to a hearing)</td>
</tr>
<tr>
<td>SY2017–18</td>
<td>ESSA requirements must be implemented.</td>
</tr>
</tbody>
</table>
Negotiated Rulemaking

• The U.S. Department of Education (ED) must use negotiated rulemaking:
  – If it chooses to develop proposed regulations regarding standards and assessments under section 1111(b)(2) of ESEA
  – To regulate the requirement under section 1118 of ESEA that Federal funds be used to supplement, and not supplant, State and local funds

• Note that accountability regulations are not required to be established through “NegReg”
Negotiated Rulemaking

• Before proposing any rules, ED convenes a panel of stakeholder constituency representatives to develop proposed regulations
• This is done through a series of facilitated meetings with the “negotiators” and ED officials
• A list of the negotiators is available on the ED website at: https://www.ed.gov/news/press-releases/us-department-education-names-committee-members-draft-proposed-regulations-every-student-succeeds-act
Negotiated Rulemaking

- Negotiators met March 21, 2016 and will meet again at least once in April 2016
- Negotiators will define “consensus” and determine protocol for discussion and agenda
- If consensus is achieved on new regulations, ED initiates the traditional rulemaking process (i.e., develops a Notice of Proposed Rule Making subject to public comment requirements)
- If consensus is achieved, the earliest possible regulations can be published in fall 2016
Negotiated Rulemaking

• If consensus is not reached, ED must submit proposed regulations to the relevant Congressional committees
• Congress has a 15-day opportunity to provide comments on the proposed regulations, which must be addressed by ED as part of the public rulemaking record
Key to Symbols

- Clear policy shift from NCLB or ESEA flexibility (underlined text indicates a specific shift)

- Language shift from earlier ESEA (implications for actual policy shifts might be unclear)

- First inclusion in ESEA of a preexisting policy (from other guidance, another federal statute, etc.)
Statewide Assessments

• Administer high-quality assessments in:
  – Mathematics and reading or English language arts (ELA) annually in grades 3-8 and once in high school
    » Accountability system to be based on mathematics and ELA at minimum
  – Science (grade-span testing) once in each of three grade bands (grades 3–5, 6–9, and 10–12)

• Aligned with challenging standards, which are aligned with:
  – Higher education entrance requirements for “credit-bearing coursework”
  – Relevant career and technical education standards
Statewide Assessments

- Valid, reliable assessments, aligned to national technical standards, that “…do not assess personal or family beliefs and attitudes”
- Provide information about higher-order thinking skills, student attainment of standards and “whether the student is performing at grade level”
- Differentiate three achievement levels (out of four possible performance levels, e.g., Unsatisfactory, Basic, Proficient, and Advanced)
- Enable disaggregation and reporting by major racial and ethnic groups, economically disadvantaged (ED), students with disabilities (SWD), English learners (EL), gender, and migrant status (where \(n\) size is sufficient)
- Same assessment is given to all students and schools
Statewide Assessment Administration

State educational agencies (SEAs) “may”:

- Allow districts to use **nationally-recognized high school assessments** (e.g., ACT and SAT) aligned to state standards and comparable to regular statewide tests.
- Be administered as a single summative assessment or **multiple interim assessments** that result in a single summative score.
- Allow **grade 8 math students** to test on grade 9-12 test if he/she then takes a higher-level test in HS and both results drive accountability for their respective administration years.
- Administer statewide assessment partly in the form of **projects, portfolios, and extended-performance tasks**.
- Allow **computer-adaptive testing** using items “above or below the student’s grade level” (e.g., to better measure growth).
- Set a target limit on aggregate **test administration time**, expressed as a percentage of instructional time.
Statewide Assessments

Students With Disabilities

- Facilitate inclusion of students with disabilities in regular statewide assessments, including accommodations
- May provide alternate assessments for students with the most severe cognitive disabilities
- Embed universal design for learning (UDL) principles in alternate assessments “to the extent feasible” (using UDL definition from Higher Education Act)
- Cap administration of alternate assessment at 1 percent of participating population per subject (not just 1 percent inclusion for accountability with unlimited administration)
- Eliminates “modified” assessments for disabled students (“2 percent” cap)
Statewide Assessments

English Learners

• Facilitate inclusion of ELs in regular statewide assessments, including use of accommodations
  – “To the extent practicable” use native-language assessments if they yield more accurate information
• Indicate languages that are present to a “significant extent” in participating student population for which assessments are needed and “make every effort” to develop such assessments (may ask Secretary for assistance)
• Use native language reading/English language arts (ELA) assessments for ELs for no longer than three years of enrollment + two extra years on case-by-case basis
• May include former ELs in the EL subgroup for accountability for up to four years, for math and ELA (three years previously)
• Administer annual assessment of English language proficiency (ELP) aligned to ELP/state standards
Statewide Assessments

“Recently Arrived” ELs enrolled in a U.S. school for <12 months:

- For one year, exclude the student from taking the ELA assessment and not count ELA, math and/or ELP results towards accountability

OR

- Include in first year participation; report on but exclude first-year ELA and math results from accountability; for the second year, include student growth in ELA and math; and for the third year, include proficiency in ELA and math
Statewide Assessments

• SEAs “may” use federal funds to:
  – Develop balanced assessment systems (i.e., formative, interim, and summative)
  – Develop competency-based assessments
  – Conduct audits to ensure assessments are necessary and high quality
  – Develop science assessments to integrate concepts related to engineering and technology
Statewide Assessments

Innovative Assessment System

- Up to seven SEAs (any consortia not to exceed four SEAs) may apply to develop and implement an innovative statewide assessment
- Pilot may be used for competency-based, performance-based, interim assessments, for accountability purposes
- Pilot development may take up to a five-year period (plus two years extension based on evidence)
- The assessment does not have to be administered:
  - To all schools or students
  - To all grade bands as identified for mathematics, ELA, and science
Statewide Assessments

Potential key issues for negotiated rulemaking

- What constitutes a “nationally-recognized” test? (e.g., ACT, SAT)
  - How ensuring comparability to statewide tests?
  - How ensuring accessibility by SWDs and ELs?
- What data around computer-adaptive testing must be reported?
  - Performance against grade-level?
- Definition of “students with disabilities” in assessment context
- Definition of “students with the most significant cognitive disabilities”
- How to operationalize 1% alternate assessment cap across districts
- Clarification on conditions for developing additional native-language assessments and “every effort” to develop

(Ujifusa and Klein, 2016)
Accountability Dimensions

- “Long-term goals”
- Academic and nonacademic indicators
- Meaningful differentiation of all public schools (not just Title I)
- Periodic identification of low-performing schools
- Supports for improvement activities
- Annual reporting
Long-Term Goals

• “Long-term” and “interim” goals for all students and disaggregated subgroups

• At minimum, proficiency, graduation rate, and progress toward English language proficiency (EL subgroup only)
  – Participation and additional academic indicator no longer discrete goals
  – Extended-year graduation rate optional, more rigorous than 4-year rate

• “Ambitious”

• No prescribed target-setting methodology
  – Targets must make “significant progress” towards closing statewide proficiency and graduation rate gaps

• Performance against proficiency and graduation goals informs annual differentiation and identification of low-performing schools (not English language proficiency goals)
Performance Against Long-Term Goals Informs School Differentiation and Identification

• This is a key mechanism for:
  – Gap closure
  – Overall improvement as a state

NCLB
Performance against Annual Measurable Objectives (AMOs) drives identification.

ESSA
“[A]nnually measure [and identify schools]...based on the long-term goals [for]...proficiency on the annual assessments...[and] four-year adjusted cohort graduation rate.”

ESEA flexibility
Many SEAs include additional factors to identify low-performing schools.
Using Performance Against Proficiency Goals for Accountability

**Example: Maryland proficiency calculation**

<table>
<thead>
<tr>
<th>2013 Achievement Calculation</th>
<th>Math</th>
<th>Reading</th>
<th>Science</th>
<th>Combined Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of Students who Scored Advanced or Proficient</td>
<td>44.12%</td>
<td>52.21%</td>
<td>44.83%</td>
<td></td>
</tr>
<tr>
<td>÷ 2013 Achievement AMOs</td>
<td>74.18%</td>
<td>73.59%</td>
<td>21.88%</td>
<td></td>
</tr>
<tr>
<td><strong>= Measure Progress Scale Values</strong></td>
<td>0.5948</td>
<td>0.7094</td>
<td>2.0493</td>
<td></td>
</tr>
<tr>
<td>× Proportional Significance Assigned to Each Measure</td>
<td>33.33%</td>
<td>33.33%</td>
<td>33.33%</td>
<td></td>
</tr>
<tr>
<td><strong>= Measure Contribution</strong></td>
<td>0.1983</td>
<td>+ 0.2365</td>
<td>+ 0.6831</td>
<td>= 1.1178</td>
</tr>
<tr>
<td>× Proportional Significance Assigned to Achievement in Calculating the School Progress Index</td>
<td>30.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>= Achievement Contribution Value</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.3353</td>
</tr>
</tbody>
</table>
Using Performance Against Graduation Rate Goals for Accountability

Example: Colorado graduation rate calculation

<table>
<thead>
<tr>
<th>Graduation Rate and Disaggregated Graduation Rate: The school’s graduation rate/disaggregated graduation rate was:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exceeds</td>
</tr>
<tr>
<td>Meets</td>
</tr>
<tr>
<td>Approaching</td>
</tr>
<tr>
<td>Does Not Meet</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dropout Rate: The school’s dropout rate was:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exceeds</td>
</tr>
<tr>
<td>Meets</td>
</tr>
<tr>
<td>Approaching</td>
</tr>
<tr>
<td>Does Not Meet</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Postsecondary and Workforce Readiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduation Rate and Disaggregated Graduation Rate: The school’s graduation rate/disaggregated graduation rate was:</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Exceeds</td>
</tr>
<tr>
<td>Meets</td>
</tr>
<tr>
<td>Approaching</td>
</tr>
<tr>
<td>Does Not Meet</td>
</tr>
</tbody>
</table>

If graduation rate “Meets” the target →
School awarded three of four possible points =
75 percent measure score
At Least Four Indicators for Annual Differentiation

- Proficiency in ELA and math
- Graduation rate (high schools) OR Growth or another “valid and reliable” statewide academic indicator (elementary and middle schools)
- English-language proficiency progress
- Nonacademic indicators of school quality or student success (must be able to disaggregate) may include the following:
  - Postsecondary readiness
  - Student access to a completion of advanced coursework
  - School climate and safety
  - Student engagement
  - Educator engagement
Annual Differentiation of All Public Schools

- Based on all students’ performance across all indicators.
- Based on subgroups “consistently underperforming” across all indicators (except English language proficiency indicator only for EL subgroup):
  - Economically disadvantaged students
  - Major racial and ethnic groups
  - Students with disabilities
  - English learners
Support for Schools With “Consistently Underperforming” Subgroups

**Comprehensive Support**
Title I schools with “Additional targeted support” subgroup that continues to underperform for a state-determined number of years

**Additional Targeted Support**
Schools with subgroup consistently underperforming across all indicators comparable to lowest 5% of Title I schools for all students (annual identification)

**Targeted Support**
Schools with subgroup that is consistently underperforming across all indicators (annual identification)
Identification for Comprehensive Support Every Three Years (Triennial)

- **Lowest performing 5 percent** of **Title I** schools based on all annual differentiation indicators, applied to “**all students**”
- **All** public schools with **graduation rates** 67 percent or lower for all students
- **All Title I** “additional targeted support” schools (i.e., identified for low **subgroup performance** comparable to lowest 5% of all Title I schools based on all students) that continue to underperform over a state-determined number of years

No category of high-performing schools (e.g., “Reward” schools)
## Support Structure for Low-Performing Schools

<table>
<thead>
<tr>
<th>Support Category</th>
<th>Common Elements of Improvement Plan</th>
<th>Other Plan Elements</th>
<th>Who Approves and Monitors Plan (Including Interventions)</th>
<th>Who Determines Exit Criteria, Including Maximum Time to Exit</th>
<th>Consequences of Failure to Exit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Targeted Support</strong></td>
<td>Informed by all indicators, including performance against long-term goals</td>
<td>N/A</td>
<td>District</td>
<td>District</td>
<td>Additional action determined by local education agency</td>
</tr>
<tr>
<td><strong>Additional Targeted Support</strong></td>
<td>“Evidence-based” interventions</td>
<td>Identifies and implements resource inequities</td>
<td>District</td>
<td>State</td>
<td>Escalation to comprehensive support status (Title I schools)</td>
</tr>
<tr>
<td><strong>Comprehensive Support</strong></td>
<td>Based on school-level needs assessment</td>
<td>Identifies and implements resource inequities</td>
<td>State</td>
<td>State (not to exceed four years)</td>
<td>More rigorous, state-determined action, such as operations-related actions</td>
</tr>
</tbody>
</table>
“Evidence-Based” Interventions

• **Strong evidence base**
  – Shows statistically significant effect on student outcomes from at least one experimental study

• **Moderate evidence base**
  – Shows evidence from a quasi-experimental study

• **Promising evidence base**
  – Shows evidence from a correlational study that makes statistical corrections for selection bias
State Role in Supports for School Improvement

- No direct support activities from states except if permission is received from local educational agency.
  - Except state-determined action for schools that fail to exit comprehensive support status, including operations.
- State-defined interventions are not precluded.
- Review resource allocation for and provide technical assistance to LEAs serving significant number of comprehensive and/or targeted support schools.
- May initiate LEA improvements if significant number of:
  - Comprehensive support schools not exiting
  - Targeted support schools
- Use 7 percent of Title I allocations for improvement activities (up from 4 percent).
Combining Measures to Differentiate and Identify Schools

• Does not explicitly require the calculation or reporting of an overall school rating (e.g., based on a composite index)
  – “Nothing…permit[s] the Secretary…to…prescribe…as a condition of approval of State plan….the specific methodology…to meaningfully differentiate or identify schools” (Sec. 1005(e)(1)(B)(iii)(V)).

• Annual differentiation
  – “Substantial” weight to each of the three academic indicators
  – “Much greater weight” in aggregate to academic indicators
  – “Nothing…permit[s] the Secretary…as a condition of approval of State plan…to prescribe…the weight of any measure or indicator used to identify or meaningfully differentiate schools” (Sec. 1005(e)(1)(B)(iii)(IV)).
## Differentiation criteria for schools in a hypothetical SEA

<table>
<thead>
<tr>
<th>Annual Differentiation Rating</th>
<th>Index Score Criterion</th>
<th>Additional Criteria Aligned to Comprehensive and Targeted Support Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>90%–100%</td>
<td>• No “A” schools can have subgroups targeted for support.</td>
</tr>
<tr>
<td>B</td>
<td>80%–89%</td>
<td>• No “B” schools or higher:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Can have graduation rate &lt;=67%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Can be identified for comprehensive support for low subgroup performance</td>
</tr>
<tr>
<td>C</td>
<td>70%–79%</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>60%–69%</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>59% or less</td>
<td>• Set “F” cut score to include bottom 5% of Title I schools</td>
</tr>
</tbody>
</table>
Indicators for Annual Differentiation

Proficiency

• Percentage of student scoring at least “proficient,” “satisfactory,” etc.
  – Based on performance against goals

Other options pending ED clarification?

• Scale scores (South Carolina)
• Proficiency index (various states)
Indicators for Annual Differentiation

Ohio Performance Index: How to measure against goals?

<table>
<thead>
<tr>
<th>Performance Index Points</th>
<th>State Test Performance Level</th>
<th>Additional Points Awarded Over Lower Performance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3 points</td>
<td>Advanced Plus (Advanced score at higher grade level)</td>
<td>+ .1</td>
</tr>
<tr>
<td>1.2 points</td>
<td>Advanced</td>
<td>+ .1</td>
</tr>
<tr>
<td>1.1 points</td>
<td>Accelerated</td>
<td>+ .1</td>
</tr>
<tr>
<td>1.0 points</td>
<td>Proficient</td>
<td>+ .4</td>
</tr>
<tr>
<td>0.6 points</td>
<td>Basic</td>
<td>+ .3</td>
</tr>
<tr>
<td>0.3 points</td>
<td>Limited</td>
<td>+ .3</td>
</tr>
<tr>
<td>0 points</td>
<td>Did not take test</td>
<td>-</td>
</tr>
</tbody>
</table>

Average the index points received by each student for overall school measure score.

<table>
<thead>
<tr>
<th></th>
<th>Current</th>
<th>Targeted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student 1</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Student 2</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Student 3</td>
<td>0.6</td>
<td>1</td>
</tr>
<tr>
<td>Student 4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Student 5</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Student 6</td>
<td>1.2</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Index Score (Average) 87% 93%

Proficiency 50% 66%

Proficiency and participation more highly incentivized

Two potential AMO approaches

Average the index points received by each student for overall school measure score.
Participation Rate

- If the participation rate is lower than 95 percent, then the denominator of proficiency calculation must be 95 percent of enrollment (Section 1005(c)(4)(E)(ii))
  - Effectively assigns score of “0” to nonparticipants.
- Provide a clear and understandable explanation of how the State will factor [the requirement to test 95 percent of students] into the statewide accountability system (Section 1005(c)(4)(E)(iii))
Indicators for Annual Differentiation

Participation Rate and “Opt-Out”

- Rule of Construction on Parent Rights (Section 1005(b)(2)(K))
  - ESSA cannot preempt state or local law regarding parental rights regarding testing participation
- School districts must notify parents that they may request information about any state or local district policy that would allow parents the right to opt their child out of any assessments (Sec. 1112(e)(2)(A))
Indicators for Annual Differentiation

Student Subgroups

- Only disaggregated subgroup data may be used for the required accountability indicators (not combined subgroups)
- It is unclear if combined subgroups may be used outside of required indicators (i.e., low-stakes indicators)
  - Example: Use of combined subgroups for nonrequired college- and career-readiness indicators that do not historically have as high a percentage of disaggregated subgroup representation.
- SEA describes $n$ size for each subgroup
Indicators for Annual Differentiation

Graduation Rate (four-year adjusted cohort)
- Four-year adjusted cohort graduation rate (required)
  - Based on performance against goals
- Five-year adjusted cohort graduation rate (optional)

Other options pending ED clarification?
- Graduation rate index
- Highest of the four-year or extended-year graduation rate
Indicators for Annual Differentiation

Student Growth (Elementary Schools)

- Common state approaches
  - Value table (e.g., learning gains)
  - Student growth percentiles
  - Value-added models
  - Growth-to-standard

- Secretary prohibited from prescribing “…the specific metrics used to measure…growth”
Indicators for Annual Differentiation

Progress in Achieving English Language Proficiency

• Lack of consensus in field regarding whether this will measure:
  – Percentage of students making progress from year to year
  – Gains in percentage of students attaining English-language proficiency from year to year
  – Combination of the two

• Some current advocacy to allowing weighting of this indicator according to the size of EL population in the school
Indicators for Annual Differentiation

**School quality or success (nonacademic indicator)**

- “May” include the following:
  - Student access to and completion of advanced coursework
  - Postsecondary readiness
  - School climate and safety
  - Student engagement
  - Educator engagement

- Valid, reliable, comparable, statewide (for each gradespan, same indicator(s) used across schools)
- Must be able to disaggregate these data
Indicators for Annual Differentiation

School Quality or Success (Nonacademic Indicator)

Other reported data required under ESSA that might be used:

- Behavior data (suspensions, expulsions, etc.)
- Participation in AP/IB coursework and tests
- Preschool participation
- College-going rates
- Chronic absenteeism (absent one month)
Indicators for Annual Differentiation

School Quality or Success (Nonacademic Indicator)

Social-emotional indicators

CORE districts (California) incorporating student self-reporting in 2015–16 school index (8 percent weighting) for:

- Growth mindset
- Self-efficacy
- Self-management
- Social awareness
Indicators for Annual Differentiation

School Quality or Success (Nonacademic Indicator)

Social-emotional self-reporting (CORE)
Growth mindset dimension

Please indicate how true each of the following statements is for you:

10. My intelligence is something that I can’t change very much.
12. There are some things I am not capable of learning.
13. If I am not naturally smart in a subject, I will never do well in it.

(Not At All True, A Little True, Somewhat True, Mostly True, Completely True)

(CORE, 2015)
Selected New Reporting Requirements

- Achievement of accountability subgroups (Race, ED, SWD, EL)
  - State report card: progress towards meeting interim targets
  - Federal reporting: “disaggregated [achievement] results” for all subgroups
- Participation in AP/IB coursework and tests
- Preschool participation
- College-going rates within first post-secondary year
- Chronic absenteeism (absent one month)
- Rates of suspension, expulsion, arrests, violence, bullying
- Achievement and graduation rates of homeless and foster youth, and students with parent(s) in military
- English language proficiency acquisition rates
- Percentage of “inexperienced” and out-of-field/subject teachers
Summary

• While assessment requirement changes under ESSA are relatively peripheral, there are more significant shifts with respect to:
  – Accountability measures for the differentiation and identification of schools
  – Intervention strategies to support students in meeting State’s challenging academic standards
  – Reporting requirements
References


Robert Stonehill, Ph.D.
202-403-6210
rstonehill@air.org

David English
202-403-6930
denglish@air.org

1000 Thomas Jefferson Street, NW
Washington, DC 20007-3835
General Information: 202-403-5000
www.air.org