

In a State of Flux, Part Two: Opportunities and Challenges for Innovation for State Assessment and Accountability

PART ONE: COMMISSIONED BY THE HEWLETT FOUNDATION & PUBLISHED JANUARY 2018

PART TWO: COMMISSIONED BY THE WALTON FAMILY FOUNDATION & PUBLISHED SEPTEMBER 2019

BACKGROUND: With support from the William and Flora Hewlett, Bill & Melinda Gates and Charles and Lynn Schusterman Foundations, Education First, Future Ed and Bellwether Education recently published landscape scans on state assessment systems

education**first**



In a State of Flux: The State of State Assessments

FULL REVIEW: EXAMINING STATES' COMMITMENTS TO HIGH-QUALITY ASSESSMENTS—AND CONSIDERATIONS FOR SUPPORT

COMMISSIONED BY THE HEWLETT FOUNDATION FOR ITS EDUCATION PROGRAM GRANTEEES & COLLEAGUES | JANURARY 2018

education**first**



ABOUT THIS SCAN: In spring 2019, the Walton Family Foundation asked Education First to design an invitational grant program to support innovation in state assessment systems

In a State of Flux, Part Two

This comprehensive report summarizes the challenges and opportunities for innovation, based on:

- Interviews with state leaders and experts
- Analysis of problems that need to be solved
- Opportunities for innovation permitted by federal law
- Creation and analysis of innovation ideas for next-generation assessment systems

A State of Flux, Part Two:

Table of contents

1	United States Assessment & Accountability Landscape, 1960s- Today: Trends, Lessons and Cautions <ul style="list-style-type: none">▪ <u>How does the history of testing and stakes influence where we are today?</u>▪ <u>What is the “state of the state” of today’s assessment and accountability landscape?</u>▪ <u>What are the problems “end users” experience in current assessment and accountability systems?</u>▪ <u>What additional evidence suggests that innovation is needed?</u>▪ <u>What flexibility is there to innovate with assessments under ESSA?</u>	<u>Slide 6</u>
2	Initial Findings and Recommendations: Catalog of Possible Innovation Ideas for Funder Support <ul style="list-style-type: none">▪ <u>What are the options we’ve uncovered for the next generation of assessment innovation?</u>	<u>Slide 56</u>
3	Appendix	<u>Slide 73</u>



1 | United States Assessment &
Accountability Landscape,
1960s-Today:
Trends, Lessons and Cautions

How does the history of testing and
stakes influence where we are
today?

Some remain passionately against using assessments because they can reproduce structural inequity when used poorly

WHY ASSESSMENTS CAN REPRODUCE STRUCTURAL INEQUITY

Though assessment and accountability systems today are used to champion equity (see next slide), these systems arose from systems that perpetuated inequity for historically disadvantaged students.

For example, the standard IQ test used today is based on the work of Alfred Binet, a French psychologist, who developed intelligence tests as diagnostic tools to detect learning disabilities. During World War I, standardized tests helped segregate 1.5 million soldiers into different units by race and test score. According to researchers and media reports, “these tests were scientific yet they remained deeply biased.” This army test was later adapted for use in college admission and is now commonly known as the SAT.

“If we think about **cognitive ability testing as a form of lottery**, in which the winners are those who possess a certain inherent capacity for processing and analyzing information, **without reference to morally salient criteria like goodness, mercy, kindness or courage, we are embarking on a new kind of impoverishment...**No one has come up with a perfect solution yet — and no one will — **but that doesn’t mean we should give up trying.**”

— Thomas Edsall, New York Times (2019)

Yet parents, civil rights organizations and education reformers view assessments as essential to pressure school systems to increase the outcomes of historically-disadvantaged students

WHY ASSESSMENTS MATTER FOR HISTORICALLY DISADVANTAGED STUDENTS

Assessments provide teachers—as well as school, district and state leaders—the data needed to analyze student outcomes by subgroup. This is critical in order to close achievement gaps between historically disadvantaged students—such as students of color, low income students, English learners—and their more advantaged peers.

Assessment data is also critical because it enables education leaders to identify schools and districts that are underserving all students or certain subgroups of students. With comparable data, the state or district can provide additional support and improvement strategies to the school and/or district, hopefully resulting in improved outcomes.

Education equity means three things. One is that it means we are doing everything we can to ensure that education in America helps reduce inequality of opportunity. That means **low-income students, students of color, and English learners have to be getting high-quality educational opportunities.**

They have to have the same ability to compete as their more affluent peers.

–John King, The74 (2017)

Over the last 50 years, assessment & accountability reform has largely been driven by policymakers' desire to ensure *all* students receive a quality education, ready for their next step

Many interviewees reinforced that standardized assessments and comparable school ratings are needed to ensure all students, especially historically disadvantaged students, are served well

Interviewees also surfaced a tension advocates face today: the current assessment and accountability system isn't working, yet they're reluctant to advocate for changes because they don't know what to advocate for instead, and are wary of losing ground

"In almost all of the one-hundred 1:1s I had when I became executive director, I asked, "What's keeping you up at night? Where are we as a movement? What are we missing?" The answers were choice, assessment and accountability. **They are feeling like, in advocacy in particular, they are stuck in a place of having to defend a system that they know isn't really producing the right incentives or working.** But they do not know what's better, and so they are not able to have a conversation about what is better or where to go from here. There isn't a concrete, substantive policy package or a political strategy to get there. **They're afraid of losing ground entirely, so the conversations that are needed aren't being had at all."**

– National Thought Leader

"Standardized testing was never meant to be the ceiling; it was meant to be the floor. We are going to lose what the floor is and this is **problematic for the students who have been traditionally underserved** and students that did not reach the floor."

– Assessment Expert

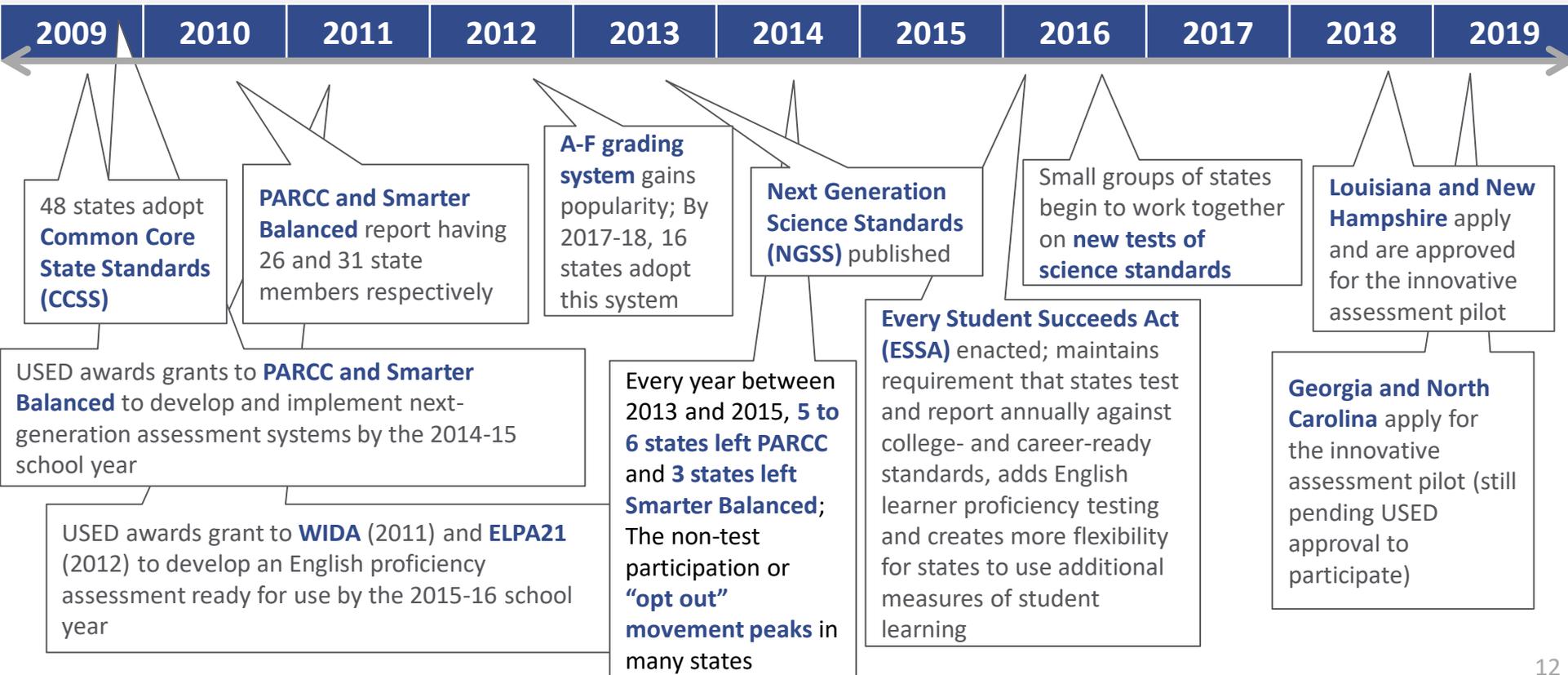
From the 1960s to the 2000s, assessment and accountability went from one extreme (low standards and student stakes) to the other (high standards, higher stakes for many actors)

1960s	<p style="text-align: center;">TITLE I ACCOUNTABILITY</p> <ul style="list-style-type: none"> ▪ 1965: Elementary and Secondary Education Act enacted: Establishes Title I, which ties federal funding to measured outcomes ▪ 1969: NAEP is implemented as an information source and neutral monitor with only nationwide findings for next 40 years
1970s	<p style="text-align: center;">MINIMUM COMPETENCY REFORM</p> <ul style="list-style-type: none"> ▪ 1970s: Several states began creating their own state assessment programs modeled after NAEP; by 1980, all states had a minimum competency testing program or a state testing program of some kind. By mandating state-administered tests and standards, legislators intended to improve the quality of schooling and “put meaning back into the high school diploma”
1980s	<p style="text-align: center;">BASIC SKILLS REFORM</p> <ul style="list-style-type: none"> ▪ 1983: <i>A Nation at Risk: The Imperative for Educational Reform</i> published; “competency examinations (now required in 37 states) fall short of what is needed, as the ‘minimum’ tends to become the ‘maximum,’ thus lowering education standards for all” ▪ 1980s: Pressure nationwide to enhance test scores; teachers begin teaching to the norm-referenced, multiple-choice tests; concerns about testing are raised (e.g. too much time and memorization)
1990s	<p style="text-align: center;">STANDARDS MOVEMENT</p> <ul style="list-style-type: none"> ▪ 1990s: National Education Goals and resulting standards movement called for the development of much more challenging standards and performance assessments ▪ 1994: ESEA reauthorized as Improving America’s Schools Act, which connected Title I funding to help students reach more rigorous statewide academic standards and mandated states must test at least once in elementary, middle and high school ▪ 1996: National Education Summit leads to the formation of Achieve, which finds over next 10 years that most state tests are still poor quality and measuring basic skills, not the more rigorous state standards or college and career readiness
2000s	<p style="text-align: center;">ACCOUNTABILITY FOR ALL</p> <ul style="list-style-type: none"> ▪ 2000s: Thomas B Fordham Institute, Achieve et al point out state tests not rigorous enough and proficiency cut scores widely vary ▪ 2002: ESEA reauthorized as No Child Left Behind Act, mandating annual state assessments in 3-8 and high school and requiring 100% proficient across all student subgroups by 2014 ▪ 2009: USED awards grants to state test consortia to develop and implement next-generation assessment systems by SY14-15

From 2010, high standards (Common Core) and quality assessments (consortia) attempted to solve for prior problems

NOTEWORTHY ASSESSMENT TRENDS from 2009 to 2019

- **Back to State-by-State:** While 46 states originally joined PARCC, Smarter Balanced or both, states' participation has eroded since 2010. As of late 2017, only 21 states report they are using items from either PARCC or Smarter Balanced.
- **Rise of SAT/ACT:** By spring 2017, the ACT or SAT assessments were mandatory for all high school students in 24 states. In 13 of these states (including six Smarter Balanced or PARCC states), ACT or SAT results also were used for school accountability decisions.



School accountability also evolved over time; most notable was the shift from NCLB to ESSA in 2016, which gave states some autonomy (within guardrails) to design accountability

Policy	No Child Left Behind Act (NCLB)	Elementary and Secondary Education Act (ESEA) Waivers	Every Student Succeeds Act (ESSA)
Requires accountability based on adequate yearly progress (AYP) toward the goal of having 100 percent of students reach proficiency in math and English	Yes	No	No
Requires state accountability system with more than two indicators	No	Permits	Yes
Requires federally-determined interventions in low-performing schools	Yes (i.e., school choice and supplemental educational services)	Yes (within the lowest-performing 5 percent of Title I schools and high schools with a graduation rate	No



Under ESSA, states have the autonomy to develop achievement goals, ratings indicators and improvement supports strategies

Even with ESSA's greater flexibility, the statute still details requirements for annual/long-term goal-setting, disaggregated data, specific multiple tests and mandated ratings

1

States **set their own achievement goals**

- No more “100% proficient” with annual targets: **States set their own goals from their own starting points.**
- No more “Adequate Yearly Progress,” or AYP: states are required to set **statewide, long-term goals and interim progress targets** for improving outcomes for **all students and each student group** (e.g., race/ethnicity, income, students with disabilities, English learners, homeless, foster and military youth).
- States set **four-year cohort graduation rate goals** with interim progress targets. States may set higher extended goals.

2

States **choose indicators** for school ratings

- States **choose at least 4 indicators**, with the first 3 getting a “substantial” and, collectively, “much greater weight” than the 4th:
 1. **Academic achievement**—including at least math and reading proficiency
 2. **Another academic indicator**—must include HS cohort graduation rate; for Elementary/Middle Schools can be growth
 3. **English language proficiency** for English learners
 4. **At least one other indicator of school quality or student success**— e.g., chronic absenteeism, school climate, social-emotional learning— that must be valid, reliable and available statewide for all subgroups

3

States and districts **determine improvement supports**

- States must identify schools and provide support and intervention to at least 2 categories of schools:
 1. **Comprehensive Support and Improvement Schools:** lowest-performing 5% of Title I schools and HS with graduation rates below 67%. Identified every 3 years. **State approves improvement plan.**
 2. **Targeted Support and Improvement Schools:** annual list of schools with one or more student subgroups consistently underperforming on ALL state indicators. Schools in the bottom 5% of subgroup performance for ALL indicators must identify resource inequities. **District approves plan.**

What is the “state of the state” of today’s assessment and accountability landscape?

Education First (2018) and FutureEd (2019) have synthesized the current state of state assessments

Membership in the Consortia continues to decline

- Membership to PARCC or Smarter Balanced has gone from 46 states in 2010 to 20 states today
- There's evidence that both the quality and rigor of state tests have risen as a result of PARCC and Smarter Balanced testing

More states are going at it alone on 3-8 assessments

- The marketplace for summative assessments in grades 3-8 is fragmented
- As of the 2017-18 school year, the largest vendors holding state testing contracts in these grades were AIR (9); DRC (7), Pearson (5) and Measured Progress (4)

When it comes to innovation, there's more talk than action

- There's interest among states to do something different, but there's still more talk than action
- To date, two states—LA and NH—have been approved for the federal innovative assessment pilot and two more—GA and NC—have applied but have yet to receive approval

Churn in state testing systems is a big concern

- Assessment experts note many states have changed their tests or assessment vendors multiple times in the past five years
- Constant changes in state assessment systems create a host of problems (e.g. harder to track performance overtime, create problems with accountability, etc.)

SAT/ACT are on the rise in high schools

- More states are embracing the ACT and the SAT as their high school assessments despite concerns that the tests are not fully aligned with state standards
- As of 2017-18, 27 states require the ACT, SAT or choice of ACT, but only 12 states use these assessments for accountability purposes

Comparing across states decreased in priority

- While states and the federal government still prioritize the ability to compare student test results within a state, comparing results across states has faded as a dominant goal

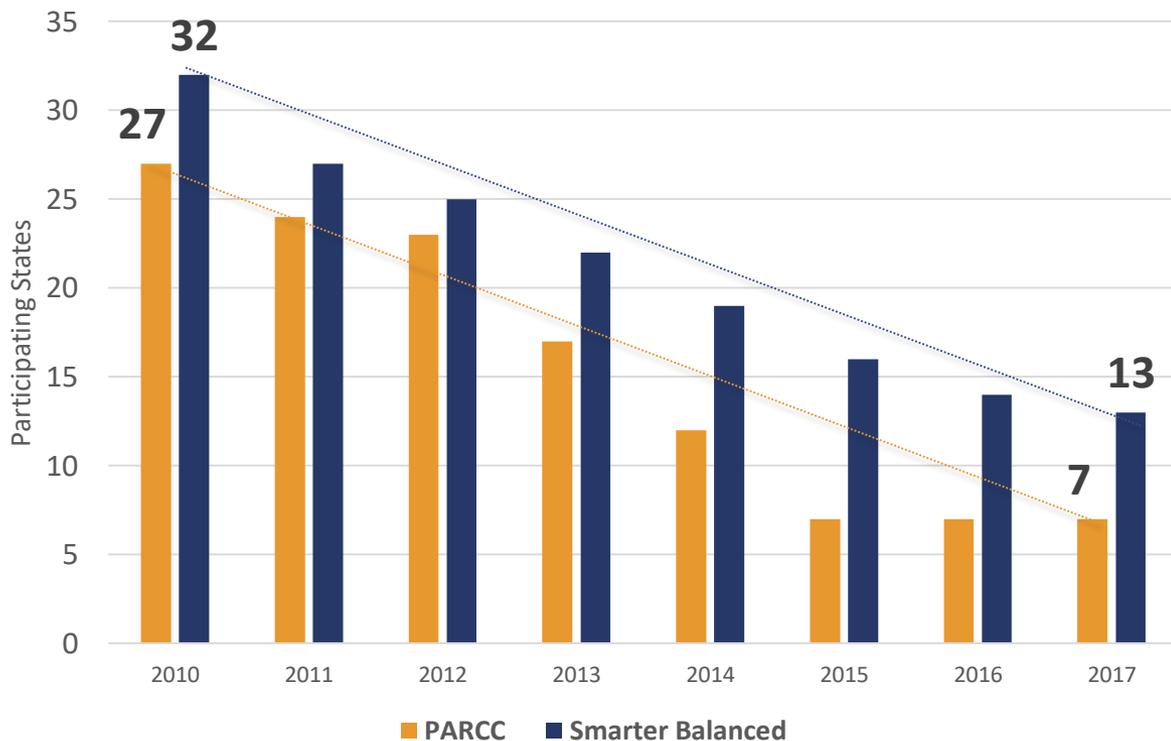
**The next slides provide more information on each of these themes*

As of 2018, PARCC/SBAC consortia assessments are rated highly on USED peer reviews, but membership is low

46 states originally joined PARCC, Smarter Balanced or both consortia, but states' membership or participation continues to fall, especially in PARCC

- In recent USED peer reviews, PARCC and Smarter Balanced assessments did better than independently developed tests: 12 of 13 (92%) consortia states substantially met criteria vs. 7 of 16 (44%) non-consortia states—plus Maryland fully met all criteria
- Total membership has gone from 46 states in 2010 to 20 states in 2017. 16 of these states fully participate in PARCC or Smarter Balanced, while an additional four states populate their assessments with consortia items. Smarter Balanced is 1-1 on adding new members this year (IN will use item bank; IA governor over-ruled participation.)

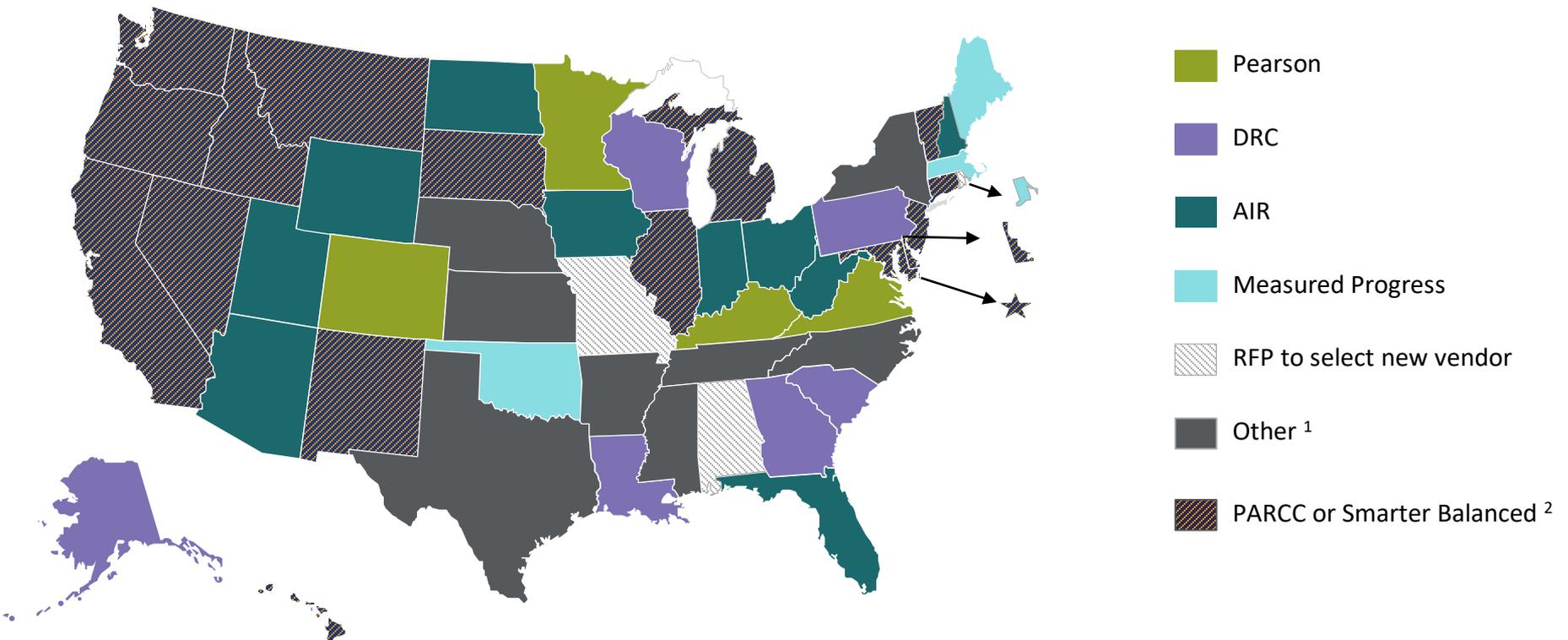
Consortia membership in PARCC and Smarter Balanced (2010 to December 2017)¹



Sources: Education Next, [The Politics of the Common Core Assessments](#) (2016); Education Week, [Which states are using PARCC or Smarter Balanced?](#) (2017) Education Week; Education Week, [State Testing: An Interactive Breakdown of 2015-2016 Plans](#) (2016); [PARCC](#); [Smarter Balanced](#); Education First interviews.

Today, most states are independently sourcing and managing their grade 3-8 tests

As of 2017, four vendors have won the majority of contracts to create new state K-8 assessments: American Institutes for Research (10), Data Recognition Corporation (6), Pearson (4), and Measured Progress (4)



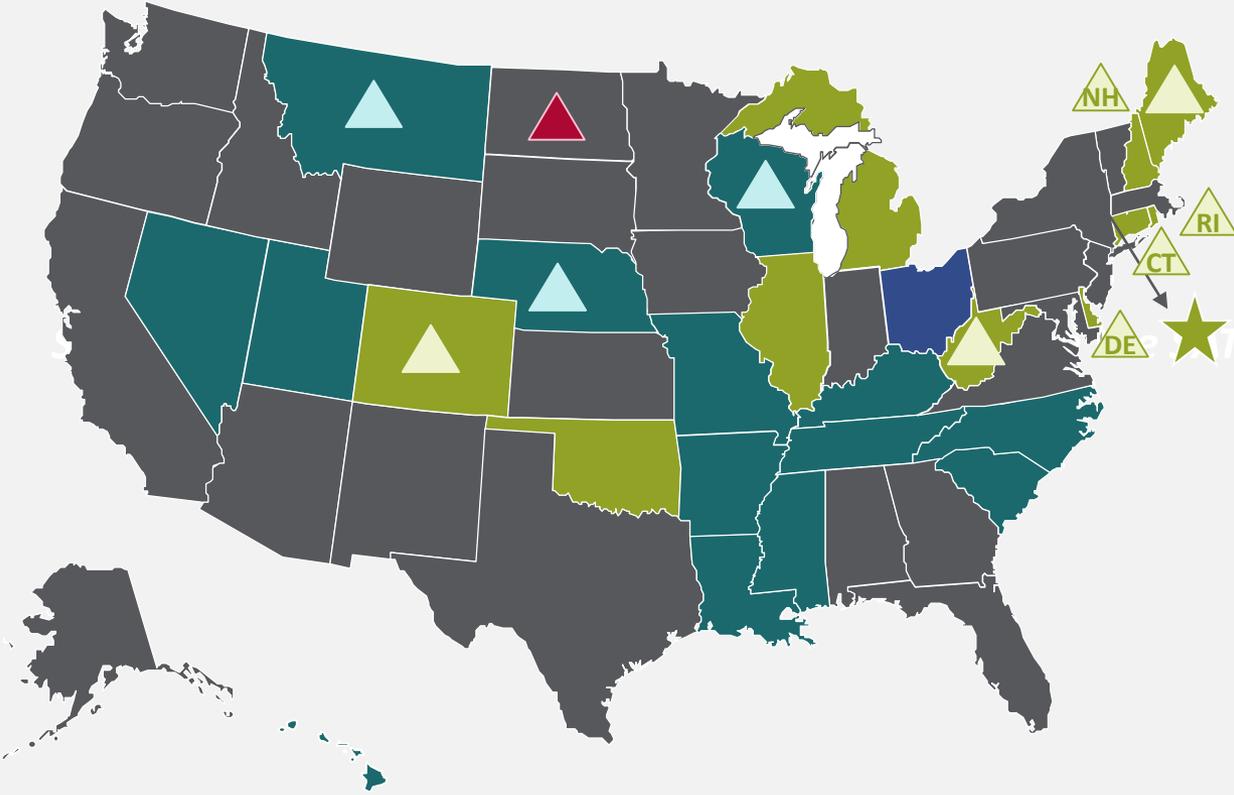
¹ Other includes: ACT, SEAs, ETS, NWEA (for 2018-19 school year), Questar, University of Iowa and University of Kansas; ² Since Louisiana and Massachusetts use a mix of PARCC items and their own items, they work with separate vendors other than PARCC's main vendor (Pearson). In Oct 2017, Indiana chose AIR as the vendor for its new state assessment; it plans to use Smarter Balanced items. Although it uses a mix of items, Michigan does use Smarter Balanced's main vendor.

Sources: SEA websites; State assessment directors; Education Week, [Which states are using PARCC or Smarter Balanced?](#) (2017).

SAT and ACT are on the rise in high schools, though most states do not use these tests for accountability purposes

Twelve states replaced their high school summative assessment with ACT, SAT or district choice of ACT

States that required all high school students to take the SAT and/or ACT in SY17-18



14 States Require ACT	
	3 states used ACT for accountability purposes: MT, NE, WI
	11 states required ACT in addition to another state assessment: AR, HI, KY, LA, MS, MO, NC, NV, SC, TN, UT
11 States Require SAT	
	8 states used SAT for accountability purposes: CO, CT, DE, IL, ME, NH, RI, WV
	3 states used SAT in addition to another state assessment: DC, MI, OK
1 State Requires SAT or ACT	
	1 state used SAT or ACT (districts choice) in addition to another state assessment : OH
1 State Offers Districts the Flexibility to Use the ACT Instead of the High School State Assessment	
	1 state is approved by USED to let districts use the ACT in place of the standard state assessment: ND

*See appendix C for more information on the high school assessment landscape

High School Assessments: States that required at least one end-of-course high school summative assessment in SY16-17

Twenty-two states required at least one, statewide end-of-course high school assessment in ELA and/or Math in the 2017-18 school year:

- Arizona
- DC
- Florida
- Georgia
- Kentucky
- Louisiana
- Maryland
- Mississippi
- Montana
- Nevada
- New Jersey
- New Mexico
- New York
- North Carolina
- Ohio
- Pennsylvania
- South Carolina
- Tennessee
- Texas
- Utah
- Virginia
- Washington

See [here](#) for more information on high school math assessments and [here](#) for more information on ELA assessments.

High School Assessments: Similar to grade 3-8 assessments, the majority of high school assessments are state developed

State required Math & ELA Assessments (2018-19)

Assessment**	Number of States	States
State-developed assessment(s)	27	Alaska, Arizona, Florida, Georgia, Indiana, Iowa, Kansas, Kentucky, Louisiana, Massachusetts, Minnesota, Mississippi, Missouri, Nevada, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Pennsylvania, South Carolina, Tennessee, Texas, Utah***, Virginia, and Wyoming
Smarter Balanced	7	California, Hawaii, Idaho, Oregon, South Dakota, Vermont, and Washington
PARCC	2	District of Columbia and Maryland
ACT Aspire	2	Arkansas and Wisconsin
Pre ACT	2	Alabama and North Carolina
ACT	13	Alabama, Hawaii, Kentucky, Louisiana, Mississippi, Montana, Nebraska, Nevada, North Carolina, North Dakota, Utah, Wisconsin, and Wyoming
PSAT 8/9	2	Illinois and Michigan
PSAT 10	5	Colorado, Delaware, Illinois, Michigan, and Rhode Island
SAT	9	Colorado, Connecticut, Delaware, Illinois, Maine, Michigan, New Hampshire, Rhode Island, and West Virginia
ACT or SAT (district/student choice)	4	Arizona, Ohio, Oklahoma, and Tennessee

All states are administering a state-developed assessment in science, with five exceptions: Alabama (ACT), Arkansas (ACT Aspire), Nebraska (ACT), West Virginia (SAT), and Wisconsin (ACT Aspire and ACT).; **Categories are not mutually exclusive.; *** Utah's assessment is 50% ACT Aspire/50% Utah-developed items.

High School Assessments: How do assessments signal readiness for postsecondary success?

Then—2013 ¹	Now—2019
<p>Three states administered high school end-of-grade or end-of-course assessments developed by K–12 and higher education leaders and had a statewide postsecondary policy for placing students into first year, credit-bearing courses based on these assessments: California, Florida, and Georgia</p>	<p>Seven states administer high school end-of-grade or end-of-course assessments developed by K–12 and higher education leaders and have a statewide postsecondary policy for placing students into first year, credit-bearing courses based on these assessments: California, Hawaii, Minnesota, Oregon, South Dakota, Texas and Washington</p>
<p>Three states that administered ACT or SAT to all students had a statewide postsecondary policy for placing students into first year, credit-bearing courses: Hawaii, Kentucky, and Louisiana</p>	<p>Ten states that administer ACT or SAT to all students have a statewide postsecondary policy for placing students into first year, credit-bearing courses: Colorado, Illinois, Maine, West Virginia (SAT); Kentucky, Louisiana, Nevada (ACT); Ohio, Oklahoma, Tennessee (ACT or SAT)</p>

See [here](#) for more information

Why have the state consortia for special education and English leaders maintained membership, while PARCC and Smarter have declined?



- Under ESSA, English language proficiency growth is included in school quality ratings, but this wasn't the case in the past, which allowed ELPA to build their brand and users confidence in it, potentially making the shift to accountability easier
- Initiatives for historically disadvantaged students and overlooked issue areas may have an easier time attracting and keeping state participants
- Benefits of working together have to be clear and strong; consortium has to permit some flexibility

- The aspirations and the reality were mismatched (e.g., a system of assessments including through-course/interims did not materialize)
- States (and other users) care more about intrastate comparability (within) than interstate comparability (across)
- SEAs drove the consortia; policymakers, parents and educators didn't understand the "why" or even the "what"
- Benefits of working together have to be clear and strong; consortium has to permit some flexibility

Many states—including Utah and Tennessee—have struggled to find assessment vendors that live up to their contractual obligations

EDUCATION

Utah Cancels \$44 Million Contract With School Test Company

BY BRADY MCCOMBS, ASSOCIATED PRESS
JUNE 7, 2019 AT 1:37 PM

Share ↗

- Spring 2018: Utah signed a \$44-million contract with Questar Assessment Inc.
- June 2019: Utah officials abruptly canceled the state's contract with Questar due to *"a flurry of technological glitches that have created uncertainty about whether this year's test scores will be validates."*

"...few vendors have the expertise and capacity to handle large-scale jobs that require building, administering, and scoring computerized tests for older students and paper versions for younger ones."

– Marta Aldrich, Chalkbeat

HELP WANTED

Will third time be a charm? Tennessee searches again for online testing company

BY MARTA W. ALDRICH, JUNE 14, 2019

- 2016: Tennessee canceled a \$108 million contract with Measurement Inc. because of *"repeated failures, including the inability of students to go online to take the test. There also were problems getting paper assessments shipped to schools on time."*
- Summer 2016: Tennessee hired Questar to take over as the state's assessment vendor. Questar struggled with computerized scoring and delivering, *"culminating last spring with days of technical disruptions that seriously undermined the test's credibility with students, parents and educators."*
- June 2019, Tennessee signed a new five-year contract with Pearson, at cost of ~\$20M in SY19-20 and ~\$17.3M each year 2-5 for total cost of ~\$90M.

Under ESSA, all states must develop and rate schools based on at least five accountability measures

ESSA requires at least five separate measures, which may consist of multiple indicators

Measure of **academic achievement** using annual assessments in **English language arts** and **math**^{2,3}

A measure of **school quality** or **student success**^{1,2,3}

4-year **graduation rate** for high schools (extended year rate may also be used)²

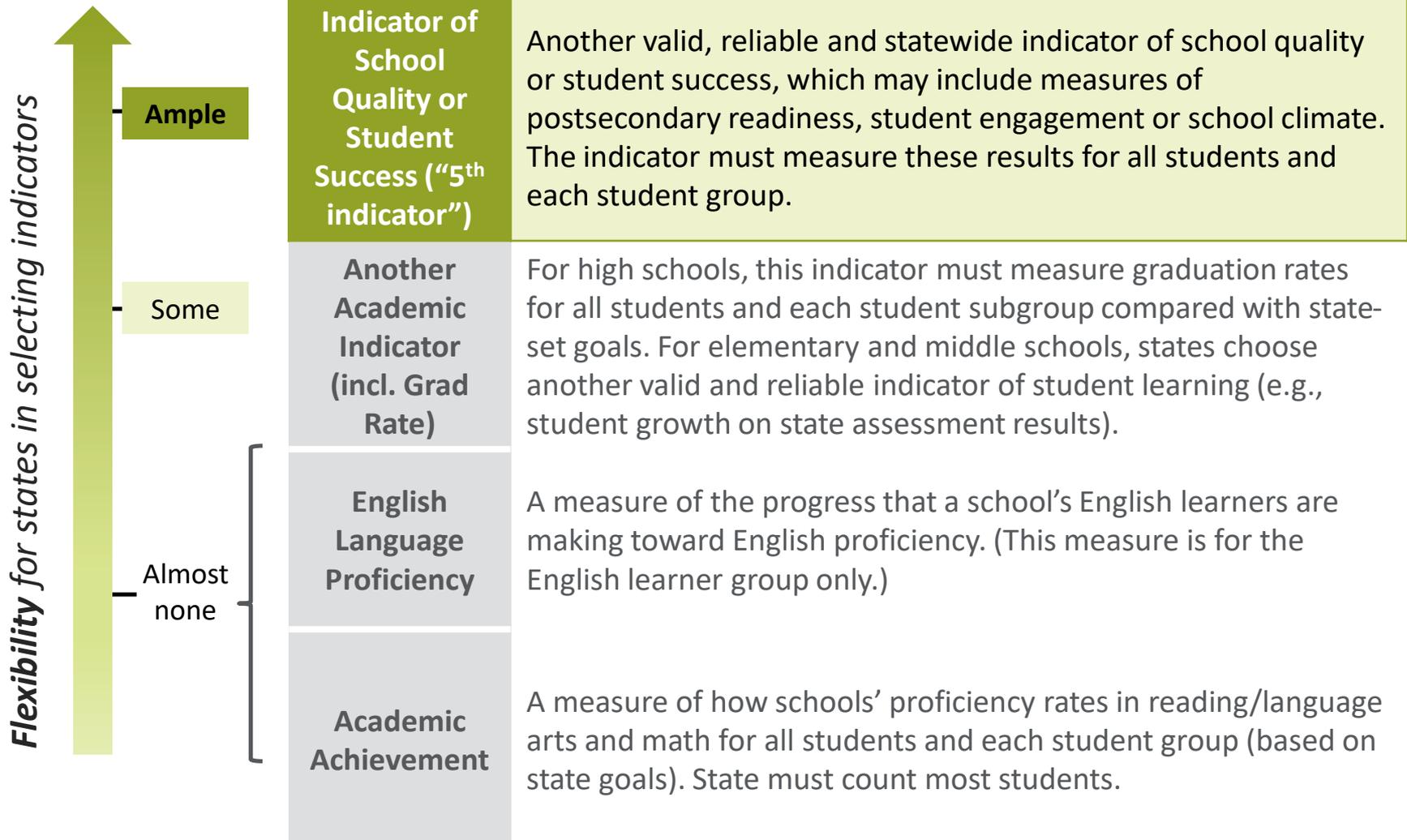
An additional academic measure for elementary schools, which can be a measure of **student growth**²

A measure of **progress in language proficiency** for English language learners²

¹May use multiple indicators; ²May use multiple measures; ³May use measures that show varied results
Source: United States Congress (2015); Conley (2017).

ESSSA allows states flexibility to select one of the required accountability indicators

Required Federal Accountability Indicators under



As a result, almost all states incorporated a growth indicator in at least K8; few states chose indicators beyond chronic absenteeism or graduation rates

STUDENT GROWTH INDICATORS (optional)

47 states & D.C.

include growth as an accountability indicator in K8

20 states include

growth as an accountability indicator in high school

21 states use a science

proficiency/progress measure

9 states use a school climate/culture measure

9 states use a social studies proficiency/progress measure

6 states & D.C. use an art access/participation or well-rounded education measure

The majority of states are using college/career readiness and/or chronic absenteeism as their “5th indicator”

SCHOOL QUALITY AND STUDENT SUCCESS INDICATORS (required)

35 states include a college and/or career readiness measure

35 states use a chronic absenteeism/attendance measure

21 states use a science proficiency/progress measure

9 states use a school climate/culture measure

9 states use a social studies proficiency/progress measure

6 states & D.C. use an art access/participation or well-rounded education measure

Greater flexibility was a key goal and several ESSA assessment provisions improve on NCLB, but ESSA's statutory requirements for assessment are still fairly restrictive

Under ESSA, all states must...

- Establish college and career ready standards and maintain high expectations when assessing students against those standards
- Assess all students; offer appropriate accommodations for English learners and students with disabilities; use principles of universal design for learning to the extent possible
- Disaggregate assessment results by the following subgroups: each major racial and ethnic group; economically disadvantaged students compared to students who are not economically disadvantaged; children with disabilities as compared to children without disabilities; English proficiency status; gender; and migrant status
- Develop tests that measure higher-order thinking skills including reasoning, analysis, complex problem solving, critical thinking, effective communication, and understanding of challenging content
- Assess no less than 95% of students in each school, as well as 95% of each subgroup; Establish a policy to govern parents "opting" their children out of state assessments

Additionally, states may...

- Develop computer-adaptive tests, as long as the assessment must measure and report assessment results against grade-level academic standards
- Administer an end-of-course high school math assessment to 8th grade students taking advanced math courses instead of the 8th grade test

The actual statutory language further constrains what states can do

According to ESSA, all state assessments must be...

"..**accessible** to English learners and students in special education. If they need accommodations, like extra time, they should get it."

"..aligned with **challenging State academic standards**, and provide **coherent and timely information** about **student attainment** of such standards and whether the student is performing at the student's grade level."

"...aligned to standards that **get kids ready for postsecondary education or the workplace.**"

"...**of adequate technical quality** for each purpose required under this Act and consistent with the requirements of this section, the evidence of which shall be made public, including on the website of the State education agency."

"..**valid and reliable, consistent with relevant, nationally recognized professional and technical testing standards**, objectively **measure academic achievement, knowledge, and skills.**"

"..tests that **do not evaluate or assess personal or family belief and attitudes** or publicly disclose personally identifiable information."

...which puts significant guardrails and constraints on assessments

- These requirements serve important purposes: privacy, access, equity, rigor, quality, utility
- But some of these requirements put significant limitations/strains on innovation, particularly
 - **...of adequate technical quality**
 - **...valid and reliable**, consistent with relevant, nationally recognized professional and **technical testing standards**

What are the problems “end users” experience in current assessment and accountability systems?

In general, systems of assessment attempt to serve two primary goals that often conflict with each other



Assessment Of Learning

Measurement to *compare* student achievement results.

Used to:

- Sort or select students for placement, grouping, or exit or entry
- Assess and compare the effectiveness of programs, schools and districts for audit, improvement or accountability purposes

Measurement to *improve* student achievement results.

Used to:

- Monitor student progress dynamically
- Diagnose student learning strengths and challenges and inform instruction strategies and choices
- Provide students with engaging assessment experiences that create actionable feedback and a clear picture of their performance relative to long-term goals
- Directly and immediately improve the education outcomes of the individuals assessed



Assessment For Learning

Most users—even accountability advocates—acknowledge persistent problems

Key constraints in today's testing landscape:

Too much **time** spent on standardized assessments

High **cost** to assess all students annually and disaggregate

Rightly or wrongly, educators and parents perceive **little value** in end of year tests

Stakeholders, especially parents, are concerned with how results will be used / **data privacy**

The focus on **validity & reliability** has inhibited assessment innovation/quality

Although **assessment technology** improved, it's still limited

While most OECD countries have sweated validity (good measures of what's important), the **US has been preoccupied with reliability** (inexpensive measures of what's measurable). The development of Common Core State Standards was a national effort to raise expectations and implement better tests. The addition of more writing made the tests longer and just added to the backlash against testing.

—National Thought Leader

Not all schools have the **infrastructure** to support assessment technology

States/ESSA bound tests to **APA testing standards**

In spring 2019, we interviewed 30 thought leaders, assessment innovators and state leaders, who largely confirmed assessment & accountability systems need to change (1 of 3)

Summative assessments aren't aligned well enough to curriculum

I'd love to get the place where there is **more formative rolls up to the summative, so there isn't a disconnect between the curriculum.** It's supposed to work with aligned standards and aligned test and the two meet to measure content and mastery. That would be common sense. But it doesn't happen in the real world. – Leading Edge Thinker

Assessments don't always accommodate the diverse array of learners

Overall, it'd be great if **assessment systems continue down the road of individual needs, not group needs.** It gets at heterogeneity at ELs and SPED. When we looked at those groups as individual groups, **we really missed the mark about how diverse these groups were.** And we totally missed ELs with special needs. – Special Populations Assessment Expert

Summative assessment results lack meaning to parents, students and teachers

Families want something better as a measure - and not something where "success" depends on being one standard deviation in this direction compared to others. – Leading Edge Thinker

Advocates want something different – and innovators want total transformation (2 of 3)

Even advocates who enacted/defended the current system are calling for overhaul....

I asked 100 members in the last three months **“where are we as a movement really vulnerable?”** The answers were choice and assessment/accountability. **They feel stuck: they have to defend a system they know isn’t really producing the right incentives or working,** but they don’t know what’s better, and they’re not able to have a conversation about where to go from here. **There isn’t a concrete, substantive policy package or a political strategy to get from here to there. They’re afraid of losing ground entirely, so the conversations aren’t being had at all.** But they know we need this. – Advocacy leader

...And some believe this overhaul needs to be a total transformation rather than reformed

If we’re right that this is about transformative change and not reform, **the things you would do to reform the system are literally the opposite of what you’d do to transform it.** Assessment is interesting: if you want A to result in B, they you do more A to get to B. Assessments play a more solidifying role once system starts to move but assessment shouldn’t lead it. Leaves me to think about whether Walton thinks of this **as assessment for accountability, assessments for learning, or other.** – Leading Edge Thinker

Regardless, interviewees acknowledged how transforming assessments is difficult under the current accountability constraints

Chicken and egg problem: We can’t do much differently about assessment when you’re constrained by accountability and vice versa. – Assessment Expert

We might **gear accountability much more toward life outcomes instead of math and reading** that we’ve been using as proxies. – Thought Leader

States will face stiff headwinds re-envisioning summative assessments, and new mindsets are needed if leaders are to have a shot at success (3 of 3)

To successfully innovate, state and district leaders need to be willing to step outside of their current boxes; focus needs to shift from what is to what could be

If we're trying to create a collaborative effort of linking education to careers in certain cities and states, we could create philanthropic prototypes that inform policy and systems change over time. **There's a risk of going too fast to statewide policy/scale.** We're still learning. It's appealing to consider trying this out as part of some program work, supporting some prototypes, with explicitness about trying it to see if it could be rolled out over time statewide. – Funder

Innovating to create new assessments also requires re-examining lay aspects of current K-12 accountability system

Does Walton want to solve for the accountability challenge? **It's the accountability challenge that is driving everything. You can't pull these things apart.** – Leading Edge Thinker

The goal should be transformative thinking as opposed to re-trying more traditional reform ideas

If we're right that **this is about transformative change and not reform**, the things you would do to reform the system are literally the opposite of what you'd do to transform it. – Leading Edge Thinker

A few interviewees articulated specific redesign and transformation assessment or accountability ideas (1 of 2)

Our statewide assessments now are the least precise instrument to figure out if schools are preparing students for life. Innovation really needs to **focus on whether we can measure the other aspects of education that will lead to the long term outcomes: success in life, civic engagement, mental health, soft skills development.** No one has a good sense for measuring and holding schools accountable for these. So instead we look at the lowest common denominator for school improvement. – Advocacy Leader

Could we look at some of the upskilling programs around the country that are connecting adults to good jobs and **figure out what are the entry-level standards for certain careers – and then create specific assessments for particular careers/sectors?** It has to be really work-based. – Funder

You have to **define the purpose first**, such as “prepare kids to live a fulfilling life.” There’s still the basics in academics; then are you preparing students with experiences and diversity of opportunity. You also want to assess aspects of student autonomy (how to learn who I am and how to make choices). – Leading Edge Thinker

We hear [from our members] they want to focus on **postsecondary outcomes, school climate data, preK-2 and early childhood data.** No one is ready to build that today. It would require a lot of research and a lot of systems building to get people prepared for a new accountability plan. – Advocacy Leader

Holistic and comprehensive assessment systems measuring skills above and beyond math and ELA that are connected to success long term

A few interviewees articulated specific redesign and transformation assessment or accountability ideas (2 of 2)

Retain key ideas but consider new approaches

I don't think there are 5 states that would keep annual testing if ESSA changes. I'm worried no matter when ESSA is reauthorized that annual testing won't stay. Testing is critical to make sure kids aren't falling through the cracks. We've done a listening tour and have **several policy ideas for states to either make the current systems more parent-friendly, or put in place measures for the future that could round out an accountability system so it's not completely test-based.** Our next phase is to **go deeper into other measures that might serve as good proxies because they're correlated with test scores,** like the 5 Essentials from Illinois. – Thought Leader

Create graduate profiles and assess with multiple measures

The Graduate Profile movement is a big step. If the profile was developed in a community in the right way (i.e., not just by central office cabal), there is great buy-in. If we have a client who has a grad profile, we can hold it up and help us push districts in the right direction. **You have this vision and profile, now what are you doing to make it a reality?** – Assessment Innovator

Compile and use cumulative measurements completed over time

We are interested in identifying approaches to assessing kids that you can do in a more continuous way. Could you in a **slightly lower stakes way ask a few questions per day woven into the daily routines of school that would have a cumulative measurement capacity** better than current instruments? And have the available data on daily basis useful to schools and less likely to have measurement flaws particular to a typical single testing day. – Assessment Innovator

We also interviewed many students, parents, teachers, principals and regional leaders from a major CMO, all of whom underscored these challenges (1 of 2)



School Leader

For our students, resilience and confidence matter—building up confidence to know they are enough. We used to do a better job of looking at character and academics. The pendulum has now swung to academics only... **If we have an equity focus, students absolutely deserve this rigorous academic piece. But our kids also live in a society where they have to be able to navigate the world. Being smart is not enough for black and brown kids.** As a school leader I feel like I'm failing every day because I can't take the lead on character anymore, because how I'm told I'm successful is how students are performing on academic tests... **What feels inequitable is when you go to affluent schools, kids have the opportunity to explore themselves as whole people. Right now, this focus on testing feels like militarism.**



Parent

I would say **stop testing them so constantly**, and get them prepared over a longer period of time so they feel confident. I want my son to feel confident and not so nervous about new learning.

It's back-to-back. We take tests every week. We take tests—say we got a test in every single class. Math, reading, and science. Then we got to write an essay. Then another essay in social studies.



Student

We also interviewed many students, parents, teachers, principals and regional leaders from a major CMO, all of whom underscored these challenges (2 of 2)



Education Director

State assessment is the one thing that will dictate everything else. I don't care about MAP or other assessments. Nothing else matters... PARCC is unfortunately the end-all, be-all. It's better for us to look good locally than nationally.



Chief Academic Officer

So much is bound to the state accountability system. What I care about most is what our schools are doing on the [A-F state] letter grade, in comparison to the neighboring district schools. I am interested in how our AP pass rates compare when you get into HS.



Regional Leader

We've heard from our leaders and teachers that we are testing all the time. **Last year, our schools did not test on only nine days in the entire year.** If you are an EL student, you have had no break from assessments... I don't know what to cut, but we need to cut.

What additional evidence suggests that innovation is needed?

Evidence suggests using multiple measures provides greater opportunity for students to demonstrate college readiness and, in particular, benefits women and students of color

The Center for Analysis of Postsecondary Readiness (CAPR) conducted a randomized assignment trial to test the impact of a multiple measures placement system across 7 community colleges in the State University of New York (SUNY) system (n=13,000 students).

Control Group: Students were assigned to remedial coursework based on a placement exam score administered by the community college.¹

Program Group: Students were assigned to remedial coursework based on **multiple measures**, including both placement test scores *and* high school GPAs

- **In math, 14% of program students placed higher** than they would have under a test-only system, while 7% placed lower.
- **In English, 41.5% of program students placed higher**, while 6.5% placed lower.
- **Women appeared to benefit more than men** from program group assignment process in math more than men
- **Black and Hispanic students appeared to benefit more than White students** from the program group assignment process in English

¹Universities determine who must take the placement exam based on prior scores on the Regents exam, AP exams and/or the SAT. Source: Barnett et al. (2018).

State assessments typically do not measure skills needed for the continually-evolving jobs of the future/the knowledge economy

KEY

Skills most standardized assessments do measure

Skills most standardized assessments do not measure

21st Century Skills Needed

Critical thinking	Complex problem solving	Judgment and decision-making
Memorization	Number facility	Getting information
Fluency of ideas	Information ordering	Thinking creatively
Estimating quantifiable characteristics	Analyzing data & information	Interacting with computers
Documenting/recording information	Updating and using relevant knowledge	Being exact or accurate

Inductive reasoning

Deductive reasoning

Research also points to numerous influential factors related to student achievement that tend to be measured only in classrooms or not at all

Hattie's (2008) Visible Learning research synthesizes findings from 1,400 meta-analyses of 80,000 studies involving 300 million students.

Potential practices and measures education systems could use to assess student readiness, in addition to summative assessments:

- | | |
|---|--|
| <ul style="list-style-type: none">■ Evaluation and reflection (high, 0.78)■ Formative evaluation (medium, 0.48)■ Goals (medium, $d=.57$)■ Feedback (student to teacher, high, $d=.73$)■ Motivation (medium, $d=.48$)■ Persistence/engagement (medium, $d=.48$) | <ul style="list-style-type: none">■ Practice testing (medium, $d=0.54$)■ Self-concept (medium, $d=.43$)■ Parental involvement (medium, $d=.51$)■ Study skills (high, $d=.59$)■ Metacognitive strategies (high, $d=.69$)■ Teacher professional development (high, $d=.62$) |
|---|--|

**For a comprehensive list of all factors influencing student achievement, see [here](#).*

What flexibility is there to innovate with assessments under ESSA?

ESSA provides all states with the option of implementing two types of assessment flexibility: through-grade assessments and district choice with high school assessments

	Through Grade Assessment Flexibility	District Choice with High School Assessment
Assessment flexibility under ESSA	<ul style="list-style-type: none"> States can use multiple assessments through an academic year, rather than just a single summative assessment 	<ul style="list-style-type: none"> States can allow individual districts to drop their state's high school exam and use a "nationally recognized" high school test such as the SAT or ACT instead <p><i>*Note, all states have the option of using SAT or ACT as their statewide high school assessment if it aligns to their state standards</i></p>
Expressed Interest	<ul style="list-style-type: none"> As of Spring 2018, zero states appeared to be seriously considering this opportunity ¹ 	<ul style="list-style-type: none"> Georgia, Oregon, Florida, California, Colorado
Denied by USED	<ul style="list-style-type: none"> As of Spring 2018, zero states applied for this flexibility¹ 	<ul style="list-style-type: none"> Arizona, Oklahoma
Approved by USED		<ul style="list-style-type: none"> North Dakota
More Information	<ul style="list-style-type: none"> See slide 7 for barriers preventing states from implementing through-grade assessments 	<ul style="list-style-type: none"> See slide 8 for a list of states that already use the ACT and/or SAT statewide

¹Based on our research, it does not appear that any states implemented through-grade assessment in SY18-19

ESSA's innovative assessment pilot allows states to experiment with new measures of student learning

Overview of ESSA's Innovative Assessment Pilot

- Up to seven states and/or consortia¹ can try out new forms of testing in a handful of districts, with the goal of eventually taking them statewide
- States **do not receive additional federal funding** for participating in the pilot
- **Approved states get five years** to:
 - develop their assessment systems
 - show tests are comparable to the state assessment
 - bring the whole system statewide
 - **If needed, **states can ask for a two-year extension***
- The Institute of Education Sciences will publish a **report detailing the initial progress** (2-3 years after state implantation begins) of innovative assessment systems carried out

ESSA's Definition of an "Innovative Assessment"

- "**Competency-based** assessments, **instructionally embedded** assessments, **interim** assessments, **cumulative year end** assessments, or **performance-based** assessments that combine into an **annual summative determination** for a student, which may be administered through **computer adaptive** assessments; and
- Assessments that **validate when students are ready to demonstrate mastery or proficiency** and allow for **differentiated student support** based on individual learning needs."

– Every Student Succeeds Act (2015)

While the ESSA pilot provides some room to innovate, it has rigid requirements designed for consistency and scalability

In order to participate in the pilot, states must abide by the following conditions...

Stakeholder Engagement	<ul style="list-style-type: none"> ▪ The innovative assessment system must be developed in collaboration with: <ul style="list-style-type: none"> + Teachers, principals and other school leaders + Local education agencies + Parents + Civil rights organizations in the state + Stakeholders representing the interest of children with disabilities, ELs and other vulnerable children
Rigor & Alignment	<ul style="list-style-type: none"> ▪ The innovative assessments must: <ul style="list-style-type: none"> + Align to state academic standards + Address the depth and breadth of state academic standards
Accessibility	<ul style="list-style-type: none"> ▪ The innovative assessment must be accessible to all students, including ELs and special education students <ul style="list-style-type: none"> ▪ E.g. By incorporating the principles of universal design for learning
Validity & Reliability	<ul style="list-style-type: none"> ▪ By the end of the pilot, districts involved must assess a representative sample of students from around the state (e.g. representative number of ELs must participate)
Reporting	<ul style="list-style-type: none"> ▪ Assessment results must <ul style="list-style-type: none"> + Provide aggregate data used for accountability and reporting + Generate an annual summative achievement determination based on the aligned state academic achievement standards + Provide teachers, school leaders and families with timely, disaggregated data to inform and improve instructional practice and student supports

Initially many states expressed interest in the ESSA pilot, but few applied; only GA, LA, NC and NH have been approved to date

“The pilot was initially one of the most buzzed-about pieces of ESSA.... More than a dozen states—including biggies like New York—initially mulled applying, but ended up deciding against it.” –Alyson Klein, EdWeek

“All of our time and resources have been focused on developing the menu of assessments that the Arizona legislature passed into law. The menu didn't really fit into the innovative pilot, so we opted not to participate.” – Stefan Swiat, spokesman for the Arizona Department of Education

	Round 1 Applications	Round 2 Applications
Formally expressed interest ¹	Arizona, Hawaii, Louisiana, New Hampshire	Georgia, Hawaii, South Carolina, Kansas
Applied	Louisiana, New Hampshire, Puerto Rico ²	Georgia, North Carolina
Approved by USED	Louisiana, New Hampshire	Georgia, North Carolina

KEY	States that did not end up applying and/or were not approved to participate in the pilot	States we provide more information on (see slides 7-12)
------------	--	---

¹Besides Arizona (see quote above), it is unclear why the other states did not end up applying—this is something we can follow-up on in our interviews; ²While Alison Klein from Education Week reports Puerto Rico applied for the pilot, USED's website does not include their application, nor state the reason why Puerto Rico was declined. Sources: Klein (2016), Klein (2018b), USDE (n.d.)

States that already have a track record or visionary leadership for innovation are pursuing the pilot opportunity

Louisiana, New Hampshire, Georgia and North Carolina are all hoping to use the innovative assessment pilot to build on work the states are already pursuing

State	Enabling Conditions
Louisiana	<ul style="list-style-type: none"> Outstanding vision and leadership from Supt White (in office since 2011) LDOE has aligned state standards, curriculum and instructional materials, interim (“benchmark”) assessments and formative assessments (EAGLE) such that 75% of students now use the same ELA curriculum LA uses A-F school and district report cards; if LDOE succeeds, other states will want to emulate this work
New Hampshire	<ul style="list-style-type: none"> New Hampshire has been working on PACE since 2011 Participating in the ESSA pilot is an extension of what they were already doing under NCLB
Georgia	<ul style="list-style-type: none"> Georgia is building on its state-led innovative assessment pilot program, established under Senate bill 362. As a result of this bill (and prior to applying for the ESSA pilot), the Georgia State Board of Education held its first statewide competition in 2018, approving three assessment pilots.
North Carolina	<ul style="list-style-type: none"> North Carolina’s innovative assessment pilot builds on work that started in 2014: State Board of Education convened a Task Force on Summative Assessments to create recommendations to reduce the amount of time students spend on standardized testing 2015-16: NCDPI develops a proof of concept on the Task Force’s assessment recommendations

Louisiana is encouraging standards-based instruction and content-rich curriculum by aligning ELA and social studies standards, curriculum and assessments

What makes Louisiana's assessment system innovative?

Louisiana's Educational Assessment Program (LEAP) system...

- Streamlines state testing by combining ELA and social studies assessments
- Measures what students have learned via passages from books that students have read in class (rather than passages they have not read as part of the curriculum)
- Provides districts autonomy to determine which assessments their students will take and which books will be included in the curriculum and assessment
- Assesses students through several brief assessments throughout the year, rather than one longer assessment at the end of the year

**See Appendix B for a list of partners helping LA with the creation, implementation and evaluation of their innovative assessment system*

Implementation Timeline	Successes & Challenges
<ul style="list-style-type: none"> ▪ Louisiana will launch its assessment pilot in SY19-20 with 20 high schools across 3 districts and 2 charter networks ▪ Total of ~21,000 students will participate in the pilot 	<p>⊕ Assessment can become more equitable, focuses on students' classroom instruction and experience, relies on teachers to build test items, and has generated positive feedback from students</p> <p>△ None in LA to date; unclear if this is replicable to other states</p>

“Our hope is that the new assessments will create an incentive for teachers to focus on the meaning of texts, to **focus on building background knowledge** rather than specific skills like summarizing or finding the main idea of a text, which really do not have a strong basis in evidence of assisting students in learning to read,” - John White, State Superintendent

New Hampshire has been developing and refining a competency-based assessment system since 2011

What makes New Hampshire’s assessment system innovative?

New Hampshire’s Performance Assessment of Competency Education (PACE) system...

- Emphasizes competency-based instruction and measurement
- Designed to ensure *all* students have meaningful opportunities to achieve critical knowledge and skills
- Reduces time for standardized assessments by integrating assessments into students’ day-to-day work
- Locally- developed performance tasks, as well as common tasks shared among all participating schools
- Aligns to grade and course competencies determined locally and to statewide academic content standards
- Provides teachers and students with formative and summative information on student progress

Implementation Timeline

- 2011: NH began developing the PACE system
- 2014-18: NH implemented PACE in a subset of schools and districts under an NCLB waiver
- 2019: NH plans to use the ESSA pilot to scale PACE statewide by the end of the 5 year pilot

Successes & Challenges

- ⊕ PACE assesses knowledge in a variety of ways, making it more accessible to students with learning differences
- △ In SY17-18, New Hampshire experienced difficulty identifying an accurate cut-score across participating school districts
- △ While NH teachers speak highly of PACE’s content, a significant percentage of teachers still do not think the time and effort of the performance tasks are worth it

“In a competency-based system, students’ opportunities are judged by the outcomes they achieve and not by “inputs” such as seat time. Therefore, students must achieve identified learning targets before moving on to the next goals and/or graduating from high school. If they do not, school districts are expected to work with families to support additional learning opportunities to **ensure that students have legitimate opportunities to master the necessary knowledge and skills.”**

– Frank Edelbulut, Commissioner of Education

Georgia gave three consortia of districts autonomy to work on formative assessments, in hopes that all districts in the state will soon be able to select one of the three to use

What makes Georgia’s assessment system innovative?

Georgia’s innovative assessment pilot program...

- Georgia ran its own innovative assessment competition in 2018, selecting three consortia of districts to work on new formative assessments
 - + The consortia of districts chosen include: (1) CTLS-Assess, (2) MAP Growth for Georgia and (3) Navy
- The objective of these new, locally developed assessment systems is to provide teachers with a better way to gauge where students are and adjust instruction accordingly
- Each assessment system will:
 - + Provide different features designed to support student learning
 - + Offer interim assessments and other features that provide formative and actionable feedback to teachers and students

Implementation Timeline

- The three consortia of districts are currently working on developing their local assessment systems
- For the five years of the pilot period, each of the three district consortia will work with stakeholders and technical experts to develop and implement innovative assessment systems
- Schools and districts not currently part of a consortium will have an opportunity to apply to join a consortium at anytime; the State Board of Education will approve applications to join
- After five years, Georgia hopes all district will be able to select one of the three assessment systems

**Note, Georgia is too early in the implementation process to determine successes and challenges*

“The intent with these locally developed assessment systems is that they have the **opportunity to do something a little more innovative that better meets their individual needs** in a way that we couldn't do at the state level.”

North Carolina plans to reduce the amount of time students spend taking tests by using a “through-grade assessment model” (rather than a summative model)

What makes North Carolina’s assessment system innovative?

North Carolina’s Innovative Assessment System will...

- Further develop, refine and scale “**NC Check-Ins**”
- Reduce the time students spend taking standardized assessments by creating a through-grade assessment model
- The through-grade assessment model will consist of three or four assessments administered throughout the school year, designed to provide teachers and parents with immediate feedback for guiding subsequent instruction

*What are **NC Check-Ins**?*

- Interim assessments aligned to the state’s grade-level content standards in ELA/reading (grades 4-8) and math (grades 3-8)
- Provides students, teachers, parents and stakeholders with immediate performance data
- While evidence suggests a strong correlation between NC Check-Ins and end-of-grade assessments (EOG), there is not validity evidence to support using results from NC Check-Ins as a predictor of student performance on the state EOG assessment

Implementation Timeline

- 2015-16: NCDPI pilots “NC Check-Ins” in 5th grade math and 6th grade ELA/reading, reaching a total of 9,000 students (4,500 per grade level)
- 2016-17 & SY17-18: Schools participated in the NC Check-Ins study on a voluntary basis
- 2017-18: Feedback from educators overwhelmingly supported NC Check-Ins, especially in terms of its ability to provide teachers with immediate feedback and detailed classroom reports
- 2018-19: NC Check-Ins was administered to over 50% of North Carolina public school students in math (grades 3-8) and ELA (grades 4-8)
- 2019 and beyond: NCDPI Applied for the Innovative Assessment Pilot; If approved, the state will enhance and expand “NC Check-Ins” by developing a through-grade assessment system in place of a summative assessment

These four states are spending \$1MM to \$5MM per year to develop and implement innovative assessment systems



<p>LOUISIANA</p>  <p>~\$4M per year</p>	<ul style="list-style-type: none"> In phase 1, Louisiana will spend ~\$4M per year on developing and implementing the pilot (SY17-18 & SY18-19) March 2019: Hewlett Foundation awarded NWEA a \$1 million, two-year grant to support the implementation of Louisiana's innovative assessment system LDOE receives over \$16 million from federal sources and ~\$20.7 million from state sources per year to support all assessment programs
<p>NEW HAMPSHIRE</p>  <p>~\$1M per year</p>	<ul style="list-style-type: none"> Projected base funding for PACE in the 2018-19 SY is \$627,7000; funding sources include: <ul style="list-style-type: none"> + \$277,700 from NH DOE + \$350,000 from a Hewlett Foundation grant to New Hampshire Learning Initiative (NHLI) to support PACE (NHLI supports funding of PACE) NHDOE and NHLI are seeking an additional \$364,00 from the philanthropic community for supplemental support. If supplemental funds are attained, the total budget for PACE in the 2019-19 SY will be \$991,700.
<p>GEORGIA</p>  <p>~\$2M per year</p>	<ul style="list-style-type: none"> The General Assembly is supporting the state's innovative assessment pilot in 2019 <ul style="list-style-type: none"> + Georgia DOE will continue to request funds from the Georgia Assembly annually The estimated FY 2020 cost to oversee the innovative assessment pilot is \$1,031,888, which covers technical assistance and personnel costs. Additionally, each consortia of districts will accrue costs to develop and implement their assessment systems: <ul style="list-style-type: none"> + CTLS-ASSESS 2019-20 budget is \$3,500,000 (funding source is not identified) + <i>Note, Georgia's innovative assessment application stated the annual budget for the Georgia MAP Assessment partnership and the Putman Consortium is still TBD</i>
<p>NORTH CAROLINA</p>  <p>?</p>	<ul style="list-style-type: none"> Currently, the NCDPI receives \$12 million in state funding and \$9 million in federal funding annually <ul style="list-style-type: none"> + NCDPI believes this budget will meet most of the cost demands for the transition to a through-grade assessment system <i>Note, North Carolina's innovative assessment application does not provide a projected annual budget for the pilot</i>

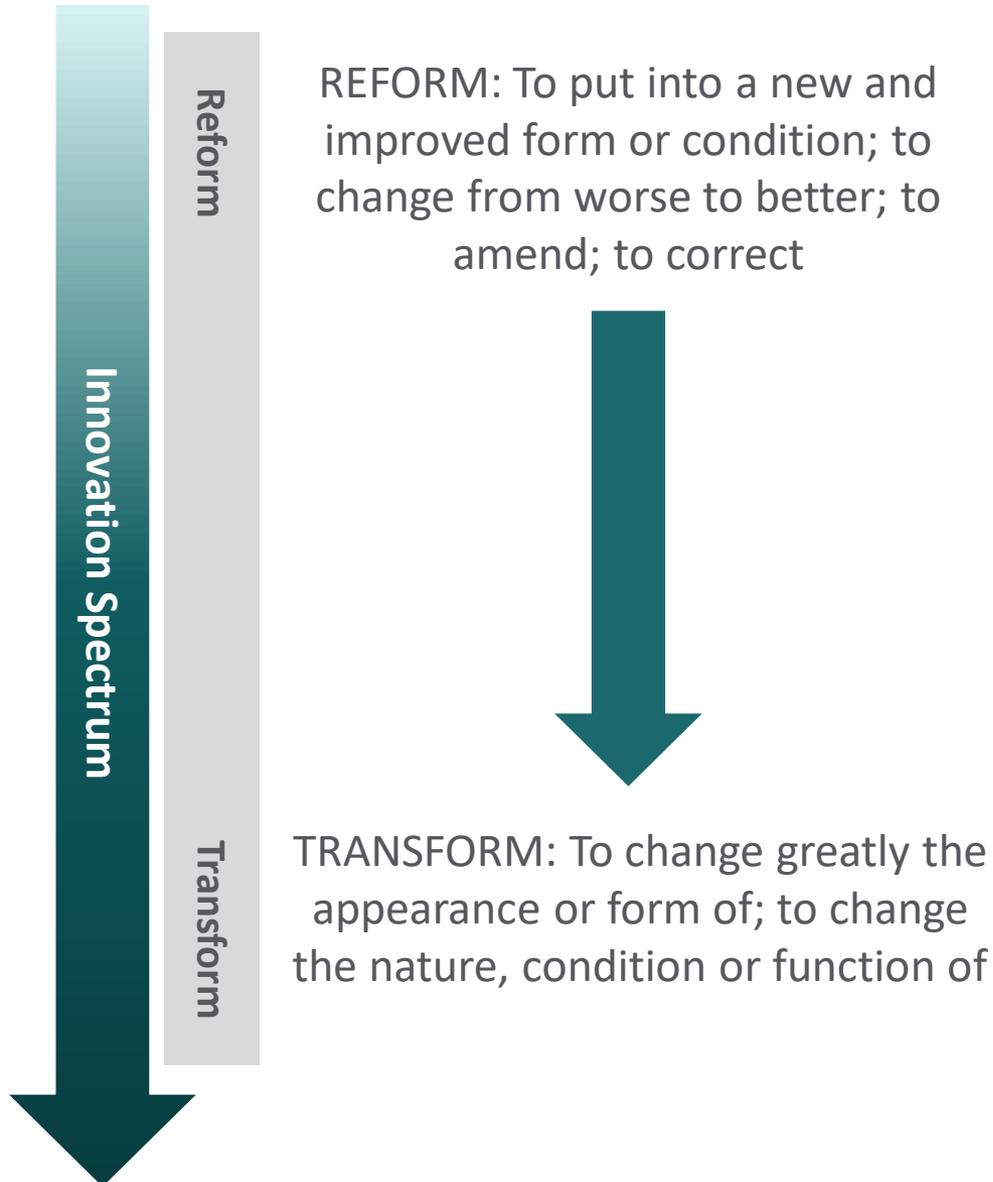


Catalog of Possible Innovation

3 | Ideas for States, Districts and CMOs to Consider

What are the options we've
uncovered for the next
generation of assessment
innovation?

Education First generated more than 25 innovation ideas and placed each on a spectrum from “reform” to “transform”



In Saudi Arabia, they have an app called “Yes, Sir” and as you know their society has an inferior perspective on women. The app tracks women because they need permission to go to store, etc. from “their men.” They can get permission by using the app with the men who are also connected to it. If for example they said they were going to the market but now need to go to another store, the app gives them permission virtually, instead of having to go home and ask. The app is a way to reform the system – because it serves the existing outcome better—the subjugation of women. But it in no way transforms the outcome or system.

This goes back to the idea of defining the purpose of the system. It’s critical to determine that up front. – Todd Rose

We evaluated each innovation's potential with these criteria

High



Low

<p>Potential for Impact: Instruction <i>Supports instructional improvement</i></p>	<p>Assessment measures provide timely, understandable, instructionally-relevant information during the current school year that educators can use to adjust/inform instruction; families can interpret score reports and can access support</p>	<p>Teachers do not see results until the next school year so assessments do not inform current instruction; families don't know how to interpret the results; assessment exists solely for inclusion in accountability; measures limited to core subjects</p>
<p>Potential for Impact: Accountability <i>Addresses barriers/constraints in current system</i></p>	<p>Indicators extend beyond ELA/math measures; measures produce a snapshot against standards (not comprehensive); measures provide <u>valid, comparable, intra-state</u> information on student performance across student groups, schools and districts; measures may be less <u>reliable</u> than in past summative systems</p>	<p>Assessments for accountability take too much testing time; purpose of the measure is to support locally-driven goals (e.g. measure a district's specific graduate profile); intra-state comparability will be difficult to achieve; measures limited to core subjects</p>
<p>Scalable <i>Can be scaled statewide or to multiple school systems in next 5-7 years; innovation will gain influence/traction</i></p>	<p>Measures can be scaled to multiple school systems or statewide within 5-7 years; innovation is "sticky" and advocates can build winning political strategy to fund and scale; won't require massive funding increases over current spending; won't require overhaul of testing vendor capacity or school infrastructure in next 5-7 years</p>	<p>Too "niche" to appeal to broad cross-section of users; too expensive to implement at scale; technology or infrastructure to implement innovation does not yet exist or is in very early stages of R&D; validity of tools used in innovation has not been tested in any way</p>
<p>Demand <i>Users are enthusiastic and willing to take risks; they find the innovation appealing and will want to try it or adopt it</i></p>	<p>Multiple users (policymakers, educators, assessment experts) on board and ready to begin; the innovation addresses problems raised by educators and families; strong, demonstrated support (and maybe existing pilot efforts already being tried) exists in a collection of school districts (who would be the point of the spear in testing out the innovation to go statewide)</p>	<p>Little is known about the innovation outside small "niche" of promoters or vendors; some users actively opposed to the innovation; innovation has too many risks to mitigate, even in a philanthropically-funded pilot project</p>

K8 Assessment Design & Reporting

Innovation Spectrum

Innovation	Definition
Through-Grade 	<ul style="list-style-type: none"> Support more states to use through-grade/through-course assessment Adapt benchmark/interim assessments for use as through-grade assessments Conduct research to identify options to compute summative score from interims/benchmark assessments or through-course tests
Meaningful Reporting 	<ul style="list-style-type: none"> Make results more instructionally useful by connecting individual score reporting to concrete, customized “what to do now” resources for teachers (who may need professional development) and parents (to help or direct support) Turn around results in formats useful to teachers and principals (e.g., last year’s scores grouped and shared with this year’s teacher)
Better Tech, More Growth 	<ul style="list-style-type: none"> Gamify assessments (and use techniques such as simulations or scenario planning) in order to measure critical thinking skills, appeal to broader range of students, include simulations, etc. Enable computer adaptivity Design new ways to measure student growth against grade-level standards
Curriculum-Embedded 	<ul style="list-style-type: none"> Pair up highly-rated instructional materials providers (e.g.) with assessment developers to adapt the materials’ embedded end-of-unit tests into measures that can be used to compute summative scores/results in accountability system
Competency- Based 	<ul style="list-style-type: none"> “Modularize” assessments so that classroom educators can decide when to give test to students when they are ready, during the school year, near point of instruction
Cumulative Validity 	<ul style="list-style-type: none"> Capture and report on dozens of formative and interim data points gathered throughout the year to calculate a summative end of year score/determination (e.g. quizzes, chapter tests, grades, labs, performance tasks, end-of-unit tests, etc.)

Middle/High School Assessment Design & Reporting

Innovation Spectrum

Innovation	Definition
Improve ACT & SAT 	<ul style="list-style-type: none"> Work with ACT and SAT to improve their math assessment portions to more accurately measure college readiness
Scale Existing Innovative Models 	<ul style="list-style-type: none"> Develop and scale more instruction + assessment systems like IB, Springboard Make middle grades and HS assessment in ELA, math and science more PISA-like, measuring learner dispositions like collaboration, complex problem-solving, global competence
Secondary Writing & Research 	<ul style="list-style-type: none"> Create a cross-state portfolio project for high school juniors that measures college-ready writing and research; include in college admissions package
High School Graduate Profiles 	<ul style="list-style-type: none"> Create state and local “graduate profiles” (to specify the cognitive, personal, and interpersonal competencies that students should have when they graduate) Use multiple measures (including community-based/outside of school credentials) to assess the competencies Create “readiness profiles” or “seals”
College Admissions Measures 	<ul style="list-style-type: none"> Work with K12 and higher education leaders to envision broader set of high school assessment measures being used as part of college admissions
Career Ready Measures 	<ul style="list-style-type: none"> Invest in career readiness measures that truly measure what employers in high-value careers need: new 11th/12th measures, transferable workplace performance tasks, CTE pathways, industry-recognized certifications

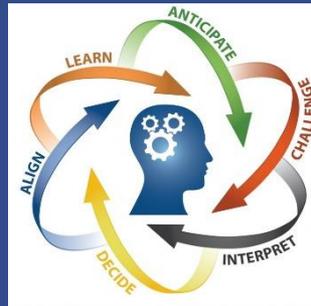
K8 and High School Accountability Policy

Innovation Spectrum

Innovation	Definition
<p>Expand school ratings and definition of school quality</p> 	<ul style="list-style-type: none"> Enhance school ratings to include expanded definition of student success (readiness for college and careers outside academic measures) Include quality inputs (such as instructional materials) as well as outputs (such as academic and non-academic/21st century skills, measures of learning environments)
<p>Redesign school accountability</p> 	<ul style="list-style-type: none"> Broaden high school accountability from ELA, math, science and the 5th indicator to expanded definitions of success in life, including on-track/off-track measures signifying persistence to HS graduation, and outcome indicators from high school, career and postsecondary institutions signifying the quality of the HS diploma (currently, at least 7 states have P20/W longitudinal data systems that could enable this)
<p>Unify K12, postsecondary and workforce accountability systems to common goal</p> 	<ul style="list-style-type: none"> Engage governors that have set ambitious statewide goals for adult postsecondary completion and employment (such as Hawai'i's "55% of adults have a credential/degree by 2025, or 55 by 25") to bring together K12, postsecondary and workforce development systems and design high school accountability around most-critical measures

We provide “A Closer Look” at four of the 25 innovation ideas; each helps illustrate the complexity and opportunities of summative measures

1. THROUGH-GRADE ASSESSMENTS



2. GRADUATE PROFILE

SHELBY COUNTY PUBLIC SCHOOLS

This Profile describes the expectations our community believes are required of successful leaders who graduate from SCPS. A Shelby County Public Schools graduate is:



3. CAREER READY MEASURES



4. REDESIGN ACCOUNTABILITY

Accountability achieve.org

HIGH SCHOOL ASSESSMENTS AND ACCOUNTABILITY DURING ESSA TRANSITION Achieve

As states consider how their accountability systems will change under the Every Student Succeeds Act (ESSA), they face important decisions about which high school assessments to incorporate, what content their assessments will measure, which tests all students will take, and how rigorous those assessments will be. ESSA requires that each state administer a statewide assessment to all students to assess proficiency in English language arts (ELA)/literacy and mathematics at least once in high school and include these results in school and district accountability systems as a measure of student achievement. States must also assess students in high school in science but are not required to include a measure of science achievement in their high school accountability systems.

The purpose of this brief is to examine which ELA/literacy, mathematics, and science assessments states are using to measure high school students' achievement within each state's accountability system and what specific content or grades have been assessed.¹ Achieve hopes that understanding the current landscape — and the vast variations — might push

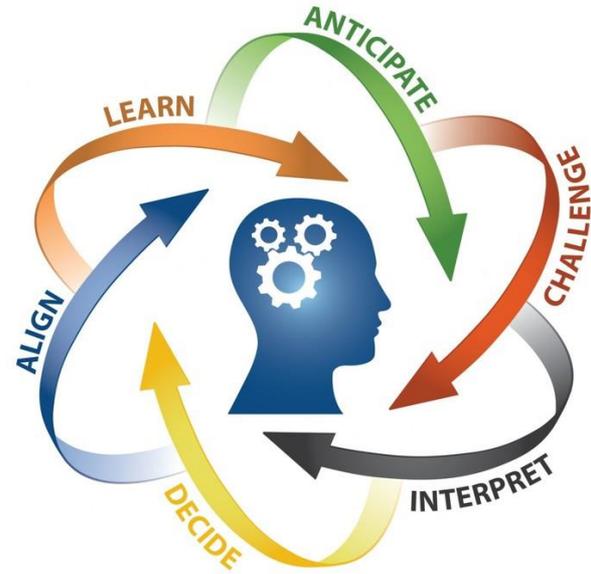
A Closer Look: What are Through-Grade Assessments?

What are through-grade assessments?

- Assessments that are administered intermittently throughout the year and roll-up to create a student summative score
- Through-course assessments cover all standards by the end of the year, with the capability to adapt content up and down based on student skill
- Through-grade assessments provide feedback to teachers and families about student progress in real time
- Through-grade assessments require all participants to use the same scope and sequence for standards (but not curriculum)

States developing or considering through-course assessments:

- Alaska, Georgia, Nebraska and North Carolina
- ESSA permits any state to use through-grade assessments



Standards, assessments, and accountability were supposed to be a feedback mechanism to evaluate instruction and that is not happening...the system has forced a focus toward the end of the year. – Assessment Expert

A Closer Look: How well do Through-Grade Assessments meet our criteria for evaluating innovations?

Summary		
		
<p>High potential</p> <p>Simple to understand but transformative innovation; moderate cost; likely scalable in 5 years; enhances equity goals</p>		
Ratings		
Impact on Instruction		<ul style="list-style-type: none"> Moves assessment closer to time of instruction and provides educators with real-time feedback on student learning Raises stakes on interim assessments, which could detract from their instructional value
Impact on Accountability		<ul style="list-style-type: none"> Enables single summative score, annual disaggregated data and comparisons across student groups, schools and districts Removes need for end-of-year standalone standardized tests
Scalable		<ul style="list-style-type: none"> Requires majority of state and/or regional networks of school systems within a state to use same curricular scope and sequence (not necessarily same curriculum, but same timing: “ratios and proportions in the first quarter”)
Demand		<ul style="list-style-type: none"> Educators and families likely to embrace this approach as long as it doesn’t substantially increase testing time during the school year or delay return of results
Equity Analysis	<p>Enhances equity goals: Brings curriculum and assessment closer in alignment; provides comparable information across student groups; provides real-time feedback to teachers on student performance</p>	

A Closer Look: What are Graduate Profiles?

What are graduate profiles?

- A document co-created with stakeholders to specify the cognitive, personal and interpersonal competencies students should have when they graduate
- Clear visualization of priority goals for teaching and learning that can be easily communicated to students, parents, faculty, and staff to align their collective efforts

A community-wide vision statement describing what a learner should know and be able to do before he or she graduates from the school. – Competency-based education expert



A Closer Look: How could school systems/states measure and use Graduate Profiles in accountability systems?



Table 9.1 Example of a readiness profile using currently available data

Key cognitive strategies	Wt.	Key content knowledge	Wt.
Independent investigation, culminating project	5	Quality of courses taken/ GPA in high quality courses	7.5
Test score and grade for AP course(s) that addressed key cognitive strategies	5	CTE pathway courses or CTE dual enrollment course(s) grades or other dual-enrollment courses or career exploration or preparation program	7.5
IB Extended Essay or comparable research paper	5		
Applied problem solving project in CTE course	5	Adjustment based on rigor of school grading practices	5
Scores on other cognitively challenging tests (e.g., IB), or other activities to develop cognitive strategies (e.g., science fair, Odyssey of the Mind)	5	Scores and subscores on state test of reading, math, and science (percentile rank & performance level)	12.5
		Participation in challenging curricular programs (e.g., Project Lead the Way)	7.5
	25.0		40.0

Key learning skills	Wt.	Key transition skills	Wt.
Attendance 95% or above, or overcoming challenges to attending regularly	5	Transition plan developed in 9th grade, updated annually	5
All elements of complex classroom-based performance task or project completed on time	5	Academic interests, career and personal goals identified	5
		Choices and decisions consistent with desired goal	2.5
Teacher rating of student learning skills	2.5	Evidence of multiple colleges and careers researched	2.5
Student rating of learning skills	2.5	Passing score on test of financial aid knowledge and strategies	2.5
Positive behavior plus contributions to classroom and school	2.5		
	17.5		17.5

How would we measure student accomplishment of the Profiles?

- Create statewide (CO) or school system (Shelby Co, Valor) Graduate Profiles
- Identify multiple tools, such as assessments menu including standardized and performance measures; SAT/ACT, IB/AP, capstone project; career certifications; micro-credentials and badges from in-school and out-of-school and work-based learning activities

Could these profiles be used for accountability?

- Possibly could include in student transcripts (“readiness seal”?)
- Difficult to report comparably on school quality; could report % of graduates who earn “readiness seal”



A Closer Look: How well do Graduate Profiles meet our criteria for evaluating innovations?

Summary		
	<p>Moderate potential</p> <p>Transformative innovation; likely 10+ years to develop and use; difficult to scale; costly; both enhances and could detract from equity goals</p>	
Ratings		
Impact on Instruction		<ul style="list-style-type: none"> Deeply meaningful to local communities (educators, families, employers) Can utilize variety of locally-available, credible measures, including in-school and out-of-school such as Girl Scouts badges
Impact on Accountability		<ul style="list-style-type: none"> Difficult to envision including Graduate Profiles in school accountability Difficult to make comparisons across communities
Scalable		<ul style="list-style-type: none"> Could be scaled to multiple school systems within a state, with a state leadership role (see e.g. Colorado) for quality/auditing/information Requires deep local expertise, access/partnerships with credentialers
Demand		<ul style="list-style-type: none"> Leading edge/innovator school systems deeply invested in this approach, but there aren't many such systems relative to US
Equity Analysis	Enhances equity goals: Provides more access community-wide to necessary competencies and multiple measures	Detracts from equity goals: Difficult to make comparisons across communities

A Closer Look: Career Ready Measures



We've been working over the past few years to help think about career readiness as not just core academic subjects...[but also] more transferable skills, [such as] collaboration, communication, time and resource management...In career exploration, transferable skills are important. – Educator

What's the opportunity?

- States are seeking new and different measures of career readiness, distinct from college-ready measures
- Need measures of 21st century knowledge and at the careers level (not specific jobs or general skills)
- Different career pathways (or clusters) have differing, unique skill requirements for reading and mathematics
- Getting more students to pursue challenging careers requires measures that let them see what they could do, not just what they can't do
- Plethora of existing industry-recognized certifications aren't included in many state accountability systems

A Closer Look: How well do Career Ready Measures meet our criteria for evaluating innovations?

Summary






Moderate potential

Transformative; Moderate cost; needs more development; could scale in 5-10 years; enhances and could detract from equity goals

Ratings

Impact on Instruction		<ul style="list-style-type: none"> ▪ If done well, signals upgrades needed to CTE curriculum in HS ▪ Could discourage college-ready coursetaking
Impact on Accountability		<ul style="list-style-type: none"> ▪ If done well, makes available more measures that have external value to families, policymakers and employers
Scalable		<ul style="list-style-type: none"> ▪ Requires moderate additional investments to create more measures of value ▪ Requires communication strategy and training with schools to discourage tracking
Demand		<ul style="list-style-type: none"> ▪ High demand from governors, chiefs, state policymakers and employers

Equity Analysis 	<p>Enhances equity goals: Increases rigor of career options compared to current state, and low-income students and students of color over-represented in current low-quality pathways</p>	<p>Detracts from equity goals: Could encourage more tracking if not done well; could encourage students to earn credentials without employment value</p>
--	---	--

A Closer Look: Redesign High School Accountability

What's the opportunity?

Broaden high school accountability from ELA, math, science and the 5th indicator to expanded definition of success in life that considers high schools' contributions to key outcome indicators:

- 9th grade on-track and off-track indicators, HS graduation rates
- Student engagement indicators
- End-of-course exams, SAT/ACT, 11th and 12th grade portfolios/performance tasks
- Meaningful college-ready coursetaking (labs, mathways, world languages)
- College credit in HS (AP/IB/dual enrollment)
- Career readiness participation (career-concentrator, work-based learning) and employer-valued credentials earned
- College enrollment, remediation, persistence and completion data
- Wages at age 25

ESSA evidence to include:

- Student engagement
- On track indicators
- Off track indicators

Accountability achieve.org

HIGH SCHOOL ASSESSMENTS AND ACCOUNTABILITY DURING ESSA TRANSITION 

As states consider how their accountability systems will change under the Every Student Succeeds Act (ESSA), they face important decisions about which high school assessments to incorporate, what content their assessments will measure, which tests all students will take, and how rigorous those assessments will be. ESSA requires that each state administer a statewide assessment to all students to assess proficiency in English language arts (ELA)/literacy and mathematics at least once in high school and include these results in school and district accountability systems as a measure of student achievement. States must also assess students in high school in science but are not required to include a measure of science achievement in their high school accountability systems.

The purpose of this brief is to examine which ELA/literacy, mathematics, and science assessments states are using to measure high school students' achievement within each state's accountability system and what specific content or grades have been assessed.¹ Achieve hopes that understanding the current landscape — and the vast variations — might push

I'm increasingly of the opinion that these essentially-NCLB systems are not salvageable. We can't hold the line. The political support for these systems has eroded, from both the left and the right. But if we could build a system that includes life outcomes measurement, I also believe you could rebuild the political support. Parents really do care about citizenship and careers. I'm saying this as someone whose entire history is as an accountability hawk. There is a sea change going on, including in my own brain.

— Advocacy Leader

A Closer Look: How well does Redesign High School Accountability meet our criteria for evaluating innovations?

Summary



High potential
Transformative innovation; could scale/get traction in 5-7 years; moderately expensive; enhances equity goals

Ratings

Impact on Instruction		<ul style="list-style-type: none"> Includes measures with strong evidence/correlation with life success Encourages focus on all grades 9-12
Impact on Accountability		<ul style="list-style-type: none"> High schools have med/high influence over most of the measures (e.g. 9th grade on-track/off-track, graduation rates, college credits, meaningful college-ready coursetaking, remediation) Represents major overhaul and shift to multiple measures with broad external value instead of state summative tests that have little value
Scalable		<ul style="list-style-type: none"> Requires significant investments to address data privacy, interoperability and quality Requires political strategy re: data privacy and cross-sector leadership
Demand		<ul style="list-style-type: none"> Would garner support from educators and families Requires significant visionary education and political leadership to accomplish major overhaul
Equity Analysis	Enhances equity goals: Comparable across schools and communities (and nationwide); Includes measures with strong evidence/correlation with life success	



5 | Appendix:
Interviewees
Sources

Interviews

Interviewee Category

Interviewee Name & Organization

SEAs

- Katy Anthes, Colorado Department of Education
- Sheila Briggs, Wisconsin Department of Public Instruction
- Jill Cowart, Louisiana Department of Education
- Jeff Cottrill, Texas Department of Education
- Cory Epler, Nebraska Department of Education
- Matt Jones, Georgia Department of Education
- Hanseul Kang, District of Columbia Office of State Superintendent of Education
- Paul Leather, Center for Innovation, which operates assessment network with Colorado, Hawaii, Illinois, Massachusetts, New Hampshire, New Mexico, Virginia
- Rick Miller, California CORE
- Gwen Perea Warniment, Tim Hand and Lynn Vasquez, New Mexico Public Education Department
- Jhone Ebert, Jonathan Moore and Jessica Todtman, Nevada Department of Education
- Jeff Riley, Julie Albino, Jass Stewart, Massachusetts Department of Elementary and Secondary Education
- Rebecca Shah, Tennessee Department of Education

Interviews

Interviewee Category

Interviewee Name & Organization

State and National Leaders and Funders

- Sara Allan, Gates Foundation
- Brad Bernatek, Gates Foundation
- Mike McGee and Ramin Taheri, Chiefs for Change
- Jen Alexander, PIE Network
- Bibb Hubbard, Learning Heroes
- Scott Norton and Kirsten Carr, CCSSO
- Jon Schnur, Bloomberg Philanthropies
- Michael Watson, New Classrooms
- Ross Wiener, Aspen Education & Society Program

Assessment Innovators and Leading Edge Thinkers

- Larry Berger, Amplify
- Jack Buckley, Imbellus
- Magda Chia, SCALE
- David Conley, University of Oregon
- Brian Gong, Center on Assessment
- Michael Horn, Entangled Solutions
- Chris Minnich, NWEA
- Heidi Ramirez, America Achieves
- Todd Rose, Harvard University
- Laura Slover and Jeff Nellhaus, CenterPoint
- Justin Wells, Envision Learning Partners

Sources

- Achieve, *2017-18 State High School Mathematics Assessments* (2019a), accessed May 1st, 2019, https://www.achieve.org/files/HS_Math_Assessments_Table_2017-18-FINAL.pdf
- Achieve, *2017-18 State High School English Language Arts/Literacy Assessments* (2019b), accessed May 1st, 2019, https://www.achieve.org/files/HS_ELA_Assessments_Table_2017-18-FINAL.pdf
- Achieve, *How do assessments matter—and for whom? Making sense of the high school student’s assessment experience* (2019c), accessed June 1st, 2019, <https://www.achieve.org/files/Student%20Assessment%20Experience%20brief.pdf>
- <http://blogs.edweek.org/edweek/campaign-k-12/2018/10/georgia-essa-innovative-test.html>
- Alyson Klein, *States Can Try Out More Than One Test Under ESSA Pilot, Trump Ed. Dept. Says*, (Education Week, 2018a), accessed May 1st, 2019, <http://blogs.edweek.org/edweek/campaign-k-12/2018/12/essa-testing-innovative-pilot-states-georgia.html>
- Alyson Klein, *Four States Want In on Second Round of ESSA’s Innovative Assessment Pilot* (Education Week, 2018b), accessed May 1st, 2019, <http://blogs.edweek.org/edweek/campaign-k-12/2018/10/essa-innovation-testing-georgia-kansas-south-carolina-hawaii.html>
- Alyson Klein, *Betsy DeVos Ok’s Louisiana Pitch to Use Innovative Tests Under ESSA* (Education Week, 2018c), accessed May 1st, 2019, <https://blogs.edweek.org/edweek/campaign-k-12/2018/07/betsy-devos-louisiana-test-innovative.html?r=1043396720>
- Alyson Klein, *Georgia Wants In on the ESSA Innovative Assessment Pilot* (Education Week, 2018d),
- Alyson Klein, *How Will ESSA’s Innovative Assessment Pilot Work* (2016, Education Week), accessed May 1st, 2019, <https://blogs.edweek.org/edweek/campaign-k-12/2016/06/how-will-essas-innovative-asse.html>
- Andre Rotherham and Bonnie O’keefe, *Why Underground Testing Consortia are Thriving* (The 74, 2017), accessed June 1st, 2019, <https://www.the74million.org/article/okeefe-rotherham-why-underground-testing-consortia-are-thriving/>
- Andrew Ujifusa, *Louisiana and New Hampshire have stepped up for ESSA’s experiment in crafting new student assessments*, (Education Week, 2019), accessed May 1st, 2019, <https://www.edweek.org/ew/articles/2019/04/03/overhauling-student-assessments-a-view-from-the.html>
- Ashley Jochim and Patrick McGuinn, *The Politics of the Common Core Assessments: Why states are quitting the PARCC and Smarter Balanced testing consortia* (Education Next, 2016), accessed June 1st, 2019, <https://www.educationnext.org/the-politics-of-common-core-assessments-parcc-smarter-balanced/>
- Ashley Zanchelli, *This Week’s ESSA News: Louisiana Supe John White on the Innovative Assessment Pilot, Alabama Report Card Delayed Again & More* (The 74 million, 2019), accessed May 1st, 2019, <https://www.the74million.org/article/this-weeks-essa-news-louisiana-supe-john-white-on-the-innovative-assessment-pilot-alabama-report-card-delayed-again-more/>
- Brady McCombs, *Utah Cancels \$44 Million Contract with School Test Company* (KSL TV, 2019), accessed June 7th, 2019, <https://ksltv.com/415660/utah-cancels-44-million-contract-with-school-test-company/>
- Caitlin Emma, *New Hampshire Leads the Way on New Testing Approaches* (Politico, 2018), accessed May 1st, 2019, <https://www.politico.com/newsletters/morning-education/2018/01/04/new-hampshire-leads-the-way-on-new-testing-approaches-064916>
- Catharin Gewerts, *ESSA Offers Testing Flexibility. So Why Aren’t States Using It?* (Education Week, 2018), accessed May 1st, 2019, <https://www.edweek.org/ew/articles/2018/04/04/essa-offers-testing-flexibility-so-why-arent.html>
- Center for Assessment, *PACE 2017-18 Standard Setting Results: Issues and Recommendation* (2019), accessed May 1st, 2019, <https://www.education.nh.gov/assessment-systems/documents/pace-standard-setting-memo.pdf>
- Chiefs for Change, *ESSA Indicators of School quality and Student Success* (2016), accessed June 1st, 2019, <http://chiefsforchange.org/policy-paper/3244/>

Sources

- Channel 3 News, *New students testing vendor to cost Tennessee \$93M over five years* (2019), accessed June 10th, 2019, <https://wreg.com/2019/06/10/new-student-assessment-vendor-to-cost-93m-over-5-years/>
- Communities in Schools and Education First, *Tool 1.7: Accountability and School Improvement Under ESSA* (2016).
- Dan Fletcher, *Standardized Testing*, (Time, 2009), accessed June 1st, 2019, <http://content.time.com/time/nation/article/0,8599,1947019,00.htm>
- David Conley and Mary Seburn, *The Promise and Practice of Next Generation Assessment: Using Multiple Measures to Enhance Student Self-knowledge and Ownership of Learning* (2019)
- Education First, Interviews (2019)
- Education first, *In a State of Flux: The State of State Assessments* (2018).
- Elizabeth Barnett, Peter Bergman, Elizabeth Kopko, Vikash Reddy, Clive Belfield and Susha Roy, *Multiple measures placement using data analytics: An implementation and early impacts report* (Center for the Analysis of Postsecondary Readiness, 2018), accessed June 1st, 2019, <https://www.mdrc.org/publication/multiple-measures-placement-using-data-analytics>
- Georgia Department of Education, *Georgia's Application for the Innovative Assessment Demonstration Authority* (2018), accessed May 1st, 2019, <https://www2.ed.gov/admins/lead/account/iada/gaiadaappdec2018.pdf>
- John Hattie, *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement* (British Journal of Educational Studies , 2011): 59, 2, accessed June 1st, 2019, <https://www.tandfonline.com/doi/full/10.1080/00071005.2011.584660?scroll=top&needAccess=true>
- Jill Mullen and Julie Rowland, *50-State Comparison: State Summative Assessments* (Education Commission of the States, 2019).
- Julie Woods, *50-State Comparison: States' School Accountability Systems* (Education Commission of the States, 2018), accessed June 1st, 2019, <https://www.ecs.org/50-state-comparison-states-school-accountability-systems/>
- Julie Woods, *P-20 Data Systems and Report Cards* (2017), accessed June 1st, 2019, https://www.ecs.org/wp-content/uploads/State-Information-Request_-P-20-Accountability-Measures-and-Report-Cards.pdf
- John Rosales, *The Racist Beginnings of Standardized Testing: From grade school to college, students of color have suffered the effects of Biased testing* (National Education Association, 2018), <http://www.nea.org/home/73288.htm>
- Krista Kaput, *How States Are Innovating With Assessments Under ESSA* (Education Evolving, 2018), accessed May 1st, 2019, <https://www.educationevolving.org/blog/2018/10/how-states-are-innovating-with-assessments-under-essa>
- Lorrie Shepard, *A Brief History of Accountability Testing* (NY, NY: Routledge, 2008), *The Future of test-based accountability* (p. 25-46).
- Louisiana Department of Education, *Louisiana Believes: Louisiana's Application for Innovative Assessment Demonstration Authority under Section 1204 of the Elementary and Secondary Education Act (ESEA)* (2018), accessed May 1st, 2019, <https://www2.ed.gov/admins/lead/account/iada/laiadaapplication.pdf>
- Lynn Olson, *The New Testing Landscape: State Summative Assessments Under the Federal Every Student Succeeds Act* (FuturEd: Georgetown University, 2019).
- Louisiana Department of Education, *Louisiana ESSA Innovative Assessment Pilot First to Receive Federal Approval* (2018), accessed May 1st, 2019, <https://www.louisianabelieves.com/newsroom/news-releases/2018/07/27/louisiana-essa-innovative-assessment-pilot-first-to-receive-federal-approval>
- Market Insider, *William and Flora Hewlett Foundation Awards NWEA \$1 Million to Support Development of Innovative Assessment Pilot in Louisiana* (2019), accessed May 1st, 2019, <https://markets.businessinsider.com/news/stocks/william-and-flora-hewlett-foundation-awards-nwea-1-million-to-support-development-of-innovative-assessment-pilot-in-louisiana-1028046478>

Sources

- National Center for Learning Disabilities, *New Hampshire's PACE Assessments: Transforming Assessments and Learning from the Ground Up* (2017), accessed May 1st, 2019, https://www.nclld.org/wp-content/uploads/2017/07/PACE.Case-Study.Fin_-1.pdf
- New Hampshire Department of Education, *Performance Assessment of Competency Education (PACE)* (n.d.), accessed May 1st, 2019, <https://www.education.nh.gov/assessment-systems/pace.htm>
- New Hampshire Department of Education, *Application for the New Authorities Under the Innovative Assessment Demonstration Authority* (2018), accessed May 1st, 2019, <https://www2.ed.gov/admins/lead/account/iada/nhapplication.pdf>
- North Carolina Department of Public Instruction, *North Carolina Application for New Authorities under the Innovative Assessment Demonstration Authority* (2018), accessed May 1st, 2019, <https://www2.ed.gov/admins/lead/account/iada/nciadaappdec2018.pdf>
- Richard Riley, *The Improving America's Schools Act of 1994: Reauthorization of the Elementary and Secondary Education Act* (U.S. Department of Education, 1995), accessed June 1st, 2019, <https://www2.ed.gov/offices/OESE/archives/legislation/ESEA/brochure/iasa-bro.html>
- Tom Vander Ark, *A Proposal to End Standardized Testing* (Getting Smart, 2019), accessed June 1st, 2019, <https://www.gettingsmart.com/2019/04/a-proposal-to-end-standardized-testing/>
- Thomas Edsall, *The Meritocracy Is Under Siege: Are we merely reproducing privilege or is there something to salvage in the system we have now?* (The New York Times, 2019), accessed June 12th, 2019, <https://www.nytimes.com/2019/06/12/opinion/meritocracy-tests-education.html?smid=nytcore-ios-share>
- U.S. Department of Education, *Innovative Assessment Demonstration Authority (IADA)* (n.d.), accessed May 1st 2019, <https://www2.ed.gov/admins/lead/account/iada/index.html>
- U.S. Department of Education, *Every Student Succeeds Act: Assessments under Title I, Part A & Title I, Part B: Summary of Final Regulations* (n.d.), accessed June 1st. 2019, <https://www2.ed.gov/policy/elsec/leg/essa/essaassessmentfactsheet1207.pdf>
- United States Congress, *1177 — Every Student Succeeds Act* (2015), accessed May 1st, 2019, <https://www.govinfo.gov/content/pkg/BILLS-114s1177enr/pdf/BILLS-114s1177enr.pdf>
- William Bryant, *The Future of Testing* (Getting Smart, 2018), accessed June 1st, 2019, <https://www.gettingsmart.com/2018/02/the-future-of-testing/>