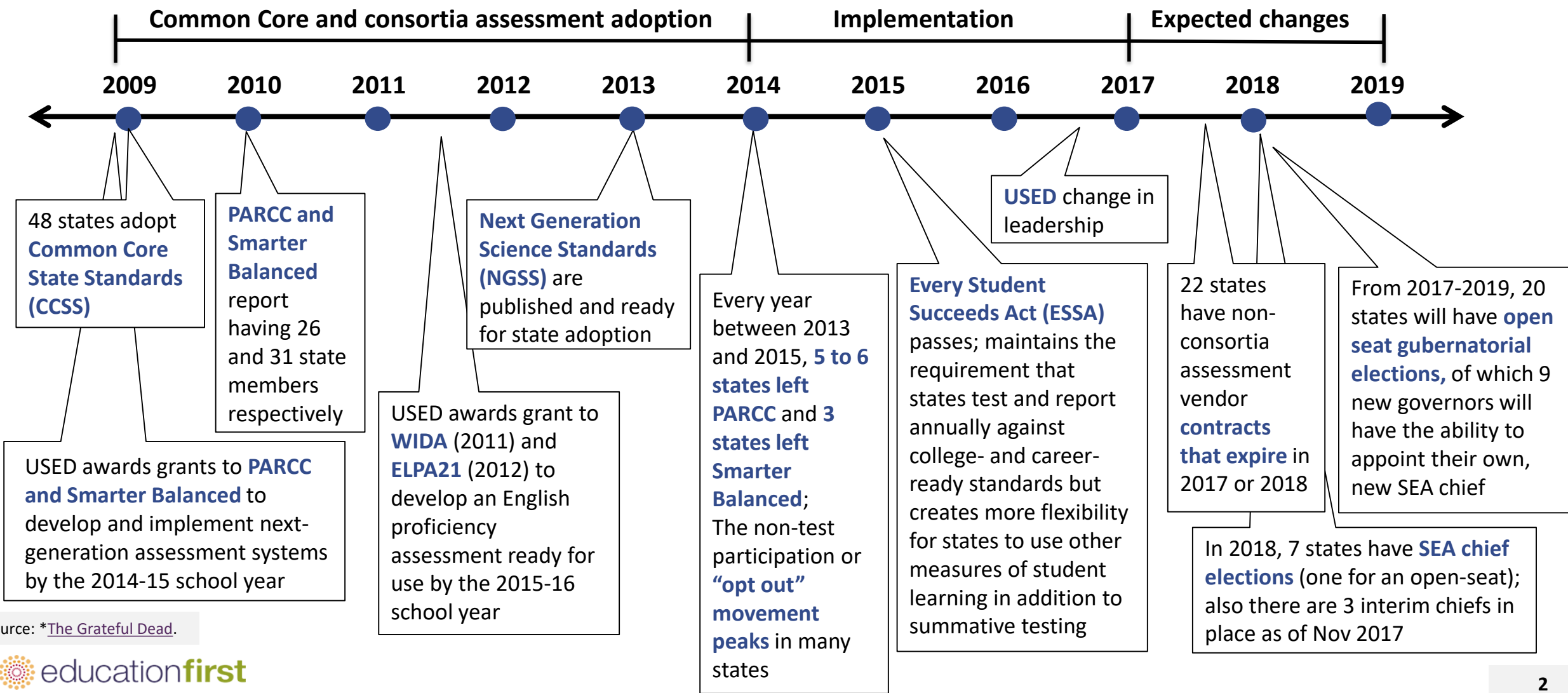


Whatever Happened to All Those New & Better State Tests?

THE STATE OF STATE ASSESSMENTS

“What a long, strange trip it’s been....”* Since 2009, a shifting policy and political environment has dramatically influenced state testing decisions—and the environment continues to evolve under new policy and education leaders



Recognizing these changes, the Hewlett Foundation asked Ed First to generate a fresh analysis of the “state of state assessments”... The big finding: 40% of states are still using consortia test items, but test quality appears not to be a prominent factor in many states

Consortia assessments continue to be rated highly...

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.

But consortia membership continues to decline...

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).

Most states are ‘going at it alone’ again, especially with grade 3-8 tests...

The majority of states are working with AIR (10), DRC (6) or other vendors (16).

SAT and ACT are on the rise in high schools, including in consortia states...

More states are moving to using SAT or ACT as their HS accountability test—now 13 states total, of which 6 are Smarter Balanced or PARCC consortia members—despite concerns about whether how well these tests measure state learning standards.

Experts are worried about these trends...

As more states venture out on their own or choose off-the-shelf tests, experts we interviewed for this project question quality, standards alignment, comparability and vendor capacity.

Key: Throughout these slides, **PARCC participating states are indicated in blue typeface** and **Smarter Balanced participating states are indicated in orange typeface**

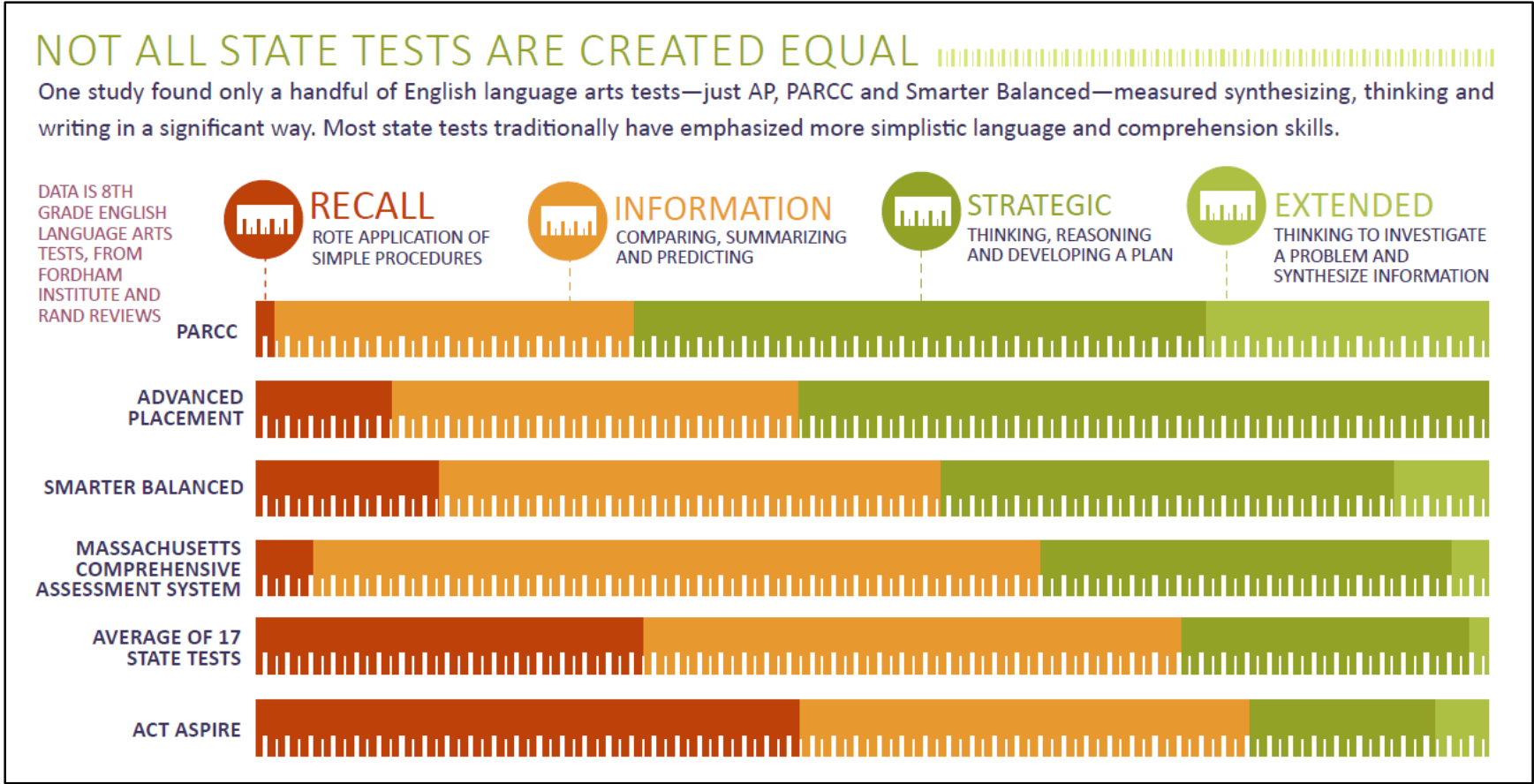
Note: See research methodology and list of expert interviewees on slide 17

What is a high-quality test? In 2014, CCSSO articulated criteria for state leaders to use in making sure their assessments matched the depth, breadth and rigor of newer state academic standards (which now emphasize problem-solving, critical thinking and writing)

CCSSO encourages state officials to use these criteria as “they develop procurements and evaluate options for high-quality state summative assessments aligned to college- and career-readiness standards.” Researchers, evaluators and advocates also have used these criteria to independently review different summative assessment options in the marketplace.

| A | B | C | D | E | F |
|--|--|--|--|--|--|
| Meet Overall Assessment Goals and Ensure Technical Quality | Align to Standards – English Language Arts/Literacy | Align to Standards - Mathematics | Yield Valuable Reports on Student Progress and Performance | Adhere to Best Practices in Test Administration | State Specific Criteria (as desired) |
| <ul style="list-style-type: none"> A.1 Indicating progress toward college and career readiness A.2 Ensuring that assessments are valid for required and intended purposes A.3 Ensuring that assessments are reliable A.4 Ensuring that assessments are designed and implemented to yield valid and consistent test score interpretations within and across years A.5 Providing accessibility to all students, including English learners and students with disabilities A.6 Ensuring transparency of test design and expectations A.7 Meeting all requirements for data privacy and ownership | <ul style="list-style-type: none"> B.1 Assessing student reading and writing achievement in both ELA and literacy B.2 Focusing on complexity of texts B.3 Requiring students to read closely and use evidence from texts B.4 Requiring a range of cognitive demand B.5 Assessing writing B.6 Emphasizing vocabulary and language skills B.7 Assessing research and inquiry B.8 Assessing speaking and listening B.9 Ensuring high-quality items and a variety of item types | <ul style="list-style-type: none"> C.1 Focusing strongly on the content most needed for success in later mathematics C.2 Assessing a balance of concepts, procedures, and applications C.3 Connecting practice to content C.4 Requiring a range of cognitive demand C.5 Ensuring high-quality items and a variety of item types | <ul style="list-style-type: none"> D.1 Focusing on student achievement and progress to readiness D.2 Providing timely data that inform instruction | <ul style="list-style-type: none"> E.1 Maintaining necessary standardization and ensuring test security | <p>Sample criteria might include</p> <ul style="list-style-type: none"> Requiring involvement of the state’s K-12 educators and institutions of higher education Procuring a system of aligned assessments, including diagnostic and interim assessments Ensuring interoperability of computer-administered items |

Independent reviews of state tests in recent years—applying the CCSSO criteria and others—have turned up wide variations in quality and depth



Mathematician Norman Webb’s “Depth of Knowledge” scale categorizes learning tasks—such as test questions or classroom assignments—into four levels, according to the complexity of thinking required by students to successfully complete them.

USED state assessment peer reviews (required by federal law) confirm these variations: As the best current proxy for assessment quality, these reviews have found less than 1/2 of non-consortia states meet criteria... The good news? With peer review approval of Maryland, PARCC now meets the criteria for alignment with standards for all states

| In examining state assessments systems, reviewers consider six criteria | |
|---|---|
| 1. | Statewide system of standards and assessments |
| 2. | Assessment system operations |
| 3. | Technical quality– validity |
| 4. | Technical quality– other |
| 5. | Inclusion of all students |
| 6. | Academic achievement and standards reporting |

States present evidence about their assessment systems against each of the six criteria, and states not fully meeting peer review requirements did so for different reasons.

- Maryland’s state test—based on PARCC—has met all peer review criteria. Importantly, this unconditional approval means any other state using PARCC meet the criteria for alignment with state standards (there could remain other parts of a state’s assessment system where reviewers need more details).
- 92% of consortia states (12 of 13 states, not including MD) that submitted a peer review in 2016 or 2017 substantially met assessment requirements (reviewers asked for more information on some aspects)
- 44% of non-consortia states (7 of 16 states) that recently submitted a peer review substantially met assessment requirements

| DOES NOT MEET REQUIREMENTS | PARTIALLY MEETS REQUIREMENTS | SUBSTANTIALLY MEETS REQUIREMENTS | MEETS REQUIREMENTS |
|----------------------------|---|--|--------------------|
| | Alabama (grades 3-8), Arizona, Florida (grades 3-8), Illinois , Iowa , Kansas, Kentucky , Minnesota, North Carolina, Texas, Virginia , Wisconsin (HS), Wyoming | Colorado , California , Connecticut , Delaware , DC , Hawaii , Idaho , Florida (HS), Indiana (grades 3-8), Missouri, New Jersey , New Mexico , New York, North Dakota, Oregon , Pennsylvania (grades 3-8), South Dakota , Rhode Island, Utah, West Virginia , Vermont | Maryland |

HOWEVER... While independent, public reviews examining the new generation of state tests consistently praise the quality of PARCC and Smarter Balanced assessments...

When I had the chance to sit down and compare content side-by-side, it became evident the substance of the PARCC and Smarter Balanced tests outshined the material from the old tests in several ways.

—Josh Parker, Maryland State Teacher of the Year, and reviewer in NNSTOY study

PARCC and Smarter Balanced tests in particular are tests that emphasize the most important content and require students to demonstrate the depth of work called for by college and career ready standards. Both tests measure a wide range of real-world skills like critical thinking, problem solving, and analysis.

—Sheila R. Schultz, HumRRO

According to our team of 40 reviewers, the PARCC and Smarter Balanced grade 3 and 5 assessments earned an Excellent or Good match to the subject-area CCSSO criteria for both ELA/Literacy and Mathematics.

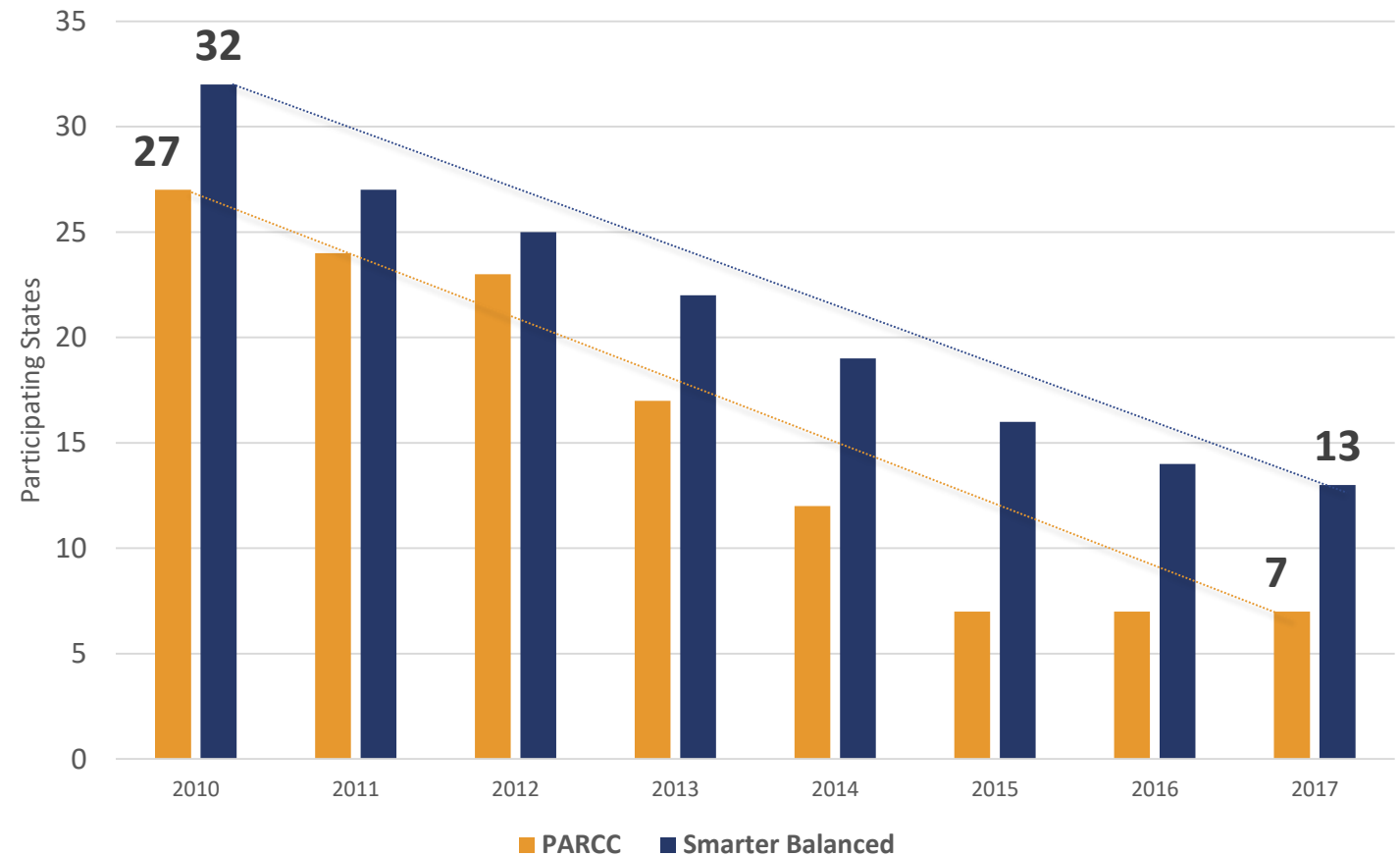
—Amber M. Northern and Michael J. Petrilli, Fordham Institute

... for most states, political considerations, costs and/or length of the test have caused them not to choose (or stick with) PARCC and Smarter Balanced

46 states originally joined PARCC, Smarter Balanced or both consortia, but states' membership or participation has eroded since 2010

- The **fastest decline** in state participation happened **between 2010 and 2015**. The tests were first administered in spring 2015.
- As of late 2017, only **20 states are members** of or are report they are using items from either **PARCC** or **Smarter Balanced**.
- Experts from the field argue that both **PARCC** and **Smarter Balanced** must adapt to the field and offer states more flexible membership options to remain viable but they are **skeptical about whether these “refresh” strategies will “save the day.”**

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



¹ Four states—Colorado, Louisiana, Massachusetts and Michigan—do not use a full or shared consortia assessment form. Louisiana, Massachusetts and Michigan use a blend of consortia questions and their own questions, but all three have remained active in the leadership of their respective consortia—and so we’ve counted them as “members” in this analysis. Colorado, on the other hand, has said it is leaving PARCC but will “likely keep using some PARCC questions” for at least the next few years; we don’t count it as a member in this analysis (*Denver Post*, June 15, 2017).

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*.

Today, 20 states are still substantially using consortia tests (or say they plan to)—although more defections and changes are possible; one expert worried, “Is everyone going to ditch the consortia tests eventually?”

As of December 2017, five states are full members of and two states populate their tests with items from **PARCC** ; 11 states are full members of and two states populate their tests with items from **Smarter Balanced**



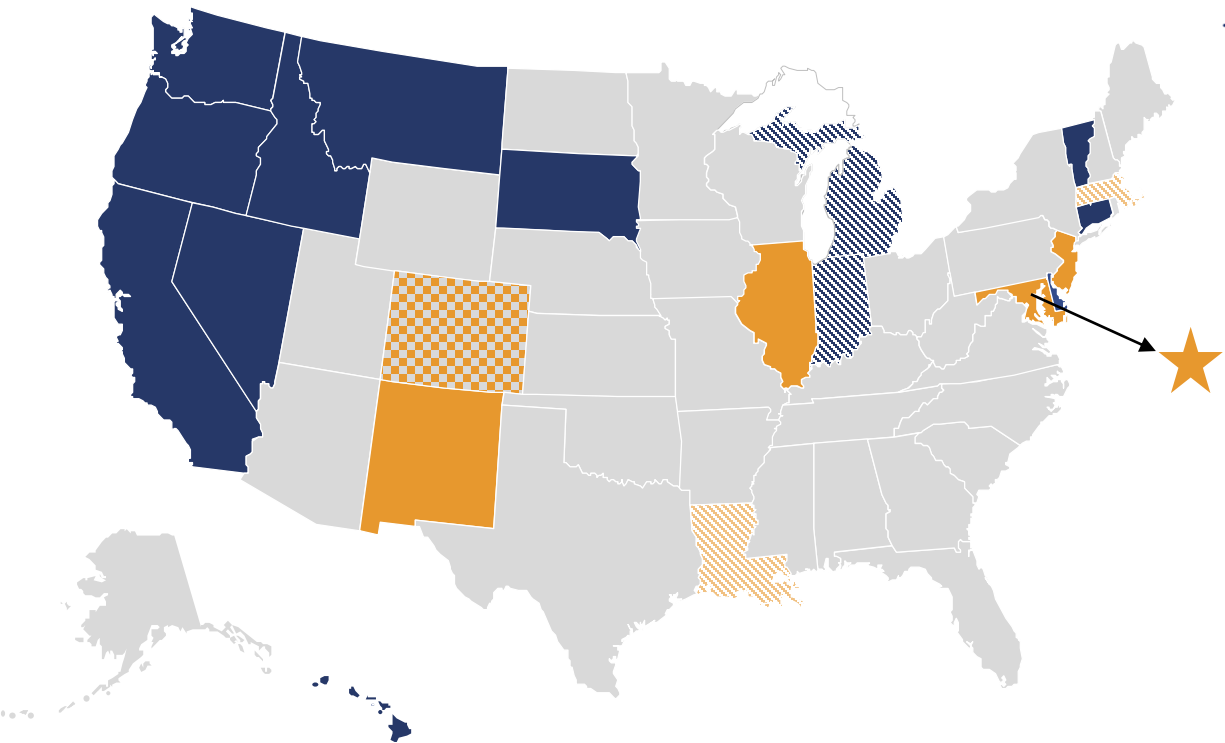
Smarter Balanced is 1-1 on growing state members this past year. While Iowa’s governor over-turned the state board’s decision to use Smarter Balanced, Indiana has meanwhile selected AIR as their new vendor; AIR intends to use Smarter Balanced items in creating the new state assessment (for 2018-19 school year).



Two PARCC states have said they plan to change.

In New Jersey, the newly elected governor opposed PARCC testing during his campaign and has pledged the state will “soon” eliminate its participation.

The Illinois State Board of Education plans to seek bids for a new test design that can be scored more quickly; the new test may still contain some PARCC items



PARCC

State assessment populated with PARCC items

Colorado has left PARCC officially but plans to continue using items for now (because of this status, we don’t count it as a PARCC member)

Smarter Balanced

State assessment populated with Smarter Balanced items

Non-consortia

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

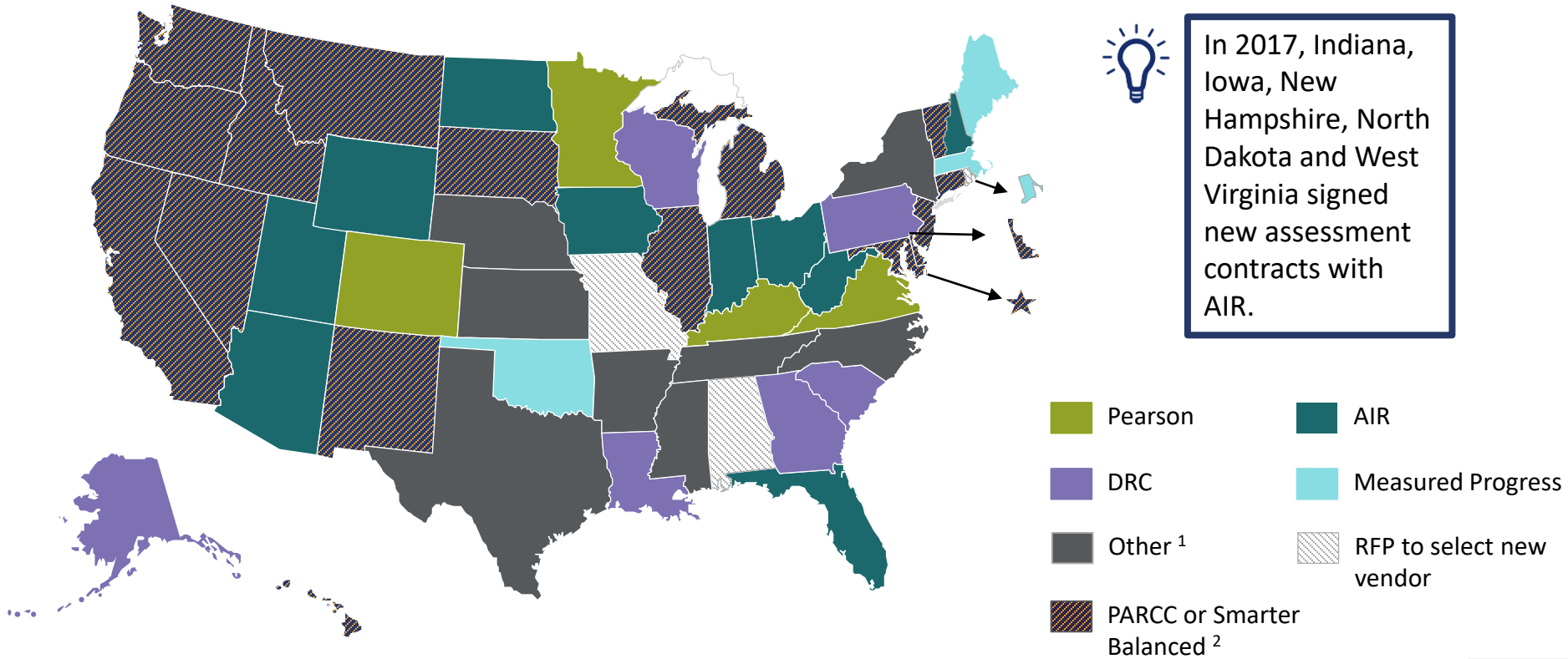
Having dropped out of PARCC or Smarter Balanced, states are venturing out and creating their own assessments once again, with the majority working with vendors AIR, DRC, Measured Progress or Pearson to create their own assessments for ELA and math in K-8

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.

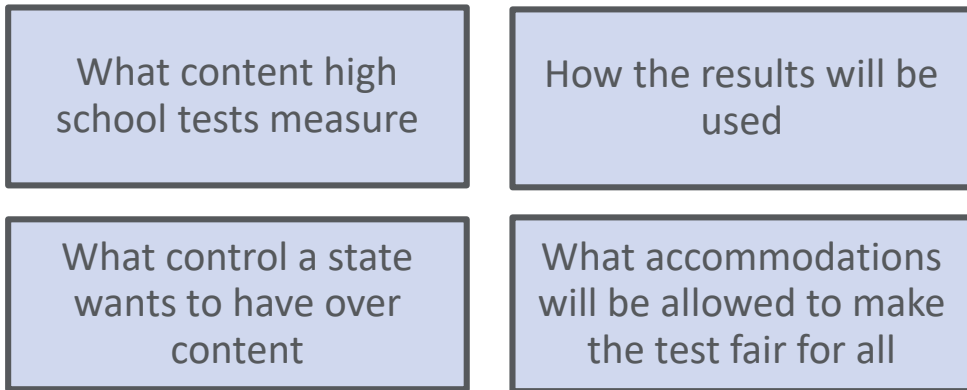
Sources: SEA websites; State assessment directors; Education Week, [*Which states are using PARCC or Smarter Balanced?*](#) (2017).



At the high school level, an increasing number of states (including consortium members) are relying on ACT or SAT, despite questions about these tests' alignment to state standards and about student access to accommodations

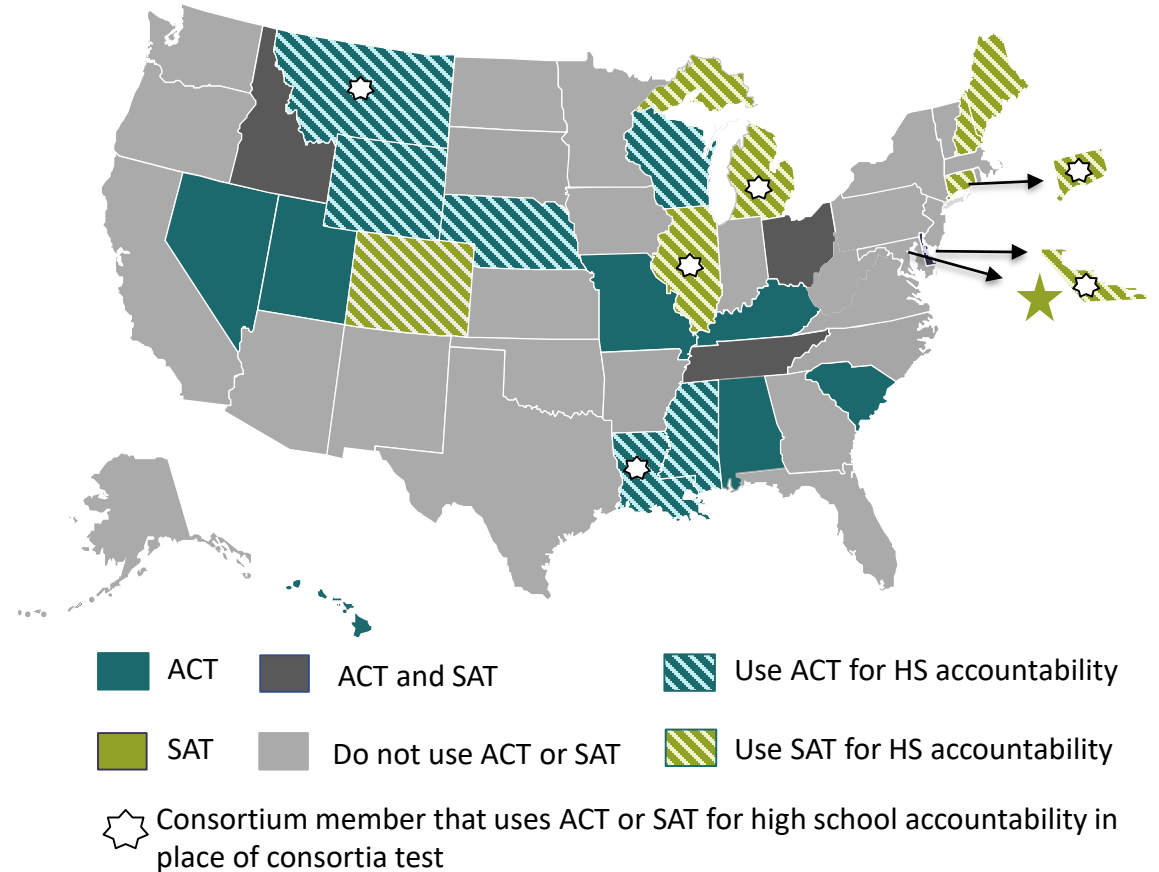
In the 2016-17 school year, the ACT or SAT assessments were mandatory for all high school students in 24 states. In **13 of these states (including six Smarter Balanced or PARCC states)**, ACT or SAT results also are used for school accountability decisions

Experts raise these potential trade-offs for using ACT and SAT for high school accountability



Note: Alabama will not use ACT for the 2017-18 school year

Sources: State assessment directors; Education Week, [Which State Require Students to Take the SAT or ACT? An Interactive Breakdown of States' 2016-17 Testing Plans](#) (2016); Achieve and Center for Assessment, [High School Assessment in a New Era: What Policymakers Need to Know](#) (2016); Erin O'Hara, [Choices and Trade-offs: Key Questions for State Policymakers when Selecting High School Assessments](#) (2016)



So what? As more states have ventured out on their own (to work with new vendors or to adopt off-the-shelf test), questions specifically around the quality and comparability of those assessments are at the top of interviewees concerns

Can a state successfully work alone to create its own high-quality test?

“States who are leaving a consortium say they are still committed to quality but the reality is they are constrained in what they can do alone by cost, by capacity, by time.”

How well do tests like ACT and SAT align to state standards?

“There is pressure on states to use ACT and SAT. It’s important for states to understand what those assessment tell us and what they don’t.”

Is it feasible to compare assessment results if more and more states are ‘going at it alone’?

“What can we say about the progress of student achievement in high school across the country? Not much, because states are all over the place in how they are measuring.”

Do vendors have the capacity to deliver high-quality assessments?

“The quality of test vendors continues to be a huge problem. It is difficult to find vendors who can consistently produce high quality materials. States are really struggling.”

However, importantly, contracts with testing vendors will be expiring during 2017 or 2018 in 22 states¹—presenting potential opportunities to influence and encourage leaders to make good decisions related to high-quality assessments

| States with One or More State Contracts Ending in 2017 and 2018 (Excluding contracts with PARCC and Smarter Balanced ³) | | |
|--|---|--|
| | Grades 3-8 E/LA and Math | High School E/LA and Math |
| 2017 (4 states total) | Alabama, Missouri, Wyoming | Alabama, South Carolina, Wyoming |
| 2018 (7 states total) ⁴ | Arizona, Minnesota, Illinois ^(2,3) , Utah | Arizona, Connecticut ³ , Illinois ^(2,3) , Minnesota, Mississippi, Utah |
| Renews annually (13 states total) | Alaska, Arkansas, Colorado, Kansas, Maine, Ohio, Tennessee, West Virginia | Alaska, Arkansas, Colorado, Delaware ³ , Kansas, Louisiana, Maine, Missouri, Nebraska, Nevada ³ , Ohio, Tennessee, West Virginia |
| Missing information | DC, Kentucky, Rhode Island | DC, Kentucky, Rhode Island |

¹ While the three rows of states in the table above total 24, Missouri is included in both row 1 (for its grade 3-8 tests) and row 3 (for its high school test) and Illinois is included twice in row 2 (for both its grade 3-8 tests and high school tests). So the sum of all states with upcoming vendor contracts is 22.

² Illinois’ SEA has announced plans to seek bids for a new test that can be scored more quickly; the new test may still contain PARCC items, state leaders say, although an RFP has not been released yet

³ Consortia members are listed in the table above only when the state has a non-consortia-related assessment contract expiring (usually contract for its high school test if using ACT or SAT).

⁴ In early 2017, Oregon’s SEA issued an RFI to seek information about possible high school assessment offerings to replace Smarter Balanced; since then, state leaders have decided not to develop an RFP for a new test at this time and instead will be further consulting with stakeholders to examine different options for making the state’s approach matches stakeholder needs.

Over the next year, there will be at least seven new governors elected with the ability to appoint their own, new SEA chief¹; also new governors were just elected in 2017 in New Jersey and Virginia, and a new governor and SEA chief will be elected in California in 2018

Five of the 9 states that definitely will have new governors take office in 2018 and 2019—and where the governor also appoints the SEA chief—are consortia members: **Connecticut, Maine, Minnesota, Nevada, New Jersey, New Mexico, South Dakota, Tennessee and Virginia**

Gubernatorial Elections (2017-2019)

☆ Consortium member

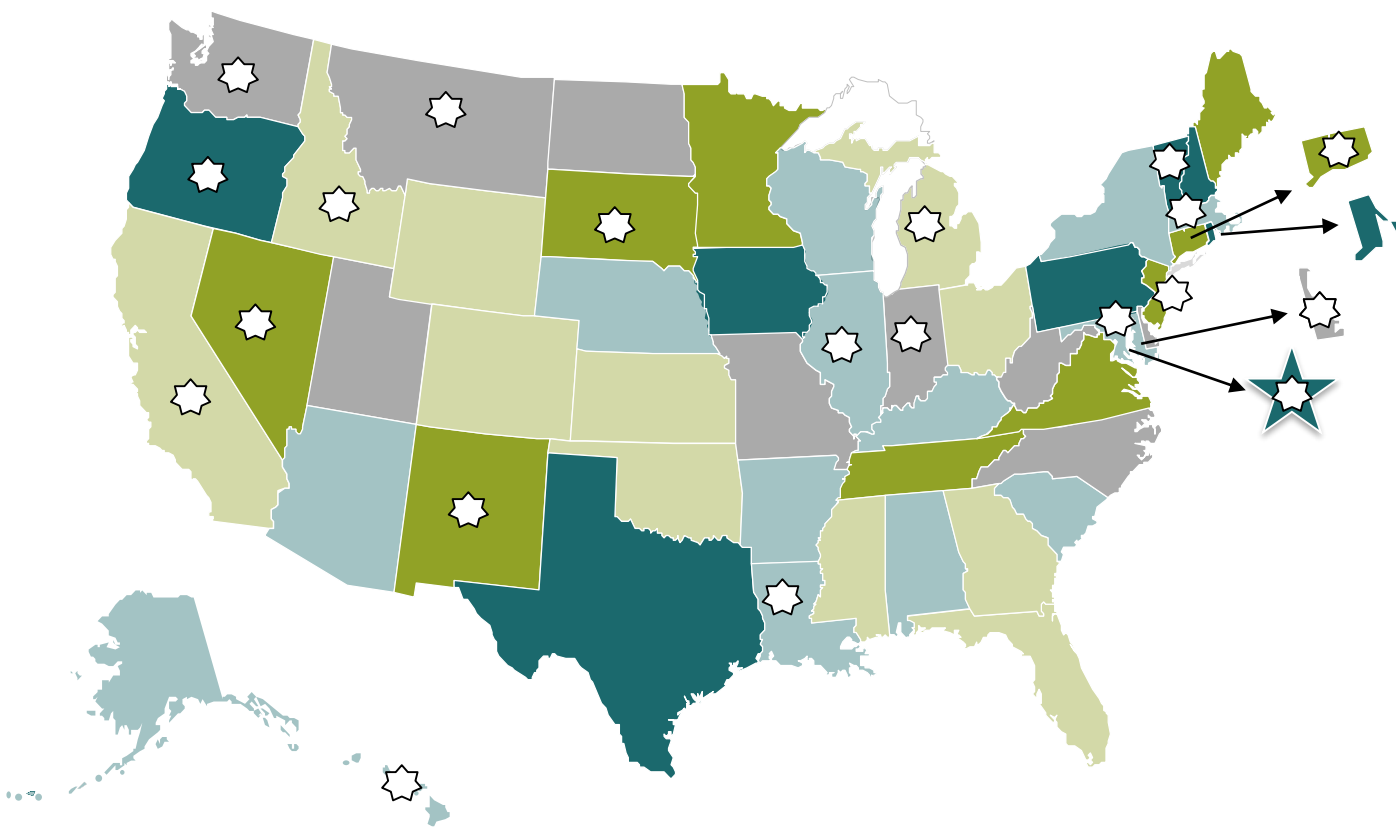
Open seat² & governor has authority to appoint SEA chief

Open seat gubernatorial election

Non-open seat & governor has authority to appoint SEA chief

Non-open seat gubernatorial election

No changes in leadership



¹ The new governors of NJ and VA—elected in November 2017—also have authority to appoint the SEA chief

² Open-seat Governor’s election due to retirement or term limits

Note: In Washington D.C. there will be a non-open seat mayoral election; the mayor has the authority to appoint the SEA chief

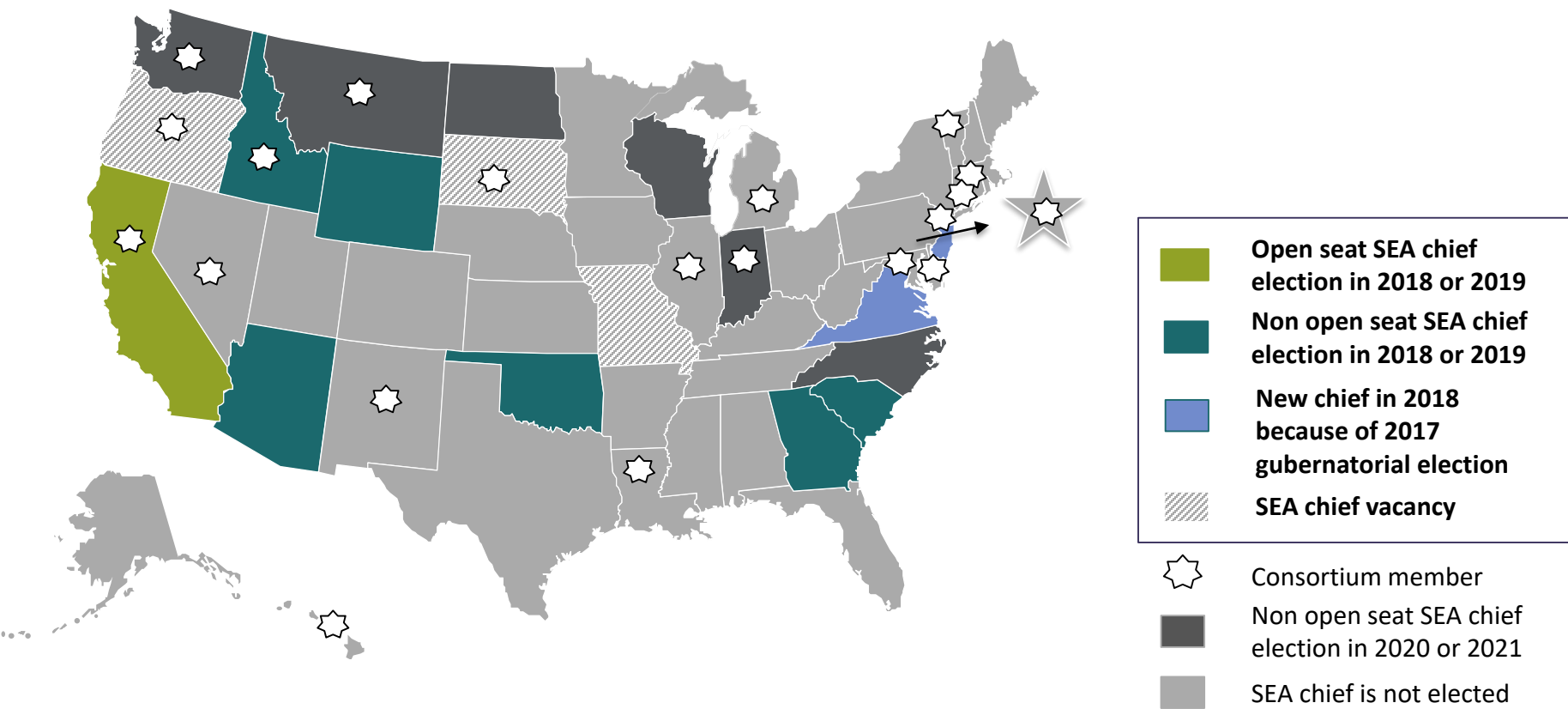
Sources: National Governors Association, *Current Governors* (2017); 270 to Win, *2018 Gubernatorial Election Map*; Education First analysis.

Additionally, 7 states have upcoming SEA chief elections, one of which (California) will be an open-seat election (no incumbent)¹

SEA chief elections in next election cycle:
Arizona, California, Georgia, Idaho, Oklahoma, South Carolina, Wyoming



The SEA chief in South Dakota is retiring at the end of 2017 and the SEA chief in Oregon was fired October 2017 by the governor. Both states have interim SEA chiefs in place. Although not in a consortia, Missouri also has an interim chief.



¹ As shown on the prior slide, NJ (PARCC state) and VA also will new chiefs because of new governors elected in November 2017.

Sources: SEA websites; State Assessment Directors; Education First analysis.

State boards of education also matter: In most states, board members have final or shared authority in approving assessment vendor contracts



37 state boards

Are the assessment authority or share the authority with the SEA or SEA chief



| |
|---|
| 13 states SEA chief is elected |
| 18 states Governor appoints SEA chief |
| 20 states SBE appoints SEA chief |



SBE is the assessment authority or shares authority in 10 of the 13 states



SBE is the assessment authority or shares authority in 10 of the 18 states

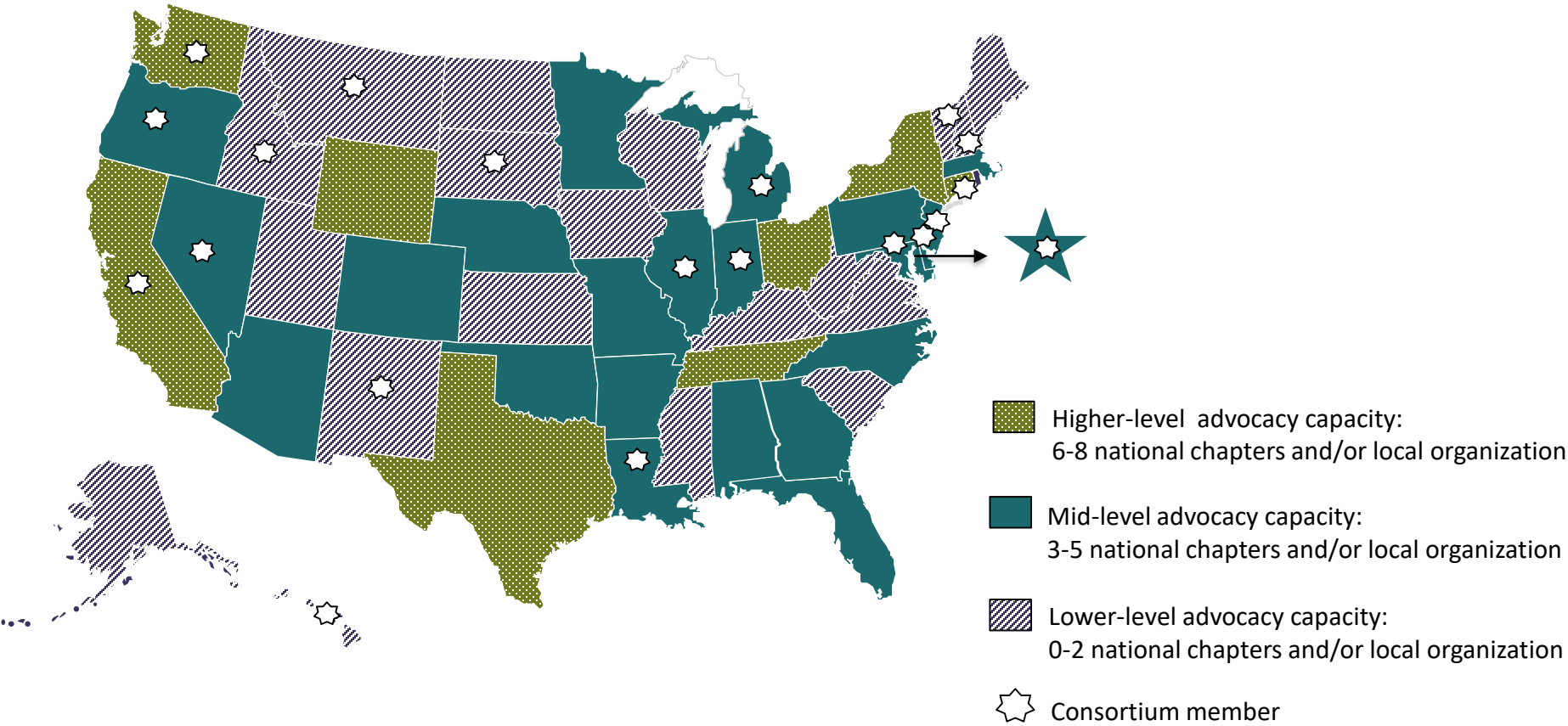


SBE is the assessment authority or shares authority in 17 of the 20 states

Sources: SEA websites; State Assessment Directors; Education First analysis.

Finally, the presence of robust education-reform and civil rights advocacy groups in a state can inform choices about assessment and other accountability decisions—but the number of these groups varies significantly by state

Only 8 states have a higher level of advocacy capacity for high quality state assessments; as a group Smarter Balanced states have varying “advocacy capacity” to tap

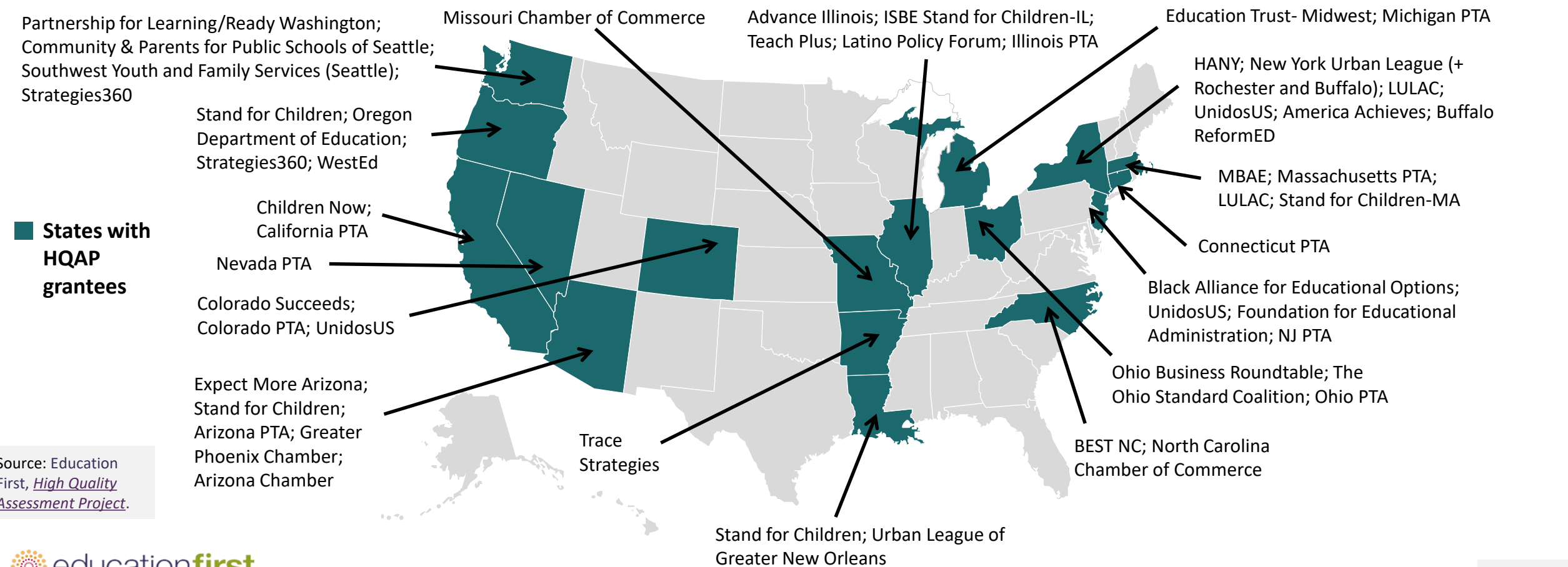


Note: We used the presence of PIE Network members, UnidosUS, Partners for Each and Every Child efforts, and local Urban Leagues to determine the level of advocacy. All of these groups’ members or chapters do not necessarily advocate for high-quality assessments, but serve as a rough proxy for state advocacy capacity on this issue.

Sources: [Pie Network](#); [UnidosUS](#); [Partners for Each and Every Child](#); [Urban League](#); Education First analysis.

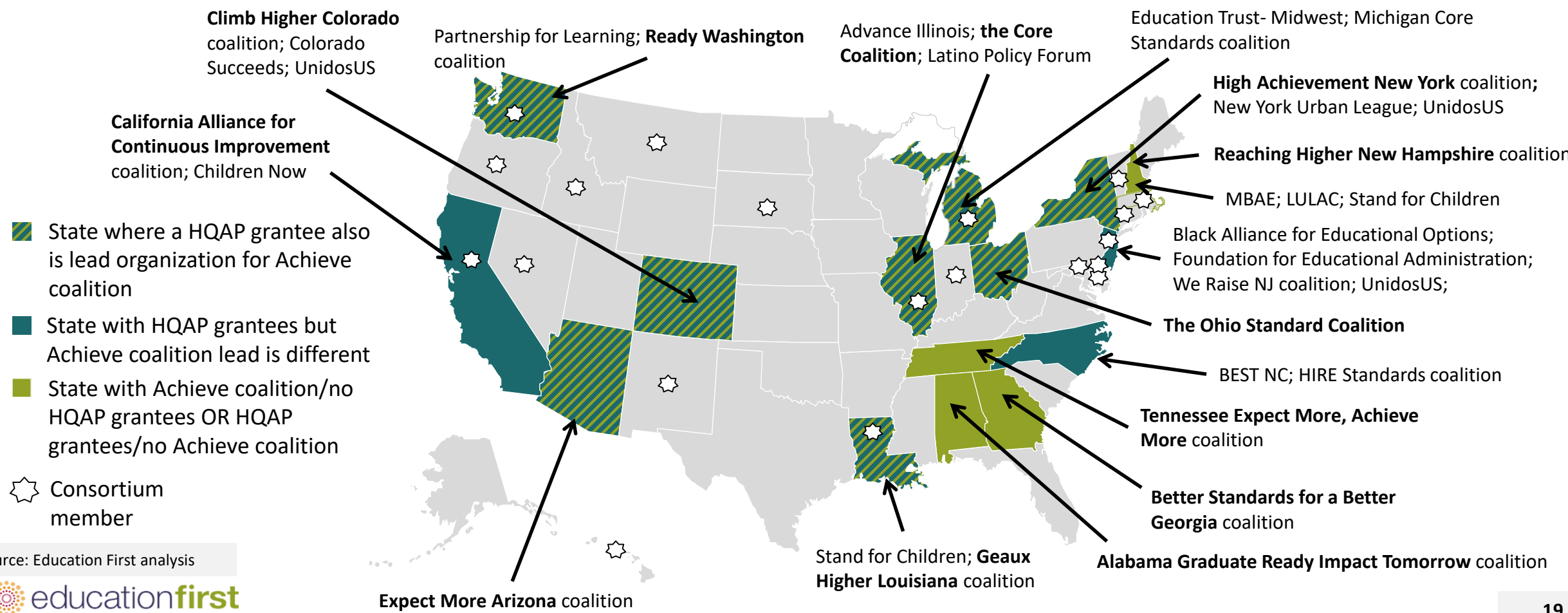
Between 2013-2016, the High-Quality Assessment Project (HQAP) made grants specifically to build state advocates' understanding of testing policies, strengthen their communications on this topic and help them engage new communities

HQAP ultimately supported nearly 50 state-based organizations and coalitions (including several also supported by Achieve) with grants, strategy advice and technical assistance.



Another way to consider in-state advocacy capacity: Ed First identified organizations and coalitions that have continued prioritizing/communicating about high-quality assessments

Specifically, we looked at the 12 state members of Achieve’s “Coalition Support Network” (**current members in bold typeface below**) and the nearly 50 state-level organizations that received support from the [High-Quality Assessment Project](#) between 2013-2016



METHODOLOGY: Our research included desk research, consultations with CCSSO/state assessment directors and discussions with key leaders in the field



Using online research, we compiled state assessment system data; then, with CCSSO, we asked state assessment directors to review and verify info



92%

Our response rate from states to confirm data and fill in gaps was 47 out of 51 states/DC (five that did not verify: **DC**, Iowa, Kentucky, **New Jersey**, Tennessee)

From this data, we compiled a dossier on every state¹

| Alabama | | | |
|---|--|------------------------------------|---|
| HEADLINES | | | |
| <ul style="list-style-type: none">AL dropped ACT Aspire in June 2017Submitted waiver to USED to use interim assessments for the 2017-18 school yearGovernor (re)elected in 2018 | | | |
| STATE CONTEXT | | | |
| Governor | <ul style="list-style-type: none">Kay Ivey (R)4/2017-4/2018; 1st term; 2 max | | |
| SEA | Alabama State Department of Education (ALSDRE) | | |
| SEA chief | Michael Serrano, State Superintendent of Education | | |
| Assessment authority | Appointed by SBE in 8/2018, term limits N/A | | |
| Assessment authority | Alabama State Board of Education | | |
| Consent membership | WIDA | | |
| Advisory groups | <ul style="list-style-type: none">Alabama School ConnectionBirmingham Urban LabAL Education PartnershipNCLD AL AffiliatesNACCT, Alabama State Conference | | |
| STATE ASSESSMENTS | | | |
| | State assessment title | Name of assessment (vendor) | Current contract timelines |
| Grades 3-8 | ELA | ACT Aspire (ACT) | <ul style="list-style-type: none">2013 - 9/2017RFP for new vendor in process |
| | Math | | |
| High school | Science | ACT Aspire Science - grade 7 (ACT) | <ul style="list-style-type: none">2015 - 9/2017RFP for new vendor in process |
| | ELA | ACT Aspire - grade 10 (ACT) | <ul style="list-style-type: none">2014 - 9/2017RFP for new vendor in process |
| | Math | ACT WorkKeys - grade 12 (ACT) | |
| | | ACT with Writing - grade 12 (ACT) | |
| | ACT WorkKeys - grade 12 (ACT) | | |
| High school | Math | ACT WorkKeys - grade 12 (ACT) | <ul style="list-style-type: none">2013 - 9/2017RFP for new vendor in process |
| | Science | | |
| ELL | ACCESS for ELLs 2.0 (WIDA) | ACCESS for ELLs 2.0 (WIDA) | <ul style="list-style-type: none">2013 - 9/2017 |
| Assessment evidence? | | | |
| ACT Aspire partially meet requirements based on the 2017 USED assessment peer review. Insufficient evidence of alignment between the assessments and standards that have been articulated. | | | |
| Expected changes for 2017-18 | | | |
| <ul style="list-style-type: none">AL dropped ACT Aspire in June 2017 (see above). AL will develop a new test to replace itOther ACT Quality Core End of Year Assessments TBDAssessment vendor in 1708 | | | |
| New high school assessment | | | |
| <ul style="list-style-type: none">New ACT Quality Core end of year assessment next year | | | |
| Opt-out | | | |
| State law requires schools to provide all students with the opportunity to take state assessments, but notes specify a statewide opt out policy (see 36-28-01) | | | |
| Assessment plans involving research | | | |
| <ul style="list-style-type: none">AL has requested a waiver from the USED to use interim assessments instead of ACT Aspire in 2017-2018. Waiver is being processed by USED; decision must be made by July 1, for approval, interim assessments must meet certain criteria under ESSA (https://ed.gov) | | | |
| Other important notes | | | |



We also interviewed **leaders in the field** to identify cross-state developments and area of need

- Beth Cocuzza and Jessica Eide, *Student Achievement Partners*
- Michael Cohen, *Achieve*
- Linda Darling-Hammond, *Learning Policy Institute*
- Catherine Holahan and Kathryn Young, *EducationCounsel*
- Scott Marion, *Center for Assessment*
- Julie Mikuta, *Charles and Lynn Schusterman Family Foundation*
- Scott Norton, *Council of Chief State School Officers*
- Hannah Skandera, *formerly with New Mexico Public Education Department*

We are grateful to these individuals for their time and contributions to this research and analysis

¹ See separate PPT file where all 51 state profile slides are compiled

Thank you!

www.education-first.com

- William Porter, partner (*bporter@education-first.com*)
- Kathleen Callahan, analyst (*kcallahan@education-first.com*)