

WHITE PAPER

Convening for a common cause: Reimagining assessment and accountability to improve student learning



Introduction

In the spring of 2020 the coronavirus pandemic and school closures upended public education in the United States, causing the annual summative assessments required under the Every Student Succeeds Act (ESSA) to be waived for all 50 states. In the months following, education stakeholders have weighed in on how best to support students' academic and social-emotional needs, including how to capture what students have learned and what learning remains unfinished.

Many of the resulting discussions reignited longstanding debates around the quality and administration of assessments and their use in school accountability systems. However, what emerged for one group of education stakeholders was an opportunity to rethink both assessments and accountability to ensure systems are driving deeper learning and providing educators and policymakers with data that is useful for supporting students, improving student outcomes, targeting resources, and guiding new instructional policies and practices.

Convening Overview

As a result, Education Reform Now and NWEA® convened over 100 individuals from more than 60 organizations to discuss the critical issues facing K-12 educational assessment and accountability. These individuals represented a wide spectrum of perspectives, beliefs, and experiences on these topics and included leaders from state departments of education, school districts, educator and parent groups, civil rights groups, universities, policy and advocacy groups, research organizations, and assessment vendors and experts. While their approaches may be different, what they have in common is a shared goal of seeing increased opportunities and improved outcomes for students. The goal was not to solve everything, but to examine the issues, hear from diverse perspectives, and see if there were critical areas of alignment across participants.

During three separate meetings held in the final months of 2020, participants were broken into small groups to discuss two topics: the immediate concerns around assessment and accountability that states and districts face due to the pandemic in the current (2020–21) school year, and a more long-term view of how to improve our current system of assessments and accountability.

This document provides policy recommendations where there was relative consensus and policy alternatives where opinions diverged. The document is intended to help federal and state stakeholders address the near-term implications of COVID-19 and inform changes in assessment and accountability systems over the long term in ways that will drive continuous school improvement and, ultimately,

better outcomes for all students—particularly students who have been the most deeply impacted by the pandemic, including students of color, students from low-income backgrounds, English learners, and students with disabilities—a goal shared by all convening participants.

Document Guide

This document is organized around four main **challenges** that surfaced in conversations across organizations. It offers **overarching policy recommendations** based on the conversations among convening participants. For each challenge, the document provides more specific **policy recommendations for federal and state policymakers**, intended to represent the general consensus of the group. However, individuals and organizations that participated in the convening may not support every individual recommendation included in this document.

Given the diversity of views among convening participants, there were also many ideas shared over the course of the three meetings that, while worthy of further investigation, did not represent a consensus view among participants. Some of these are more specific proposals that relate to the policy recommendations but on which there was not consensus. These alternatives are included in the document as **ideas under debate** to guide further thought and discussion on potential avenues for federal and state policy. This document also includes articles (see pages 6, 8, 12, and 14) that provide additional context on a number of the policies under review. Any policies described in these articles may also include non-consensus ideas unless specified otherwise.

Challenge

Overarching Recommendation

SPRING 2021

States and districts are struggling to understand how students, educators, and schools are doing and how they can most equitably apply interventions and distribute resources.

Provide federal and state support to states and districts for identifying gaps in learning outcomes and growth, implementing appropriate evidence-based interventions, and targeting resources to students who have been the most deeply impacted by the pandemic—including students of color, students from low-income backgrounds, English learners, and students with disabilities—in an equitable manner that reflects local context and need.

BEYOND 2021

How can we envision accountability systems so that the data they provide are more actionable for a variety of stakeholder groups in providing high-quality and effective learning opportunities and resources? How can we make sure they offer responsive support to educators and families, and drive the improvement of student outcomes?

Reframe approaches to accountability, including the role of assessments in accountability systems, to better recognize growth and the narrowing of gaps in opportunity and outcomes over time.

ESSA places a great deal of the onus for change at the level of individual schools when it may be that systemic policies at the district and state levels—including determinations of allocations of funds, staff, and other resources—are most in need of change.

The United States Department of Education (USED), Congress, and education stakeholders should work together to reform accountability and improvement systems so that states and districts are held accountable, respectively, for appropriate actions that influence student success. A feedback loop between schools, districts, and states should be created so that goals and requirements are met, with the provision of sufficient, effectively targeted resources and the enactment of evidence-based policies and practices.

Annual statewide summative assessments were established in part to serve as an indicator of whether students are receiving a high-quality education and to identify schools most in need of support.

Develop assessments that provide students with more equitable opportunities to show what they know and produce results that are more timely and actionable for policymakers, educators, families, and students.

However, many states are struggling to close gaps, suggesting that for some states data from statewide assessments may not be as actionable as intended. Many assessments also lack cultural relevance.

Spring 2021

Across the country, families, students, and educators have had to make significant adjustments in order to keep learning going, from the wide use of remote and digital learning to how school meals are provided. Well into the 2020–21 school year, it is clear that states, districts, and schools need more guidance and support from federal and state governments, in alignment with the significant adjustments that have been made. This is particularly true for accountability and assessment requirements. While some of the participants thought that states should be required to administer assessments this spring, this was not a unanimous opinion. However, given the recent [USED decision](#) not to waive assessment, and their guidance on some flexibilities states may consider, recommendations from the group that may help contextualize data to provide better guidance and support student learning and development are included below.

Challenge

In a school year unlike any other, states and districts are struggling to understand how students, educators, and schools are doing and how they can most equitably apply interventions and distribute resources.

Overarching Recommendation

Provide federal and state support to states and districts for identifying gaps in learning outcomes and growth, implementing the appropriate evidence-based interventions, and targeting resources to students who have been the most deeply impacted by the pandemic—including students of color, students from low-income backgrounds, English learners, and students with disabilities—in an equitable manner that reflects local context and need.

FEDERAL RECOMMENDATIONS

Assessments

- While USED has already provided some initial guidance, it should provide states with additional guidance on how to appropriately contextualize, interpret, and use assessment results, given the likelihood that fewer students will be assessed and the mode

may be different. It should also describe any federal support that will be provided to aid these efforts.

Accountability

- USED should encourage states to expand and monitor opportunity to learn (OTL) indicators—both inputs and outcomes (see the article on page 6)—and, to the extent possible, temporarily introduce new indicators reflective of the unique circumstances created by the pandemic into their support systems. This may include providing guidance on ways to facilitate the collection of these data without increasing burden.
- Consistent with existing guidance, USED should consider flexibility around the use of accountability for identifying new schools for improvement, while maintaining requirements for disaggregated and publicly reported data by individual student group.
- USED could provide guidance on how states can temporarily change the weight of summative assessments in their accountability systems—disaggregated and publicly reported by individual student group—and support a role for OTL indicators to contextualize data from these systems.



Appropriate Use of 2021 Summative Assessment Data

Any data from the 2021 summative assessments should be interpreted and used cautiously as students—even within the same class—may have very different learning experiences. To support their appropriate use, states should:

- Use data to determine resource allocation of relief funds and future investments in equitably educating students with the highest need, particularly students who have been the most deeply impacted by the pandemic, including students of color and students from low-income backgrounds.
- Collect and use critical OTL indicators to contextualize any assessment data.
- Report, as required by ESSA, the assessment participation and non-participation rates disaggregated by subgroup.
- Avoid the use of assessment data for purposes such as school identification or punitive action.

STATE RECOMMENDATIONS

Assessments

- States should work with districts to provide support and flexibility around how to appropriately interpret and contextualize results.

Accountability

- States and districts should use state assessment results in conjunction with other contextual data to inform planning, interventions, and resource allocation to students who have been the most deeply impacted by the pandemic, including students of color and students from low-income backgrounds, instead of using results in connection with decisions such as school identification, educator evaluations, grade promotion, or graduation. States should clearly communicate how assessments will be used to districts, educators, parents, students, and the community.
- State leaders could partner with districts, and seek input from educators and families, to identify the most critical and efficiently collected OTL indicators—disaggregated and publicly reported by individual student group—that should be used to provide context for student assessment data. These critical OTL data can be used to help identify schools and students most in need and inform the appropriate intervention to

support students who have been the most deeply impacted by the pandemic, including students of color, students from low-income backgrounds, English learners, and students with disabilities.

Ideas for State Action Under Debate

- States that use summative ratings (such as A-F) could put their rating on hold for a year and provide other relevant data, or adjust their systems to provide a more holistic understanding of the conditions for learning and student achievement, such as a dashboard approach.

Opportunity to Learn

Opportunity to learn (OTL) indicators include measures of the resources and conditions that matter most for student learning, providing information to drive school, district, and systems-wide improvement. Whether or not states choose to formalize the collection and reporting of [a set of well-chosen OTL indicators](#)—disaggregated and publicly reported by individual student group—OTL indicators can be a tool for equity during and beyond COVID-19, with districts and states using this information to close gaps in students' access to the resources and conditions necessary for success.

Prior to the pandemic, a number of states were including a variety of contextual data, such as OTL and equity measures, in their accountability and reporting systems as indicators of school quality. These indicators often included: school climate, access to advanced coursework, student engagement, the provision of well-rounded curriculum, suspension rates, chronic absenteeism, and more.

To support student learning through the pandemic and beyond, states and districts can identify and act on closing longstanding gaps in OTL, including those OTL and equity [indicators](#) that may be more specific to learning in the time of COVID-19. Such indicators may include:

- access to adequate connectivity and devices
- mode (in person, hybrid, or distance), length, and delivery (synchronous or asynchronous) of instruction
- COVID cases
- equitable distribution of experienced, in-field, and effective educators
- adherence to health and safety guidelines
- services for specific student subgroups (e.g., English language learners, students with disabilities)
- expanded learning time
- changes in district and school enrollment
- participation in district and/or state assessments
- suspension and expulsion rates
- funding within and across districts
- supports for historically underserved students (e.g., students experiencing homelessness, students of color, students from families with low incomes)

It is important to note that OTL indicators are best suited for reporting purposes and to inform improvement efforts and should not be used for accountability purposes. Additionally, while some OTL measures can inform school-level actions, others should inform district-level actions and decisions and, as such, shouldn't be included as school-level indicators.

In short, OTL indicators are essential to supporting student learning, closing equity gaps, and contextualizing academic outcomes during and beyond COVID-19.

Beyond 2021

The pandemic has illuminated persistent disparities in educational opportunity and outcomes while simultaneously necessitating new ways of approaching teaching, learning, assessment, and accountability. Policymakers can either address these circumstances as temporary, with the intent to eventually return to “business as usual,” or they can embrace the opportunity to effect longer-term change so that historic challenges can be addressed. While there are diverse opinions on exactly what should change and how, many agree that now is the time to revisit and reevaluate longstanding policies and assumptions to identify what has been working well and what should change—and to take action to transform the system and its outcomes.

Challenge

State accountability systems have evolved since ESSA was ratified in 2015. ESSA provided states with much more flexibility around how to identify and intervene in schools that are identified as in need of support—especially the degree to which those decisions are based on assessment results—and allowed the use of many new measures to identify and support schools. Some advocates assert that these changes have greatly lessened the degree to which schools identified as in need of improvement are pressed to address persistent achievement and opportunity gaps, especially for students from historically disadvantaged groups. Others feel that state accountability systems, even under ESSA, still rely too heavily on assessment results and have too many of the same triggering mechanisms and consequences that they see as punitive and counterproductive. It is our intent that airing and exploring these different viewpoints will lead to more informed, thoughtful, and nuanced policymaking.

How can we envision accountability systems so that the data they provide are more actionable for a variety of stakeholder groups in providing high-quality and effective learning opportunities while still providing comparable data to policymakers for targeting resources to the students and schools most in need? How can we make sure they offer responsive support to educators and families, and drive the improvement of student outcomes?

Overarching Recommendation

Reframe approaches to accountability, including the role of assessments in accountability systems, to better recognize growth and the narrowing of gaps in opportunity and outcomes over time. This should include integrating the use of OTL and other equity indicators that are evidence-based,

relevant, and able to leverage improvement in student outcomes into public reporting systems. In conjunction with academic achievement data, and without compromising high expectations for all students and educators, these indicators could produce school report cards that are more comprehensive, representative, and actionable than the accountability systems that exist in many states.

FEDERAL RECOMMENDATIONS

- Congress should provide funding to support states in developing data systems that include OTL indicators, as well as monitoring and reporting OTL indicators (disaggregated and publicly reported by individual student group). This should include helping states determine which indicators are evidence-based, which are the most relevant, and which can leverage the most significant improvements in student outcomes.

STATE RECOMMENDATIONS

- States should offer support to districts to include the use of additional OTL indicators—disaggregated and publicly reported by individual student group—that are relevant to improving student learning to help contextualize academic data and root improvement plans in equitable access to learning opportunities.
- States should assess and consider revisions to their ESSA plans to ensure a better balance between the weight of growth and achievement as academic indicators in their accountability systems.

The Role of Academic Growth in Teaching, Learning, and Accountability Systems

ESSA helped to establish multiple measures in state accountability plans, and student academic growth stands out as one of those critical measures. Forty-eight states and the District of Columbia have included a growth measure as part of their ESSA plans, marking a significant improvement in efforts to gain a better understanding of student and school performance. There is [variability](#) in how states measure growth, but one aspect that is common across all of the measures is that they are rooted in student performance on annual state summative tests, which assess grade-level proficiency. Growth measures based on this data are important in supporting efforts to understand how schools are doing in helping all students make progress in reaching or exceeding rigorous academic standards. However, states may wish to expand these measures to include growth across all levels of performance—rather than just measuring differences in levels of “proficiency”—and across longer timespans; they may also wish to reexamine the [weight applied to growth](#) and achievement in their accountability systems.

States looking to rebalance and refine the role of indicators such as growth in their accountability formulas should take a systemic approach to the process. This is because the weights assigned to each indicator (such as growth and achievement) in an accountability system may not represent their actual influence on the composite school score used for school ratings. The score distributions of the indicators can result in one having a much greater impact than the other, even if both are weighted evenly.

In addition, states may also want to evaluate their growth indices to ensure a match with intended use and interpretation. Criterion-referenced growth systems importantly focus on growth to proficiency, but may prevent recognition of schools with many rapidly progressing students who started far behind, or miss schools where proficient students are making little growth. Some critics suggest that these systems incentivize a narrow focus on the growth of students who are “on the bubble” of being proficient.

Conversely, norm-referenced growth measures, which compare the growth of students to each other (either at the state or national level), can provide insight into where systems are making “more than normal” growth as they work to address gaps, but they can lose sight of whether that growth is enough to get students to proficient or beyond over time. Critics of this approach suggest that it can provide a false sense of achievement for schools, educators, and students and result in misidentification of schools as successful when they are not growing students enough or at all.

Both criterion and norm-referenced growth provide valuable information to help address system needs but should be considered with awareness of their unique drawbacks. Including both types of growth measures to provide a well-rounded view without compromising efficiency and clarity is an area for exploration in the realm of assessment and accountability innovation.

States should also consider exploring the timespans they are using to measure growth. Circumstances created by the pandemic, such as the suspension of summative tests in spring of 2020, have surfaced discussions about using “[skip-year growth](#)” to understand how schools are doing in supporting student progress. Some view this as a [temporary approach](#) in response to pandemic conditions, while others have suggested, even prior to the pandemic, that a more

fair and accurate way of assessing school performance would be to look at [growth over periods longer than two years](#) to better align existing short-term accountability models with the longer time periods required for schools to prepare students for success. Some systems have already started to explore some of these opportunities such as value added systems that explore student and system growth over multiple years. Lessons learned from these types of growth systems can help inform future conversations, and states using such systems should be supported in understanding how to best contextualize data in light of pandemic-related disruptions.

Growth measures are important not only at the state level but also at the district level. Many districts measure a different kind of growth in student progress by using interim or benchmark assessments administered in the fall, winter, and spring. In aggregate, this information helps administrators understand how much learning is occurring during the school year, across all achievement levels, to inform district program and policy decisions. At the individual student level, it can support educators and students in setting challenging learning goals.

As states embrace balanced systems that include support for assessments beyond the annual summative test, they may wish to consider including within-year growth measures in addition to those rooted in summative performance. These measures may help to more effectively account for [summer learning loss](#) and for growth of students who are not yet proficient as well as those who are already proficient. These measures help to reveal schools with low proficiency rates but high growth and those with high proficiency rates but low growth.

Challenge

ESSA places a great deal of the onus for change at the level of individual schools when it may be that systemic policies at the district and state levels—including determinations of allocations of funds, staff, and other resources—are most in need of change.

Overarching Recommendation

USED, Congress, and education stakeholders should work together to reform accountability and improvement systems so that states and districts are also held accountable, respectively, for appropriate actions that influence student success. A feedback loop between schools, districts, and states should be created so that goals and requirements are met, with the provision of sufficient, effectively targeted resources and the enactment of evidence-based policies and practices.

FEDERAL RECOMMENDATIONS

- Congress and USED should work together to condition new funding on how well states and their districts target their own funding and policies, based on poverty and student needs, availability of select OTL, and other outcomes.
- USED should apply oversight to ensure that states are in compliance with the equity provisions included in ESSA. While ESSA is not perfect, it was a step forward in supporting the use of multiple measures to: provide a more accurate and holistic view of student and school performance; hold all students to high standards; provide information on how subgroups of students are doing via disaggregation of data; include OTL indicators; help English learners make progress toward proficiency; and increase student access to experienced, in-field, and effective educators and fiscal accountability provisions, like supplement not supplant. Ensuring that states are complying with these aspects of the law is an important part of turning the system into one that improves outcomes for students.
- USED should explore ways to ensure that states are following the spirit of the law not just in word but also in deed, such as ensuring

that data collected are used in support of evidence-based interventions identified and selected in collaboration with stakeholders (including educators and families) that provide needed supports, particularly the oversight of schools identified for Targeted Support and Improvement. USED could also evaluate if and how states are engaging with the community stakeholder groups outlined in their ESSA plans.

Ideas for Federal Action Under Debate

- Congress and USED could put forth new options for accountability and improvement systems, such as layered or differentiated accountability systems that go beyond the school level to incorporate state or district levels of accountability. These could include a state-level equity check centered on driving resources and district-level systems that include additional measures of school climate.
- In a reauthorization of ESEA, Congress could include equity indicators in reporting for grades K-2—such as access to high-quality pre-K; effective, experienced, and in-field educators; and equitable funding—given the paramount role of early learning in later success in school and the relevance of ensuring that students and schools are resourced properly to facilitate student learning during this critical period.

STATE RECOMMENDATIONS

- State and district leaders should use OTL, equity, and academic data to undertake and inform funding reforms.
- States should partner with districts to identify core district-level OTL and equity indicators (disaggregated and publicly reported by individual student group), support them in the efficient collection of these data, and provide additional resources needed to help districts in taking action based on indicators, in conjunction with academic data.
- State and district leaders, as well as assessment providers, should authentically incorporate community voices—especially educators, families, and community-based, culturally specific organizations representing groups facing inequities—into the conversation and policy development.

Challenge

Annual statewide summative assessments were established in part to serve as an indicator of whether students are receiving a high-quality education and to identify schools most in need of support. Assessments should be designed to provide systems-level data to school leaders, district administrators, and state policymakers to inform the appropriate policy-level decisions. Statewide summative measures are key equity guardrails and they shine a light on inequities, particularly for students of color, students from low-income background, English learners, and students with disabilities. However, many states are struggling to close gaps, suggesting that for some states data from statewide assessments may not be as actionable as intended.

Overarching Recommendation

Develop assessments that provide students with more equitable opportunities to show what they know and produce results that are more timely and actionable for policymakers, educators, families, and students.

FEDERAL RECOMMENDATIONS

- Actively promote assessment innovation and allocate more funding to support the development of high-quality assessments that provide meaningful data that can be disaggregated by student subgroups. Provide more guidance and technical assistance to states in interpreting what the law and requirements allow, including examples of feasible innovative models that states could pursue, with an emphasis on models that drive deeper learning and are efficient, cohesive, and actionable for state and district stakeholders.
- Incentivize assessment innovation and make it less burdensome for states to pursue. This means making the following changes to the Innovative Assessment Demonstration Authority (IADA) program:
 - Create a funded planning grant opportunity for states that are interested in pursuing IADA so they can consult and engage with stakeholders before participating in the pilot.

- Provide sufficient funding and technical assistance to IADA program participants.
 - Lengthen the timeframe for implementing and scaling new assessments.
 - Reevaluate technical/peer-review requirements in relation to innovative assessment goals and advancements in the field (see article on page 14 for more information).
 - Convene states pursuing innovation to share learnings and problem-solve together.
 - Require a description of the professional learning and assessment literacy that will be provided as part of innovative assessment plans.
- Provide funding and support to incentivize innovation efforts beyond IADA and to improve engagement with communities in the innovation process.
 - Increase funding available to states to improve their assessment systems through activities such as assessment audits and statewide implementation of innovative assessments—approaches that do not require participation in the IADA program.
 - Allocate funding for educator training and professional development on the effective use of assessment data and design to empower educators to meaningfully participate in and benefit from the innovation process.



Ideas for Assessment Innovations

As states, districts, communities, and assessment vendors and organizations collaborate to develop new assessment systems, they will need to resolve the tension between the need for instructionally relevant data for educators and parents and comparable systems-level data for administrators, parents, and policymakers. States may ultimately need to build systems that include multiple innovations or design summative tests that provide systems-level data in conjunction with more localized assessments that provide data to inform instruction.

Innovations that aim to bridge district and state assessments could include:

- *Performance-based assessments* (as well as competency- or mastery-based assessments) that promote and measure deeper learning.
- *Curriculum-aligned and interdisciplinary assessments* that limit the influence of students' background knowledge and better align results with teaching and learning.
- *Multifunction assessments* such as through-year models (whether computer-adaptive or performance-based) that produce more than one type of data to different stakeholders at multiple points during the school year.
- *Allowing off-grade items* in summative assessments to better gauge students' level of achievement and the prior-grade building blocks needed to attain grade-level standards while retaining sufficient levels of grade-level content.

It is important to note that many advocates have concerns that some or all of these approaches may negatively affect test validity, limit the availability of grade-level individual scores, and undermine the disaggregation of achievement data for each school by student subgroups (by reducing "N" sizes), a bedrock principle of assessment systems for many civil rights and disability groups.

Innovations focused on reducing the burden of summative tests while keeping them separate from assessments designed to inform instruction include:

- *Shortened summative assessments* that reduce the number of items and/or standards tested to what is most critical.
- *The matrix-sampling approach to summative testing*, which tests all students against grade-level proficiency in all subjects, and ensures that across a state, school, or district, sub-score reporting representative of the full depth and breadth of standards is possible.

STATE RECOMMENDATIONS

- Take advantage of existing opportunities to improve assessment systems and/or implement innovative approaches, including:
 - ESEA's assessment audit option, which can help eliminate low-quality, redundant, or misaligned tests.
 - The Competitive Grants for State Assessments Program to enhance the quality of assessment instruments and systems with areas of focus such as improving approaches to measuring growth and achievement and improving assessments for English learners and students with disabilities.
 - Implementing innovative assessments statewide, outside of the IADA program (when assessments are implemented statewide rather than with a limited number of districts or schools, participation in IADA is not required).
- Pursue assessment innovation with a focus on one or more of the following (see article on page 12 for examples):
 - Streamlining and increasing the efficiency of assessments (understanding that there is no single assessment that can fulfill all purposes).
 - Developing assessments that are fully aligned with curriculum and measure deeper learning.
 - Avoiding challenges to local control of curriculum and pacing, preserving district choice while at the same time promoting high-quality, standards-aligned curriculum.
 - Improving the utility, including timing, of assessment reports for educators, parents/caregivers, and students so that they are easy to understand and feature clear guidance on actions that support student learning.
 - Engaging educators, families, culturally specific organizations, and assessment experts in the creation of systems that effectively address cultural and local relevance without compromising the provision of data that enables comparisons across student subgroups and between schools and districts.
- Engage educators in the innovation process and fund educator training and professional development in conjunction with innovative assessments so data is used and applied to actions that support student learning.
- Communicate with and engage educators, families, policymakers, and other public stakeholder groups in discussions about the purpose of assessment, the desired outcomes of assessment innovation, and the relevance of data in informing teaching, learning, and systems-level decisions to support better outcomes for students.
- Highlight the proper use of data and tangible examples of how it can inform and drive improvement, including equitable funding, resource allocation, and evidence-based policies.
- Support educators in designing, implementing, and using assessments and data to improve instruction.

Aligning Technical Requirements and Innovation Goals

The goal of assessment innovation is to make improvements and effect positive change. Standards for educational testing that ensure valid, reliable data are a critical part of this effort. Some suggest, however, that certain aspects of the technical requirements such as “comparability” may merit revision in connection with innovation. Others suggest that revising comparability requirements would impact the validity and reliability of the data. This is a complex discussion, as there are [multiple types of comparability](#) to be considered. Two aspects that are often discussed in connection with innovation challenges include the following:

- Comparability across groups: Within an assessment, student scores mean the same thing for all groups of students, regardless of student characteristics like race, gender, ability status, etc., and groups are consistently defined so that they can be compared across jurisdictions.
- Comparability across different assessment systems: How well two different assessments measure the same construct and how well scores can be compared.

Many advocates consider the latter as a precondition for the former. Others assert that requiring comparability across different assessment systems—or demonstrating the “interchangeability” of student scores between new and old tests—makes true innovation challenging, as it is akin to an attempt to change things while keeping them the same.

To create space for innovation, USED could convene technical experts to examine whether a different approach to demonstrating comparability across new and old systems could be possible, while still providing assurances of comparability across groups—i.e., ensuring that students (and schools/districts) participating in pilots are being held to similar expectations as students (and schools/districts) in other parts of the state.

Again, some advocates and psychometricians believe that comparability across assessment instruments is a requirement for comparability across different groups. Others suggest that considering the theories of action for each innovative pilot could help inform determinations of what is “comparable enough” and that to support these conditions, USED could extend the IADA pilots to allow for two initial funded “planning” years. During the planning phase, states could focus on developing a theory of action that clearly articulates how the innovation fits into and advances that state’s vision for teaching and learning.

Appendix: Resources for Further Reading from Convening Participants

Assessment in Spring 2021

- [American Succeeds—Student-Centered Policy Priorities for 2021](#)
- [CCSSO—Restart & Recovery: Assessments in Spring 2021](#)
- [Center for American Progress—Student Assessment During COVID-19](#)
- [Center for Assessment—This Is Not a Test, This Is an Emergency](#)
- [Data Quality Campaign—COVID-19 Recovery Requires Education Data](#)
- [Education Reform Now—Why Local Tests Should Not be Used in Lieu of Annual Statewide Assessments—This Year, or Any Year](#)
- [FutureEd—Blueprint for Testing: How Schools Should Assess Students During Covid](#)
- [Migration Policy Institute—Educating English Learners During the COVID-19 Pandemic](#)
- [National Academy of Education—Educational Assessments in the COVID-19 Era and Beyond](#)
- [New Classrooms—The First 100 Days and Beyond](#)
- [NWEA—K-12 Assessment Policy Recommendations for the 117th Congress](#)
- [SchoolHouse Connection—Statewide End-of-Year Assessments and Students Experiencing Homelessness](#)
- [Seek Common Ground—Families and Communities Key to Assessment and Accountability Decisions During the Pandemic School Year](#)
- [Tennessee SCORE—COVID-19 Impact Memo 10: Assessment and Accountability](#)
- [Understanding America Study](#)

Accountability in Spring 2021

- [Bellwether Education—Making Next Year Count: Equity in School Accountability](#)
- [Data Quality Campaign—COVID-19 Recovery Requires Education Data](#)
- [Data Quality Campaign—Measuring Growth in 2021: What State Leaders Need to Know](#)
- [Fordham Institute—Bridging the Covid Divide](#)
- [Fordham Institute—Testimony Given Before the Ohio Senate Education Committee](#)
- [NWEA—Beyond Back to School: Equity & Innovation in the COVID-19 Era](#)
- [Seek Common Ground—Families and Communities Key to Assessment and Accountability Decisions During the Pandemic School Year](#)
- [Tennessee SCORE—COVID-19 Impact Memo 10: Assessment and Accountability](#)

Rethinking Assessments

- [American Academy of Political and Social Science—What Use Is Educational Assessment?](#)
- [Aurora Institute—Education Policy Issues for the COVID-19 Era](#)
- [Center for Assessment—The Challenges and Opportunities of Balanced Systems of Assessment](#)
- [Chiefs for Change—How Should Education Leaders Prepare for Reentry and Beyond?](#)
- [Education Reform Now—Trends in Innovative Assessment Pilots Reveal Opportunities for Actionable Data as Well as Equity Concerns](#)
- [EdTrust—Equity-Driven Approaches to Measure Student Learning](#)
- [Educational Measurement—Using Learning and Motivation Theories to Coherently Link Formative Assessment, Grading Practices, and Large-Scale Assessment](#)
- [Educational Measurement—Educational Assessment: What to Watch in a Rapidly Changing World](#)
- [FutureEd—The Big Test: The Future of Statewide Standardized Assessments](#)
- [FutureEd—Tough Test: The Nation’s Troubled Early Learning Assessment Landscape](#)
- [Knowledge Works—How the Next President Can Reshape K-12 Education](#)
- [Learning Heroes—Parents as Change Agents](#)
- [Migration Policy Institute—Native Language Assessments for K-12 English Learners](#)
- [National Association of State Boards of Education—How a Crisis Can Transform Learning, Teaching, and Assessment](#)
- [National Academy of Education—Comparability of Large-Scale Assessments](#)
- [National Education Association—Learning Beyond COVID-19](#)
- [New Classrooms—The Iceberg Problem](#)
- [Next Generation Science Standards—Assessment Resources](#)

Rethinking Accountability

- [Alliance for Excellent Education—When Equity Is Optional](#)
- [Aurora Institute—Education Policy Issues for the COVID-19 Era](#)
- [Data Quality Campaign—Growth Data: It Matters, and It’s Complicated](#)
- [Fordham Institute—Resetting School Accountability, From the Bottom Up](#)
- [Knowledge Works—How the Next President Can Reshape K-12 Education](#)
- [National Academies—Developing Indicators of Educational Equity](#)
- [New Classrooms—The Iceberg Problem](#)
- [NWEA—How a Crisis Can Transform Learning, Teaching, and Assessment](#)
- [TNTP—Student Experience Toolkit](#)

List of Convening Participants

Below is a list of participants who attended one or more of the convening sessions in October, November, and December of 2020. Being listed as a participant does not mean that an individual or organization endorses all the recommendations contained in this document.

Liz Cohen; 50CAN	James W. Pellegrino; Learning Sciences Research Institute, University of Illinois Chicago	Katie McClarty; Renaissance
AASA The School Superintendents Association	Robert Schaeffer; National Center for Fair & Open Testing (FairTest)	Sara Kerr; Results for America
Alliance for Excellent Education	Meghan Whittaker; National Center for Learning Disabilities	Dr. Melody Schopp; SAS Institute
America Succeeds	Julie Sugarman; National Center on Immigrant Integration Policy, Migration Policy Institute	Edynn Sato, PhD.; Sato Education Consulting LLC
Sonja Santelises; Baltimore City Public Schools	Melissa Lazarin; National Center on Immigrant Integration Policy, Migration Policy Institute	Tennessee State Collaborative on Reforming Education (SCORE)
Collaborative for Student Success	National Education Association	The Education Trust
Dale Chu; DC Strategies	National Parents Union	Keri Rodrigues; The National Parents Union
Education Reform Now	National PTA	UCLA CRESST
Bethany Little; EducationCounsel	Cory Epler; Nebraska Department of Education	Amalia Chamorro; UnidosUS
Evan Stone; Educators for Excellence	Dr. Jeremy Heneger; Nebraska Department of Education	Roxanne Garza; UnidosUS
Randy Bennett; ETS	Amanda Aragon; NewMexicoKidsCAN	Barnett Berry; University of South Carolina (ALL4SC)
Lynn Olson; FutureEd	NWEA	Morgan Polikoff; University of Southern California
Thomas Toch; FutureEd	PIE Network	Michaela Miller, EdD; Washington Office of Superintendent of Public Instruction
Dana Rickman, PhD; Georgia Partnership for Excellence in Education	Christine Pitts	Joanne L. Jensen; WestEd
HCM Strategists		
Laura Jimenez		
KIPP		
Bibb Hubbard; Learning Heroes		

The remaining participants have chosen to remain anonymous.