

THIS IS NOT A TEST, THIS IS AN EMERGENCY: Special Considerations for Assessing and Advancing Equity in School-Year 2020-21

KEY TAKEAWAYS

State policymakers are confronting well-documented intersecting crises – medical, economic, and racial – with especially dire implications for educational equity. State education leaders face a moral urgency to both understand and respond to the challenges students are experiencing and to do so in ways that address burgeoning equity gaps. Education assessment can play a crucial role in identifying these learning and related challenges, allowing policy leaders to direct resources to where the needs are the greatest. It will be incredibly challenging, however, to collect, interpret, and use high-quality state standardized test data this school year. This brief recognizes this conundrum and offers recommendations for state leaders regarding assessment in 2020-2021 including:

1. **Separate assessment from accountability.** There are serious threats to producing valid data to support accountability decisions this year. Changes to accountability systems will require a waiver to state ESSA plans and possibly change in state law or regulation.
2. **Plan now!** State standardized tests operate on a long planning and quality-control-cycle, therefore, the time to act is now. Planning must begin almost immediately. Waiting until January or February will be too late to adapt testing systems to best understand and act on opportunity gaps and learning progress. Plans should also account for various contingencies, particularly due to remote and hybrid schooling, and allow for adjustment as conditions change.
3. **Collect Opportunity-to-Learn (OTL) information.** States must design a system for collecting data to document and understand students' access to the resources, tools, and experiences they need to learn for two main reasons:
 - a. **State summative assessments do not provide information necessary** for policymakers to understand students' learning context. Indicators such as whether students have access to devices and broadband, the type (e.g., synchronous/asynchronous) and amount of instruction available, and the level at which students are engaged are essential for shining a light on inequitable opportunities and directing additional resources where they are most urgently needed.
 - b. **State summative assessment results cannot be interpreted** this year without a better understanding of the circumstances and opportunities facing students. Did a student or group of students get a question wrong because they did not know the concepts tested, were not taught the concepts as part of planned instruction, or did not have the technology to perform their best on the test? Without OTL data this year, it will not be possible to accurately interpret test score data.

State summative assessments do not provide information necessary for policymakers to understand students' learning context. Indicators such as whether students have access to devices and broadband, the type (e.g., synchronous/asynchronous) and amount of instruction available, and the level at which students are engaged are essential for shining a light on inequitable opportunities and directing additional resources where they are most urgently needed.

INTRODUCTION

Providing quality public education is among the most solemn and consequential of all state responsibilities. Regrettably, the education enterprise faces grave threats this year, requiring a lot of attention to the fundamentals of safety, stability and supportive relationships for students, who are experiencing unprecedented disruptions and stressors. State education leaders have responsibility to document the extent of the challenges and their influences on students, educators and schools. The state standardized tests have been an important tool for gathering data on achievement in English language arts, math and science for the last 20-30 years. We support state testing to monitor equity and other long-term educational trends. Unfortunately, there are many threats to defensible (i.e. valid) interpretations of test scores *this year*. State leaders are facing seemingly intractable choices this year, none of which is close to perfect. Unfortunately, state leaders are getting advice from the extremes —“test at all costs” versus “do not test under any circumstances”—and neither extreme is helpful to state leaders who are trying to do what’s best for students while dealing with a myriad of uncertainties. Therefore, this document attempts to present a balanced approach for state education and policy leaders regarding state assessment and OTL data collection efforts in 2020-2021.

The U.S. Department of Education’s September [letter to chief state school officers](#) indicated that the Department was not likely to grant a blanket waiver for Spring 2021 state assessments. The Secretary’s letter allows states to consider alternative data collection approaches in cases when it is either not possible to administer the regular state assessment or when it is likely the test cannot produce trustworthy results. States cannot and should not wait to begin their planning to determine if or how to assess this year because state assessment planning and implementation operates on a long time cycle. Policy leaders must communicate with educators and families clearly and honestly about the challenges they are confronting and to share the steps they are taking to support educators, students and families right now.

The Secretary’s letter allows states to consider alternative data collection approaches in cases when it is either not possible to administer the regular state assessment or when it is likely the test cannot produce trustworthy results.

THREATS TO VALID USES OF STATE TEST SCORES IN 2021

Policy leaders need trustworthy data to maintain moral and legal commitments to all students and especially to allocate resources to close equity gaps. State assessments have served an important monitoring function over the past 20 years to help shine a light on differential learning outcomes for specific groups of students. Therefore, many leaders are feeling pressure and responsibility to continue relying on state assessments in spring 2021.¹ But what if the assessment data this year are not capable of supporting the important decisions facing educational leaders? Some say, “It is better to have low-quality data than none at all.” But that’s not true. There is a long-history of students, particularly low-income students and students of color, being over-identified for special education, subject to over-remediation, denied grade promotion, and kept out of advanced classes based on tests scores that have not been validated for such uses. In the current context, acting on weak data carries considerable costs and risk, such as misrepresenting actual achievement, encouraging inordinate focus on remediation rather than access to grade-level content, and exacerbating stress and anxiety being experienced by all students and educators, but especially by those furthest from opportunity.

Policy leaders need trustworthy data to maintain moral and legal commitments to all students and especially to allocate resources to close equity gaps.

¹ Note: We refer only to state summative assessments in this document and not interim or other assessments generally under the control of local school districts.

As tempting as it might be in what seems like a data desert, grasping at invalid test scores may be like lunging toward a pool of water in a mirage. It may intuitively feel like the right thing to do, but due to this year's circumstances, such actions may lead to serious unintended consequences, especially when time, attention, and other resources are scarce.

As tempting as it might be in what seems like a data desert, grasping at invalid test scores may be like lunging toward a pool of water in a mirage.

Valid interpretations of test scores in the 2020-2021 school year are subject to many more threats than in *any* previous year in the modern testing era. Validity is always evaluated in the context of use. State test scores are used to support accountability determinations, but also serve important monitoring and evaluation uses. There are a [range of schooling scenarios](#) (in-person, hybrid, remote) currently employed and unless conditions change dramatically between now and next spring, we believe federal school accountability requirements will be either waived or modified because the threats to meaningfully interpreting the scores are so substantial that traditional accountability uses will be indefensible. Thus, our discussion here focuses on other uses of test score results to understand student achievement at individual and aggregate levels.

There is no question the coronavirus pandemic and the sudden shift to remote schooling last March has had a dramatic impact on students' opportunities to learn. Worse, there is considerable [evidence](#) that as many as one-third of students, many of whom already experienced significantly fewer educational opportunities, were essentially disengaged from school all last spring. With a significant proportion of students starting this school year in remote or hybrid situations, we expect similar challenges to learning even as we acknowledge that educators and leaders have considerably improved the quality of online learning compared to last spring. Some might suggest these disruptions are the reason *for* giving state assessments this year. When students lack opportunities to learn the content on the test, however, users risk drawing invalid conclusions about the quality of school programs or the efforts of teachers. While trustworthy student achievement data would certainly be helpful in understanding the scope of the crisis, using that information as a basis to support claims or make attributions about the effectiveness of educators, programs, or schools this year is irresponsible.

The more substantial threats to accurate interpretations of test scores are tied to the considerable likelihood that many or most students will be unable to test in schools this year. Some assessment companies have made considerable progress in administering tests remotely. However, none of these companies have produced evidence that scores from remote- and in-school -administered tests meet technical comparability requirements. In other words, just because a test can be delivered to a student remotely does not mean educators and leaders can treat the scores interchangeably with tests administered in traditional in-school settings.

...just because a test can be delivered to a student remotely does not mean educators and leaders can treat the scores interchangeably with tests administered in traditional in-school settings.

Security issues are typically the first concern regarding remotely-administered tests because the level of remote proctoring necessary to prevent cheating for an examination taken outside of school would require a level of intrusiveness that is not likely to be acceptable and may not be legal in many states. Therefore, security is a considerable threat to score meaning for remotely administered tests. Even if security concerns can be overcome, there are many other threats to the interpretability of scores from remotely administered tests. Differential access to the internet, devices, and quiet spaces for instruction and assessment threaten the comparability of scores from remote and in-person tests. Even relatively stable internet access is no guarantee that testing will not be interrupted; all of us have experienced failed Zoom connections during an important meeting. Additionally, we must consider how differences in test administration conditions will affect students. For example, one student may be able to test in a quiet location with little disturbance (e.g., a private bedroom), whereas another student must complete the test in a location filled with distractions (e.g., at the kitchen table with multiple siblings). This variability in environment influences focus and motivation differently compared to student testing in traditional classroom settings.

Finally, most current remote administration platforms do not offer the full range of test accommodations characteristic of standard administration platforms for English learners and students with disabilities. Further, accommodations requiring the participation of special educators, translators, or other specialized personnel may not be available at all or have to rely on an untrained adult (e.g., parent) to help. Beyond the obvious comparability threats when some students receive appropriate accommodations and others do not, there are potential legal issues under the Individuals with Disabilities Education Act (IDEA) and Title III of the Every Student Succeeds Act (ESSA) if students do not receive necessary accommodations.

There are likely many more threats to accurate interpretations than these examples. We have serious doubts whether testing under these conditions can produce an accurate picture of what students know. Taken together, these challenges make it doubtful that test scores from remote and in-person administrations can be combined such that valid interpretations are supported. Being able to aggregate and compare results across schools, districts, and subgroups is the cornerstone of current school accountability systems, but it is also critical for accurate monitoring of score trends.

We have serious doubts whether testing under these conditions can produce an accurate picture of what students know.

WHAT SHOULD EDUCATION POLICY LEADERS DO?

Opportunity-to-Learn Indicators

The learning conditions last spring and into this year warrant an expanded indicator system. Policymakers should focus on designing and implementing opportunity-to-learn (OTL) data collection efforts. There is a long tradition of collecting OTL data as part of surveys of learning conditions as well as embedded in large-scale testing programs. These data have been critical in helping to understand varying levels of educational inputs (e.g., access to curricular and instructional materials, safe and healthy conditions for learning), but have also provided valuable information for interpreting test results. Both functions for OTL data are necessary in any year, but even more so this year.

Validly interpreting the result of the 2021 state assessment, even if administered somewhat normally in schools, will still be challenging ([Keng, Boyer, & Marion, 2020](#)). Drawing credible inferences about student and school performance if a critical mass of students test remotely will be more tenuous. OTL data can contextualize test results to help test users better understand what test scores mean when there is more than typical uncertainty. Additionally, the richness of the OTL data can supplement the information gained from test scores so educational leaders have a more complete picture of the educational system than from test scores alone.

What if, for any number of reasons, states do not test this year? High-quality OTL data across multiple levels of the educational system can provide rich information, so state leaders can direct resources to places where students do not have meaningful and equitable learning opportunities. Therefore, state leaders should prepare a robust OTL data collection system whether statewide standardized tests happen this year or not. Thinking long-term, states should be collecting OTL data regularly, so this could be a good example of not letting a crisis go to waste and design a longer-term strategy for regular OTL data collection and use.

Most previous OTL data collection efforts focused on student- and teacher- level variables. The current crises have brought systematic and structural inequities to the forefront; therefore, state leaders should collect data at multiple levels of the system—such as states, districts, teachers, and students—and consider how these indicators interact with race and class. For example, indicators at the district and even state level would include such factors as internet and device availability because individual schools should not be solely responsible for providing access to these essential materials. Additionally, students, teachers, and parents, via surveys, can provide [key insights on the conditions of learning](#) in remote settings. Critical indicators at the school and classroom level would include records of “attendance,” including in online learning, and

engagement with grade-level content judged by completion of assignments or some other approach. State leaders should solicit information from schools about the types of instructional approaches (synchronous, asynchronous) and how much time students are afforded to interact with each other and with caring adults.

It is beyond the scope of this document to provide detailed data collection plans. While it may be challenging for some states to design and establish new OTL data collection systems, this is not a new field so states can work with technical advisors and other partners to prioritize a few key measures to start with this year, while planning for a more robust OTL data collection and reporting system for the future, along with training so key users understand how to accurately interpret and use the results. Regardless of state decisions on assessment administration, each state can and should collect and report some data on opportunity to learn this year and they should build OTL indicators into their long-term reporting systems.²

Large-scale testing

We support large-scale testing, but are concerned with the many threats to justifiable interpretations to test scores this school year. Therefore, we offer the following suggestions regarding state standardized testing for this year:

1. If conditions do not markedly improve, and especially if they deteriorate, state leaders must ask if administering statewide standardized tests is warranted. They should ask their assessment specialists to articulate the specific purposes and uses for testing in spring 2021 and then require an honest appraisal about whether the threats to valid interpretations have been addressed such that the test can fulfill the intended purposes and uses. Negotiating with the U.S. Department of Education on these issues will be required, but federalism requires a robust dialogue and due regard for the considered judgments of state education and political leaders, especially in responding to multiple emergencies with highly constrained resources.
2. We recommend using information about prioritized standards³ and curriculum to shorten and adjust test blueprints, whether tests are administered in person or remotely. For example, by eliminating the requirement for producing “subscores” (i.e., scores assigned to specific domains within the topic, such as “numbers and operations” on a 5th grade math test) in addition to overall achievement, tests could be considerably shorter. Most technical advisors agree that such subscores add very little value even in the best of times, so this could be a good time to eliminate them for good.
3. If almost all schools and students in a state remain in school for in-person instruction throughout this academic year, then using the existing state assessment may be justified taking into account recommendation #1. Wherever possible, students should test in school, so testing conditions are most similar to past approaches. This must be thoughtfully balanced with the health and safety concerns for students, families, and staff.
4. If a state moves forward with remote administration, it is imperative to decouple the test from any accountability uses and to eschew comparability claims. There have been no studies documenting the comparability of remote and in-person test scores. Therefore, if the state administers remote assessments, they must conduct comparability studies and issue clear guidance for how to properly interpret and use these results and how to avoid misinterpretations and use. Further, state leaders will have to be ready to address and deal with lower than typical rates of participation.

If conditions do not markedly improve, and especially if they deteriorate, state leaders must ask if administering statewide standardized tests is warranted.

² See: Floden, R. E. (2002). The Measurement of Opportunity to Learn. in Porter, A. C. & Gamoran National Research Council. *Methodological Advances in Cross-National Surveys of Educational Achievement*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/10322>.

³ See: Student Achievement Partners' [2020-2021 Prioritized Instructional Content in English Language Arts/Literacy and Mathematics](#).

5. If the goal is to take the pulse of the educational system, there is no need to test every student on all of the content. States could employ numerous sampling strategies such as testing a representative set of students in every grade or testing only a few grades like what is done with state science tests; these strategies can produce group data to track equity gaps but not produce scores for every individual student. Sampling approaches would have the added benefits of reducing costs and clearly signaling that this year is different and ameliorating accountability fears. As we noted, any such changes would require a waiver from ESSA testing requirements and considerable planning time, so states must undertake such planning now.

If the goal is to take the pulse of the educational system, there is no need to test every student on all of the content.

In putting forward a testing plan, all of these options must be balanced against the potential misuse of tests—e.g., labeling students, excessive test prep, and mind-numbing remediation activities—and potentially lost opportunities for making sure every student is seen, supported, and engaged in learning this year. Regardless of the scenario, state leaders must work with their assessment providers and technical advisors early and often to address the logistical, technical, and contractual ramifications of their policy. Waiting until spring 2021 to make changes will foreclose many meaningful options for adaptation.

Waiting until spring 2021 to make changes will foreclose many meaningful options for adaptation.

What about accountability?

While the U.S. Department's stance on assessment sounded firm, the door to accountability waivers appears to be at least partially open. We believe that school, teacher, and student accountability, as we know it, should be suspended for 2020-2021. However, we strongly support a focus on reporting key indicators for which the state, district, and communities—as well as schools—are responsible, such as OTL indicators that can shine a light on which students have—and have not—been afforded access to meaningful learning opportunities.

FOCUS ON WHAT MATTERS MOST

In this historic year, states are forced to prioritize among competing uses of limited time, money, and attention. We urge an emphasis on wellness of students and staff first, and then addressing assessment issues in light of the need for data to understand the scope of the crisis. We pointed out that OTL measures may provide data to help decide where needs are greatest without the potential negative consequences associated with pushing forward state standardized tests.

We fully expect and support state standardized tests remaining a prominent feature of public education in the future. No matter what, decisions regarding testing should be communicated early and clearly to educators, students and families. Additionally, state policymakers can and should communicate the ongoing commitment to equity and the measurement of learning to avoid strengthening the hand of testing opponents.

Some advocates argue that unless state leaders continue their assessment programs this year, they risk never being able to return to the current system of assessment and accountability. But recklessly pushing ahead with statewide testing could be a case of winning the battle only to lose the war. When many students and caregivers are experiencing adversity and intense stress, being forced to change routines and commit days to state testing could inspire new levels of resistance that would re-energize and expand the opt-out movement of a few years ago.

Recklessly pushing ahead with statewide testing could be a case of winning the battle only to lose the war. When many students and caregivers are experiencing adversity and intense stress, being forced to change routines and commit days to state testing could inspire new levels of resistance that would re-energize and expand the opt-out movement of a few years ago.

State and district educational leaders must always focus on creating excellent and equitable learning opportunities for all children. The conditions this year require these leaders to focus even more intensely on the core mission of addressing pervasive and unconscionable opportunity gaps while supporting local educators and students with resources necessary to engage in meaningful learning. Wise use of robust OTL measures along with state assessment results, if they can be collected, can provide information educational leaders need to support their students this year. Policymakers need to begin planning now and, regardless of decisions on state summative assessments, all states need to move immediately on collecting and reporting OTL data.

Suggested citation:

Marion, S. F., Gonzales, D., Wiener, R., & Peltzman, A. (2020). *This is Not a Test, This is an Emergency: Special Considerations for Assessing and Advancing Equity in School-Year 2020-21*. Center for Assessment (www.nciea.org) and The Aspen Institute (www.aspeninstitute.org/education)



ABOUT THE CENTER FOR ASSESSMENT:

The National Center for the Improvement of Educational Assessment, Inc. (The Center for Assessment) is a not-for-profit, non-partisan corporation comprised of assessment and accountability experts with a keen awareness of policy and practical considerations. Center professionals provide customized support to states and districts in designing, implementing, and improving fair, effective, and defensible assessment and accountability programs. The Center leads and/or serves on technical advisory committees for more than 35 states. For additional information about the Center for Assessment, please visit our [website](#).



THE ASPEN INSTITUTE

EDUCATION & SOCIETY PROGRAM

ABOUT THE ASPEN INSTITUTE:

The Aspen Institute is a global nonprofit organization committed to realizing a free, just, and equitable society. The Aspen Institute Education & Society Program improves public education by inspiring, informing, and influencing education leaders to take action across policy and practice, with an emphasis on achieving equity for students of color and children from low-income backgrounds. Aspen Education supports leaders at all levels – from urban superintendents and their teams, to state chiefs and their cabinets, to elected officials and their staffers, to education support organizations, associations, nonprofits, and philanthropy. To learn more about the Education & Society Program, visit our [website](#).

