



RATIONALE FOR STATES CONSIDERING ASSESSMENT FLEXIBILITY

Scott Marion, Chris Domaleski, and Juan D'Brot
National Center for the Improvement of Education Assessment¹

March 16, 2021

INTRODUCTION

This document is meant to support states who find they need additional flexibility beyond the three options listed in the U.S. Department of Education's (ED) [February 22, 2021 letter to state chiefs](#). This rationale accompanies the [Assessment Flexibility Guidance](#) posted by the Center for Assessment on March 9, 2021. State leaders should select and modify aspects of this document as they see fit to help make their case for their waiver request.

The Center for Assessment believes that decisions about statewide summative testing during the spring of 2021 should be informed by local contexts and state policy priorities. Some state leaders may conclude it is appropriate and feasible to move forward with testing, while others will determine conditions suggest another alternative as the best course of action. We believe state leaders should be afforded flexibility when it is determined that the risks outweigh the potential benefits.

PURPOSES AND USES OF STATEWIDE SUMMATIVE ASSESSMENT

One of the truisms in testing is that tests are designed and validated for a very limited and specific set of uses. State assessments required under ESSA must be used as the foundation of ESSA-required school accountability systems. With the immense challenges to viable accountability in 2021, some have questioned the purpose of status quo statewide summative assessment.

ESSA also mandates testing for public reporting, particularly in terms of reporting disaggregated results for identified student groups, and we agree that public reporting is another important function for statewide testing. Therefore, many have argued that even if accountability is waived or greatly curtailed, states should administer their statewide assessments to fulfill this purpose, especially this year. We support this use-case in normal school years, but this year is anything but normal.

¹ We are grateful to the Council of Chief State School Officers (CCSSO) for their support in producing this document. We are also thankful to Scott Norton and Kirsten Carr of CCSSO for helpful comments on earlier versions. Of course, any shortcomings or errors are our own.

Many advocacy groups and policy leaders are urging that tests should be administered this spring because “the data are critically needed.” There are competing ideas, however, about how the data will be used and whether these uses are technically defensible. There are also conflicting ideas about how tests can be safely and appropriately administered under the COVID circumstances that are likely still to be extant this spring. We expand on these issues below.

TRUSTWORTHY AND USEFUL RESULTS

In order to serve public reporting or accountability purposes, state tests must be administered under standardized conditions. If that is not possible, the ability to draw valid inferences from the results is severely hampered. First, in spite of the admirable push to get all students back in school for in-person learning as soon as possible, it is unlikely that all students are attending in-person schooling during the spring testing windows, which generally takes place in March through May. If students are not in school for learning, states are left with two unsatisfying choices for administering tests this spring: (1) students could be required to come into school buildings on a schedule that supports safe social distancing to complete tests, or (2) tests could be administered remotely.

Testing in School

State and district leaders would face a contradictory position by requiring students to come into school buildings (or other facilities) to take state assessments when students are not in schools for learning. In fact, ED made clear in the February 22nd letter that students should not “be brought into school buildings for the sole purpose of taking a test.” Policymakers are also mindful of the opt-out movement, at its height a few years back. Any testing protocols that called for special cooperation from families would likely prompt a spike in refusals as parents elect to keep their children at home due to health or related concerns.

Remote Administration

Large-scale remote proctoring is without precedent in K-12 state testing. It is a departure from responsible professional practice to implement an unproven approach with so many potential risks. Moreover, the kinds of remote proctoring requirements necessary to ensure secure test administrations would violate most states’ student privacy rules.

Early results from assessment providers and states have shared results from remotely-administered interim assessments last spring and this fall. While the results might be useful on a local level, the results highlighted enough concerns such that no more than a handful of states are considering administering summative tests remotely. Finally, most technical experts doubt that test scores from normal, in-person test administrations and remote test administrations can be combined automatically as if they were equivalent due to the many threats to comparability.

Reduced Content

Most schools, either by design or circumstances, taught a reduced set of content this year compared to previous years. This is true even for the schools that have largely had their students in school most of the year. Organizations such as Student Achievement Partners recommended prioritizing [high-leverage content standards](#) during the 2020-2021 school year to help school personnel focus on what matters most. However, we know of no states that changed their test blueprints to match a reduced set of learning targets. This could cause a systematic mismatch between what students were taught and the content of the test. Equally as concerning, but harder to detect, is the fact that

many teachers will simply have less time to teach the grade level standards, so they likely are making non-systematic decisions about what content to include and exclude. In both cases, this can lead to challenges in interpreting student test scores.

POTENTIAL NEGATIVE CONSEQUENCES

Even in the unlikely event that essentially all students are back in schools early this spring, we question whether states should be required to administer statewide assessments as if everything was normal. Yes, the data may help estimate shortfalls, even disproportionate losses due to COVID, but these data will not be comparable across jurisdictions or to 2019 data. Further, we are concerned that the potential unintended negative consequences present more compelling concerns with state testing this year.

Accountability Concerns

Those who advocate for administering state tests, even with formal accountability consequences removed, must recognize that results would still be publicly reported. This is accountability to most school personnel. Test results in spring 2021 will reflect a mixture of contributing causes including, but not limited to:

- Support from state and district offices,
- Teachers' efforts and skills,
- Remote learning platforms,
- Wi-Fi access and device availability,
- Extended student absences, and
- Schooling supports from parents,

In spite of all these confounding factors, any negative perceptions associated with test scores will fall on the shoulders of school personnel alone.

Interpretation

The challenges associated with appropriately interpreting test results this year shifts the equation. Most users will interpret the test scores through an "in-person lens" because that is all they know. Unfortunately, without a comprehensive understanding of the factors influencing test results this year—many discussed herein—the risk of misinterpretation likely outweighs the chances of appropriate interpretation and use. This topic has been discussed at length in technical advisory committee (TAC) meetings and at other technical meetings this winter, particularly the Council of Chief State School Officers' Technical Issues in Large Scale Assessment State Collaborative. Several experts at the Center for Assessment (see many of the resources [here](#)) and elsewhere (e.g., see Andrew Ho's recent [paper](#)) have written about the challenges of validly interpreting test scores this year, both at individual and aggregate levels.

Lost Instructional Time

Educators, parents, and students likely will resent giving up a week or more of precious instructional time to participate in a state test, especially if students have only returned to in-school learning shortly before testing. In our interactions with district leaders around the country, many emphasized their need for every bit of instructional time they could have. Besides the potentially demoralizing consequences of having students take tests for which they might only be partially prepared, giving up any precious instructional time will likely make matters worse.

MEETING FEDERAL REQUIREMENTS

The Every Student Succeeds Act requires “comparable” large-scale test data to document the levels of student learning across a state. As noted above, we question the degree to which these data will be as comparable as intended. However, we would, in this limited case, support review and aggregation of already administered interim assessment data from established commercial providers or those administered as part of state assessment systems. The studies from the major interim assessments provide the type of information, along with opportunity-to-learn (OTL) information, useful for guiding policy decisions. This does not mean we support replacing state tests with interim assessments post pandemic or adding new tests, but it makes sense to capitalize on the tests districts are already administering this year. This is because, as we have illustrated throughout this memo, statewide tests will be, at best, approximate indicators of student and school performance.

Shorter tests and other changes to test design

Finally, some are calling for alternative designs for state summative tests to provide a more representative picture of student learning than could come from aggregating data from multiple interim assessment systems. The only practical option at this point—using short versions of the state summative assessments—is a practical option for this year only if the state and its assessment provider have already engaged in substantial redesign work and *only if essentially all students can test in schools in somewhat normal conditions*.

Planning Instruction and Informing Parents

The U.S. Department of Education has expressed concern about reporting student performance information to parents. Knowing that the percent of students who are proficient on the state test has declined since 2019 does not provide any substantive information about how to plan instructional interventions nor does it tell parents what their students have mastered or what additional help is needed. Assessments embedded in high-quality curriculum already being used are the best way for teachers to plan instruction and to share information with parents. District and school leaders without existing assessments could identify key assignments reflective of grade-level expectations and could share exemplars of student work to help parents understand how their students are performing relative to grade-level standards. While these alternatives may be imperfect, they are better suited to provide instructionally useful feedback and less invasive to implement.

Accessibility and Accommodations

The Individuals with Disabilities Act (IDEA) and ESSA have justifiably strong requirements for assessing students with significant cognitive disabilities through the use of an alternate assessment on alternate achievement standards and for providing the full range of allowable accommodations for students with disabilities and for English learners. Most commercial interim assessments do not provide accommodations for low-incidence disabilities (e.g., blindness) and none of the major providers offer alternate assessments. Therefore, states must either, through the state assessment system or other means, provide alternate assessment and/or appropriate accommodations for all students who need them.

FINAL THOUGHTS

Documenting “COVID-19 learning loss” is one of the main arguments for testing this year, but such testing advocates rarely specify the necessary degree of precision or the level of the system where the information is needed. For instance, an end-of-year test rarely provides enough information at the student level to guide instruction. Would the use be focused at the school and/or student group level? Who would take action (district leaders, state policymakers, or others)? Would the information garnered from an end-of-year state test provide enough added precision beyond what could be gathered from opportunity-to-learn (OTL) data to shed light on how policymakers should address pandemic-related inequities²? If the learning losses are anywhere near what we suspect and are being revealed through recent analyses, many justifiably posit that we already have enough information to know we need major structural interventions (e.g., funding a month or more of summer school for all students) to address these issues.

² See [Marion, Gonzales, Weiner, and Peltzman, 2020](#) and [Marion, 2020](#).