Restart & Recovery: Assessment Considerations for Fall 2020

July 2020

Introduction

Many district and school leaders are wondering how to assess students at the beginning of the 2020-2021 school year. Assessment and learning specialists wrestle with many of the same questions. There is much talk of "diagnostic" assessments, but little clarity regarding how to use these assessments or how best to design them. The term *diagnosis* implies a degree of specificity that many instructional systems do not support, certainly not at the individual level. This can lead to the use of potentially misleading labels. Instead the focus should be on specific articulation of use-cases and theories of action as the foundation for assessment design and implementation.

This paper discusses assessment considerations for fall 2020, focusing on specific use-cases associated with key education stakeholders. The focus is generally on two levels of the educational system for different purposes, both of which are very important. Assessment directors will be interested primarily in how assessments can best support teachers and other front-line educators, to ensure that students are learning grade-level content. However, district and state leaders have a critical responsibility to monitor student performance in terms of the equality of opportunity-to-learn (OTL) and to address whatever inequities they find. Thus, in addition to assessing grade-level content competency, these tests could play a role in helping state and local leaders to identify concerns about OTL equity.

This paper is based on the guiding principle that all students are capable of grade-level learning, regardless of their starting points; all students are capable of progressing to the next grade level this fall and mastering that content. The goal is to avoid over-remediation, focusing on below grade-level work only insofar as it is needed to help each student understand the grade-level work to be done.

It starts by outlining three re-entry scenarios, followed by an overview of several cross-cutting issues: timing (i.e., when to assess and make decisions); organizational and resource issues; and content area and grade-span differences. It concludes by discussing the implications for assessment design early in the 2020-2021 school year, with tables showing critical issues that state leaders are likely to confront, along with use-case scenarios for evaluating the implementation of different types of assessments. Finally, it offers 10 recommendations, which should be adapted to whichever re-opening scenario the school opts to implement.

Re-Entry Scenarios

Given the unresolved nature of the COVID-19 pandemic, any discussions of fall 2020 assessments must remain open-ended. It is unlikely that school will return to normal, with students and teachers interacting in classroom settings without disruption throughout the 2020-

2021 school year. Any approach to instruction and school organization—and therefore assessment—should consider at least three re-entry scenarios:

1. <u>Fully in-place</u>. School resumes normally in the fall, and the school year unfolds normally.
2. <u>Blended, or partially in-place</u>. School resumes in-person classes, but social distancing necessitates some sort of alternative scheduling. For example, to maintain proper spacing in classrooms, half the students attend class in person while the other half attend remotely, alternating weeks.
3. <u>Remote</u>. A final possibility is a return to full-time remote schooling, with school buildings remaining closed, as was the case for most U.S. schools in spring 2020.

Possible permutations of these scenarios might include earlier or later re-openings and cyclical returns to remote schooling. Some states are preparing for early openings of the school year (some as early as July) in anticipation of possible disruptions to schooling during 2020-2021. These different scenarios, along with other variations, have implications for the instructional and organizational strategies employed to fulfill the many functions of schooling. These scenarios undoubtedly will affect assessment strategies, as well.

There is not enough time to make wholesale changes to local assessment systems and significantly increase the assessment literacy of teachers and leaders on a large scale. Educators need to have modest expectations about what can and cannot be accomplished in this timeframe, recognizing that school systems are dealing with many competing demands. To make the biggest impact under such extreme circumstances in this limited time frame, district and school leaders are encouraged to focus on accomplishing a few, high-leverage goals.

Virtually the entire education enterprise has been racing to address the implications of this unprecedented and rapidly evolving situation. School leaders have had to deal with the immediate and pressing task of continuing to educate students during the spring 2020 lockdown, while confronting the longer-term ramifications for assessment and accountability. When the crisis first exploded in early March, many stakeholders assumed that children would be able to return normally to school next fall. By April, however, it was growing clear that this may not be the case. Educational leaders, service providers, and other stakeholders have since shifted their focus to reopening school next fall, planning for multiple contingencies for educating students in each of the three general scenarios listed above. With determination and adaptability, the nation's educational leaders are wrestling with numerous cascading issues related to health and safety and student learning.

How and when can leaders use assessment evidence to inform school organizational strategies and resource issues to maximize educational opportunities next fall? A typical school improvement cycle is based on state assessment results from the prior spring, along with a range of other data collected throughout the preceding school year, to support planning efforts over the summer. Waiting until the new school year commences to collect and interpret assessment data allows leaders no time to plan and implement major structural or scheduling changes, such as providing additional mathematics learning time during flexible periods for students or allowing additional common planning time for teachers. To make such organizational changes possible, school and district leaders need the kind of actionable data that enables them to begin planning now.

In the absence of such data, school and district leaders may need to adopt a hypothesis-testing mindset and commence with planning based on certain assumptions and currently available data, as noted in Recommendation 8. Assessment information collected upon school re-opening could then be used to confirm or refine their hypotheses. For example, if they have current attendance and engagement data, school and district leaders could predict which students are likely to have fallen further behind in math, then arrange schedules to provide targeted additional time for those students in the fall. Once fall assessment results are available, leaders could use the findings to triangulate such decisions or allow for other adjustments to the students' schedules.

Even without attendance or engagement data, however, it is reasonable to predict that most students will begin the 2020-2021 school year further behind than their peers from prior years. Many educators are concerned that achievement gaps will be exacerbated. In addition to the well documented racial and socioeconomic achievement gaps, we could see test-score gaps between households where the parents/guardians are working full time and households where at least one

Assessment Considerations for Fall 2020 7/14/20

adult has the time, skill, and temperament to support their children's educational needs. This is especially true for younger students.

In formulating organizational and instructional plans, it will be important for district and school leaders to base their hypotheses on local conditions, by connecting with the people most directly affected by the pandemic: the students themselves, families/guardians, and teachers. Informed by teacher-based data and other information collected during spring 2020, for example, school leaders might hypothesize that students entering fifth grade mathematics will have a weaker-than-expected grasp of word problems that involve fractions. Using this hypothesis, they might organize instruction in the fall so that students can engage with such content. It also will be important to have contingency plans that allow for flexible adjustment, depending on the extent to which a given working hypothesis is confirmed by the results of fall assessments conducted soon after students return to school. Such data, however, is only one source of evidence and must be considered in context, along with other relevant contextual information.

Organizational and Resource Issues

The challenges facing all educators this fall center on instruction and school organization. Stakeholders may recommend assessments for a variety of purposes, but school leaders' decisions will need to focus on what type of assessment is necessary and feasible in the context of the school's organizational and instructional systems. For example, does the district/school have the resources and organizational capacity to provide intensive, student-level remediation? Does the state have such resources? If not, designing and administering assessments to identify student-level learning loss may not be offset by the time such testing would take away from actual, grade-level instruction. No doubt, many stakeholders are interested in monitoring educational trends, but if test results cannot result in meaningful action, the time spent on assessments should be minimized.

Educator Practices

Teachers at all grades and across all content areas will need time to plan for and implement pre-assessment activities into their first instructional units for this fall—which is good instructional practice in any year but particularly important in 2020-2021. Teachers may need to scaffold students who have pandemic-related learning loss into grade-level learning. This is does not necessarily mean schools should "drill" incoming students on prior-year standards. In English language arts (ELA), for example, it may be more helpful to assess students' vocabulary and contextual knowledge skills for an upcoming text, rather than assessing specific aspects of the prior year's standards, such as being able to identify the author's purpose. For incoming high school science students, it may be more important to assess their understanding of sixth-grade

proportional reasoning than their mastery of certain concepts in biology that are needed for learning chemistry.

Educators could weave such pre-assessment into the regular classroom assessment systems. Doing so would tie assessments to the skills and knowledge built into the local curriculum. It is important to note that a general-survey assessment would not provide the level of information that teachers need to make specific, just-in-time instructional decisions. Moreover, commercial assessment products with an overreliance on selected-response items could lead teachers to feel obliged to remediate each student on every missed question or aligned standard. This kind of instruction is unnecessary in all content areas, and the implications for potential OTL inequities is concerning.

Expecting teachers to plan pre-assessments and instructional units at the individual-student level would be onerous and unfeasible, and it could introduce unacceptable variability in terms of assessment quality. Instead, districts might consider obtaining existing guidance from providers of high-quality instructional materials or facilitating opportunities for school- and grade-level teams to collaborate prior to fall instruction to develop their assessments. Granted, such arrangements would not constitute formal assessments, but they could be designed to align with specific, precursor knowledge and skills deemed necessary for success in the current unit(s).

It is useful for district leaders to first ask their curriculum providers if these assessments already exist. Finally, in order to meet the unique challenges of remote learning in spring 2020, some students may have improved skills such as self-regulation, persistence, and creativity. Conversely, some students might have been frustrated in their development of these skills, due to lack of support or various structural issues. Teachers should be attuned to how to assess and capitalize on or shore up\ these skills, which are critical for resiliency.

<u>Content Areas and Grade Spans</u>

Many of the discussions about fall 2020 assessments may lack the nuance educators have come to associate with the expectations for different content areas and grade levels. Determining what students have learned in a given domain such as mathematics, for which certain material builds on the acquisition of prior-level knowledge and skills (even if there is no consensus on the ideal order for those levels), is more critical than domain-based assessments may be in science or social studies, where learning can shift from topic to topic and sequentially-acquired knowledge is less critical.

However, this generalization does not apply to all grade levels. In high school, for example, there may be a general understanding that students need to have learned key physical-science concepts

Assessment Considerations for Fall 2020 7/14/20

before transitioning into chemistry or should have a solid understanding of geography before taking a world history class.

On the other hand, ELA poses an interesting challenge when planning assessments for the fall. Most districts likely have more reading and language-use assessments than in any other content area. In this case, rather than administering reading assessments just because they are readily available, schools would be better off selecting only those tests whose results will serve an intended use.

Providers of ELA assessments should avoid using a standard, domain-based sampling approach to test development and instead engage in more purposeful design approaches, with special focus on content demands and curricular alignment. Except for learning how to read and developing early literacy skills, much of reading is tied to the knowledge and vocabulary that students acquire through text. Therefore, once students know how to read fluently, there is no generic "most important skill" students should have learned during the previous grade; identifying an appropriate scope for each higher-level ELA assessment will depend entirely upon the curricula a given district is using.[1]

Use-cases

Table 1 outlines the most likely use-cases for each of the main classes of assessment users, organized by the granularity of information needed, from the most general to the most specific. For instance, teachers need the finest-grain information for all their use-cases, whereas state and district leaders generally need only coarse-grain information. The table also indicates the appropriate level of reporting (e.g., student- or school-level) to meet the intended uses. Timing and referent (norm- or criterion-referenced, relative to presence or absence of performance expectations) requirements are also shown. Also included are potential interpretations which are either justifiable or likely to be inappropriate. State, district, and school leaders are encouraged to engage in a use-case exercise for their specific contexts, thinking through the processes and mechanisms to bring the use-case to fruition.

---

[1] We recommend relying on Student Achievement Partners' *2020–21 Priority Instructional Content in English Language Arts/literacy and Mathematics* as guidance for essential knowledge and skills from prior grades.

Table 1: Users and Use-cases	General Assessment Requirements	Appropriate Interpretations	Interpretations to Avoid
State and district leadership • Monitor and communicate quantifiable information regarding equity and access in the state and/or district about the effect of COVID-19 on overall achievement and achievement gaps. • Direct discretionary funding and resources to schools and districts most in need. • Support monitoring and program evaluation needs. • Forecast spring 2021assessment results to anticipate potential accountability issues. • Monitor screening and diagnostic needs for students referred to special services.	• Granularity: Most coarse-grained • Reporting: School-level and higher, by student group • Referent: Norm- or criterion-referenced • Timing: Fall 2020 • Grade/Content Priorities: Focus on K-2 ELA Foundational Skills and then sample other grades and content areas. The assessment is either one that is normally given early in the school year (e.g., commercial interim) or tied to the existing state summative assessment (e.g., custom interim). The assessment is quite short (e.g., one class period per subject area). Assessments can rely on some type of student- and/or item-sampling to minimize testing time while collecting rich, school-level information. Note: it may be challenging to introduce sampling for the first time in fall 2020. Document COVID-related achievement gaps by either norm-referencing (comparing to student performance to prior years) or criterion-referencing (tied to a defined expectation, such as state-defined proficiency).	Make school-level (but not student-level) results public and communicate transparently. Compare student performance with that of similar students in the past (e.g., entering 2020-2021 fifth graders, compared to entering fifth graders in prior years). Use assessment results to evaluate gaps in achievement among various student groups. If the fall assessment is already linked to the summative assessment scale, use aggregate-level results to forecast anticipated changes in spring 2021 assessment results.	Avoid any comparisons not supported by existing linking characteristics of the test. Linking designs implemented for the first time in fall 2020 will be subject to misinterpretations. Consider carefully when the test was administered (e.g., first week back in school) when interpreting results. Avoid making instructional or organizational decisions based on sub-score reporting (domains or strands). These scores are neither reliable enough nor targeted enough to support decisions at the school or student levels. Approach any changes in achievement gaps very carefully. Avoid reporting achievement gaps at granularities finer than the content area. Avoid making long-term predictions regarding student performance.

Table 1: Users and Use-cases	General Assessment Requirements	Appropriate Interpretations	Interpretations to Avoid
Parents/Guardians • Provide information about how their students compare to same-grade peers from prior years.	• Granularity: Coarse- grained • Reporting: Student level • Referent: Norm-referenced • Timing: Fall 2020 Present the general assessment information as described for state and district leaders to parents/guardians and students in user-friendly formats.	Parents/guardians are able to judge how well their students learned prior-grade content, as compared to how well students performed at that school/district/state in prior years.	Consider carefully when the test was administered (e.g., first week back in school) when interpreting the results. Avoid over-interpreting or misinterpreting small changes in results. Avoid making long-term predictions about student performance
Parents/Guardians • Provide information on what supports their students will need to be successful when beginning the 2020-2021 school year.	• Granularity: Relatively fine-grained • Reporting: Student level • Referent: Criterion-referenced tied to curriculum expectations • Timing: Fall 2020 Link specific assessment results to specific supports that parents/guardians can provide.	Parents/guardians receive guidance regarding whether students will need out-of-school support to achieve grade level expectations in 2020-2021, and what kind of support.	Avoid making vague suggestions or recommending supports that parents/guardians cannot provide (or advocate for) on behalf of their student.

Table 1: Users and Use-cases	General Assessment Requirements	Appropriate Interpretations	Interpretations to Avoid
School leadership Plan for the most optimal school organization strategies, including: • Screening for potential intervention needs • creating school schedules to best meet learning needs • marshalling resources to support remediation and other supports for individual and small groups of students	• Granularity: Coarse-grained • Reporting: Student level and student group level • Referent: Norm- or criterion-referenced • Timing: Spring/summer 2020 (planning); fall 2020 (adjustment) Use established tools for screening reading and related challenges Because of the COVID-19 school disruptions, most schools will not have formal assessment information to meet the planning use-case in the necessary time frame. School leaders will need to rely on existing information, such as teacher records of content covered in-place and remotely, student course grades, and attendance records, to support their planning use-cases.	Use screening tools according to published guidance but consider the results in the context of school disruptions. Using attendance and engagement data from spring 2020 for planning and organization, adopt a hypothesis-testing approach that can be adapted if formal fall assessment results and/or teacher reports do not support working hypotheses. If district-level assessments were administered in spring 2020, schools may very cautiously use that information for planning purposes, as described in the use-case.	Avoid over-interpreting and acting on any screening results that are close to published "cut scores." Avoid making decisions that could keep students from learning grade-level content. Avoid making inflexible plans – for example, making it difficult to move a student from a targeted intervention placement after teachers observe that the student is better served in a mainstream setting. Avoid making any firm decisions based on spring 2020 test results.

Table 1: Users and Use-cases	General Assessment Requirements	Appropriate Interpretations	Interpretations to Avoid
School leadership • Organize professional learning activities during 2020-2021.	Granularity: Moderate grain size Reporting: Student level Referent: Criterion-referenced Timing: Fall 2020 Any school-level assessment exclusively used to support this use-case should occupy a limited footprint, be tied to the specific curriculum taught at the school, and be focused on (a) the highest priority prior-grade content and (b) the precursor skills and knowledge for successful engagement with grade-level content in 2020-2021. Leaders elevate the formative assessment practices and classroom evidence gathered by teachers as worthy information on which to base instructional decision making.	Focus professional learning on the priority areas flagged by the assessment. Find ways to celebrate and capitalize on relative strengths identified. Demonstrate when and how teachers can put into practice the findings from the assessments, highlighting just-in-time support for grade-level instruction. Prioritize, support, and provide space for teachers' pre-assessment and formative assessment practices.	Avoid instructional approaches that exclusively address gaps as a prerequisite to teaching grade-level content. Avoid grading students based on pre-assessment and formative assessment results.
School leadership • Support monitoring and program evaluation needs.	Granularity: Course-grained Reporting: Depends on the program context Referent: Flexible Timing: Spring 2020 Support this use-case with an assessment that meets the requirements for state and district leaders. (See above.).	Along with other data, use assessment results to support monitoring and program evaluation theories of action.	Avoid individual-level and/or curriculum-related reporting.

Table 1: Users and Use-cases	General Assessment Requirements	Appropriate Interpretations	Interpretations to Avoid
Teachers • Gain knowledge of the key concepts and skills students need to succeed in the first major instructional unit of the school year. • Develop student profiles on those key skills.	Granularity: Finest grain size Reporting: Student level Referent: Flexible Timing: Fall 2020, or as needed Tie short pre-assessments to the knowledge and skills required for success in the first or first few instructional units of the school year. Avoid conducting a survey of all possible course content for the upcoming school year. Design the assessment to yield useful information for each of the limited, major-curricular skills and concepts specific to content area and grade level. Make these assessments very low stakes tests that can be administered in a range of conditions.	Use student performance on the curriculum-embedded pre-assessment to provide feedback and/or supplemental instruction to students to help them progress to grade-level content. If appropriate, use flexible, short-term groupings based on student profiles.	Pre-assessments should not be graded. Avoid using the results to support extensive remediation or retention initiatives that delay student exposure to grade-level content.

Table 1: Users and Use-cases	General Assessment Requirements	Appropriate Interpretations	Interpretations to Avoid
Teachers • Check for understanding of and strategies for addressing conceptual and skill gaps in the context of grade-level curriculum, to foster student learning.	Granularity: Finest grain size Reporting: Student level Referent: Performance expectations Timing: Just-in-time Employ a range of regular formative assessment practices (daily) tied to the curriculum, to provide feedback to students regarding their learning of the current unit and how to improve learning. Information gathered from the large-scale assessments described above will be of very limited use in this case.	Teachers are able to make initial instructional decisions based on their examination of student work and other information gleaned from (a) the pre-assessment and (b) checks for understanding embedded within units and lessons. Teachers use the results of ongoing formative assessment practices to adjust instructional approaches and to target emerging incorrect and/or partial understandings. The focus is on near-term supports to help students meet the expectations of the current instructional units.	Avoid grading any pre-assessments or formative assessments. Avoid using the results to support extensive remediation or retention initiatives that delay exposure to grade-level content. Avoid using information from large-scale assessments to "over-rule" the curriculum-embedded assessments.

.

<u>School/Teacher Use-Case: An Example</u>

Table 2 offers a more specific examination of assessments designed to support and integrate with the approaches to curriculum and instruction recommended throughout *Considerations for Teaching and Learning*. This examination follows from the use-cases cited above, elaborating on the characteristics of assessments which either support or do not support the recommended curricular and instructional use-case.

The types of assessments needed to support students' successful re-entry in fall 2020 depends critically upon the design of curriculum and instruction. As noted in *2020–21 Priority Instructional Content in English Language Arts/literacy and Mathematics*, it will be key for curriculum and instruction to be focused on moving students forward through grade level content and designed to minimize the need for extensive remediation prior to commencing with grade-level instruction The essential characteristics of this assessment (listed below) are contrasted with those of assessments designed for other use models.

Table 2

Assessment designed to support and be integrated with recommended approach to curriculum and instruction	
Has These Characteristics	**Does Not Have These Characteristics**
Integrated with vision and model of curriculum and instruction that moves students to learning essential grade-level content as quickly as possible.	Uses an assessment model in which all or most of the previous set of content standards and/or curriculum is re-taught prior to going on to grade-level content—i.e., using an extensive, "remediation first" instructional strategy.
Thin and lean: focused on the essentials of the grade-level content to be learned in specific instructional units and coordinated with the grade-level curriculum.	Takes place in a survey assessment covering the previous year's entire set of content standards.
Administered "just in time" within a lesson, section of learning, or unit of study.	Assesses students once for the entire year at the beginning of the year, or infrequently throughout the year.
Largely formative in nature, administered by and for the teacher to provide instructional action, ideally informed by high quality instructional materials	Is largely summative in nature, e.g., based on standardization and/or external controls.
Designed to yield information about student knowledge and performance in content-specific ways	Is designed with summative uses prioritized, e.g., yields a scale score, supports comparability, or applies a unidimensional construct
Coordinated with other assessment-related aspects of the classroom and district	Is free-standing and not coherent with other assessments being administered.

assessment systems (e.g., unit-based tasks, interim assessments).	

Assessment Implications

Several assessment implications emerge from this paper's discussion:

Start with pre-assessments for the first few units

Some leaders might question the advisability of focusing pre-assessments only on the knowledge and skills necessary for the first few units, rather than for the whole year. The fact is, trying to drill students on *all* of the necessary precursor skills before moving on to "new" learning is not the most effective approach. Spending weeks or months on remedial teaching is instructionally inappropriate—and unnecessary for most students, even in this uniquely challenging school year. It also substantially reduces the time available for on-level instruction and learning, which can lead to significant opportunity costs for both teachers and students. Focusing remediation efforts on just the first few units will enable teachers to move students more quickly on to grade-level content.

In addition, testing students early in the school year on skills they might need for units to be taught much later in the school year would be premature; grade-level learning unfolds in real time, and teachers shore up important knowledge and skills all along the way, even as they are introducing new material. The results of an assessment given at the beginning of the year presumably would be obsolete once schools have been in session for a few months.

For example, in fourth grade students learn how to multiply and divide multi-digit (large) numbers. To do this, they first must understand place value and be able to confidently complete the four basic operations (addition, subtraction, multiplication, division) using single-digit numbers. An assessment conducted at the beginning of the year might reveal students' weaknesses regarding these concepts or skills. However, an early fourth-grade unit involves adding and subtracting multi-digit numbers, which enables teachers to shore up students' understanding of place value and their fluency with multi-digit numbers. Students then can use their improved understanding and fluency to master the more complex mathematical operations they will be introduced to later in the year.

Formal assessment issues. Two major aspects of more formal assessment practices are directly impacted by the in-place versus remote school scenarios: (1) whether the tests contain secure materials and (2) whether proctoring is needed to prevent cheating.

Secure test materials. Consider, for example, whether a test contains secure materials, such as test questions that the states and test vendors (for district- or state-wide assessments) need for linking the current test scores to an established test scale or when teachers reuse mid-term and

final exams each year, which they want to keep secure. Such a test could not be given to students in a remote setting without a high risk that test items will be compromised.

Need for proctoring. If a test has high stakes for students, such as a classroom test that is graded, teachers may be concerned about cheating. In these cases, schools need to find ways to schedule and administer the test on-site. Teachers may have to prepare multiple forms of such tests, if CDC requirements necessitate having only a certain number of students physically present at one time. Alternatively, schools might set aside special days specifically for testing, in which all the students for half the grades covered are on-site.

Another way to address this challenge is to administer assessments that are rich learning experiences and not easily gameable, to reduce or eliminate the need for proctoring. This would be a perfect opportunity to move toward assessments which require answers that cannot be discovered with a quick Google search (see Marion and Snider, 2020 for an elaboration of these ideas). Of course, even using rich, performance-based items in remote assessment does not prevent outright cheating of the type where one student either simply copies the response from another student or confers with another student or adult to produce the answer. Thus, even for this type of assessment proctoring or some other type of security might still be necessary.

Informal assessment issues

Activities that require synchronous connections to students, such as group or one-on-one checks for understanding, will require access to high-quality remote connections. There are several good tools to facilitate online conferencing, whiteboarding, graphing, polling, asking yes-or-no questions, and pacing impressions, but they all depend on the quality of hardware, software, Internet connections, and troubleshooting capabilities at both ends of the connection. With regard to trouble-shooting technical difficulties in real time, the physical classroom setup has more options to fall back on: if the whiteboard doesn't work or there is a technology or power failure, the teacher can use other methods to complete testing in the moment. Similarly, in a classroom setting, practices that depend on teachers noticing facial expressions suggesting negative or positive affect, and thus potential dissonance/puzzlement or understanding, will be difficult to implement effectively or across an entire classroom in remote settings. On the other hand, reading nonverbal cues will be a challenge even with in-place settings, if students are wearing face masks.

Strategies for adapting assessment practices to blended scenarios. For formal testing, schools may need to develop scheduling/organizational strategies to fulfill security and proctoring requirements. For less formal assessments, such as classroom quizzes and subject-matter tests, teachers could establish agreed-upon methods for students to ask questions or express concerns in remote settings. In such situations, the burden of responsibility for "checking in" shifts from the teacher to the student (and/or parents/guardians). This shift may be problematic: in a

classroom setting, students often communicate learning-related affect spontaneously and unconsciously. To manage this dilemma, teachers might need to use electronic polling or similar tools, in addition to more frequently and explicitly pausing to ask whether any students have questions or problems.

<u>Recommendations</u>

1. **Prioritize assessments to support instructional actions.** It will be important to closely align assessments with the local curriculum, or at least design them to assess the state's content standards. Note: in some cases, particularly when testing early literacy and numeracy, the assessments could be tied to general learning and developmental progressions while still providing instructionally useful information.
2. **Direct and prioritize resources to help teachers/schools create or use existing pre-assessments that are tied to specific instructional units**. As noted in this paper, the workload required to prepare for the next school year is extraordinary. State and district leaders have an opportunity to prioritize CARES Act and other funding for summer professional development opportunities for teachers and leaders, to help them develop pre-assessment and formative assessments tied to specific units of instruction early in the school year. Note: teachers and school leaders are encouraged to examine their high-quality instructional materials before creating new assessments from scratch.
3. **Avoid administering the spring 2020 summative assessment in the fall of 2020**. Using the spring 2020 summative assessment to meet this fall's assessment needs could pose great challenges. For a more complete discussion of the rationale for this recommendation, see Landl and Boyer's 2020 blog post[2] asserting that spring 2020 summative assessments would be at the wrong grain size, take too long to administer, and carry the potential for misinterpretation.
4. **Employ short, large-scale assessments sparingly—and use them to support an equity agenda**. These tests could include commercial interim assessments or custom-designed, state interim assessments. This paper outlines potential use-cases for state, district, and school leaders, as well as for parents/guardians, but does *not* recommend creating and administering separate assessments for each of these users. There could be a case for a short, large-scale assessment at the state level or at the district level, but not at both. The choice regarding whether this should be the focus of the district or state depends on the intended use-cases, values, and potential actions. Note: It will be key to base any decision on the state's or district's capacity and likeliness to actively use assessment results to monitor, address, and report on equity and access.
5. **Avoid administering large-scale assessments during the first week of school**. Students will have been out of their school buildings for six months. It makes little sense to

---

[2] https://www.nciea.org/blog/school-disruption/summative-state-assessments-can-wait

welcome students back to school with a test while they are still trying to get their feet under them. It is reasonable to expect considerable differences in performance between students tested during the first week of school and students tested a few weeks after they have had a chance to resume the school routine and brush up on key concepts and skills.

6. **Nothing might be better than something**. There will be a temptation to administer whatever assessments states and districts have at their disposal, without regard to whether the test results would serve the use-cases outlined above. Many might think that something is better than nothing, but if that "something" leads to misinterpretations and missed opportunities, then "nothing" might very well be the better choice. In other words, school, district, and state leaders might be better off relying on information gathered from teachers than by being distracted by "official" assessments that are not elucidating.

7. **Think through the use-cases**. There have been many complaints over the years that large-scale accountability tests have little or no instructional value. This is due, in large part, to a lack of understanding regarding the specific purposes and uses of various assessments. If the use-cases do not lead to defensible theories of action, the assessments should not be administered. For example, if leaders say they want to use the fall 2020 large-scale assessment results to direct additional resources to those most in need—but they have no discretionary resources or support—they would be better off not testing.

8. **Employ a hypothesis-testing mindset**. District, school, and state leaders cannot afford to wait until they receive assessment results this fall to begin planning. It would be best for these leaders and educators to hypothesize what students learned during 2019-2020 school year and what sorts of support they will need for the 2020-2021 school year, then use their hypotheses to begin planning organizational and instructional strategies for this fall. They are also encouraged to think through the types of assessment results that would require them to reject or modify their hypotheses. This *a priori* exercise is critical to avoid a situation in which data is misconstrued to confirm preconceptions, rather than interpreted objectively.

9. **Consider the implications for different grades and content areas**. As a follow-on to the use-case recommendations, educational leaders need not feel compelled to use the same assessment strategy for all grades and content areas. They are encouraged to tie their assessment choices to the specific needs and potential actions for each grade/content area combination.

10. **Do no harm**. Like physicians who take the Hippocratic Oath, assessment users and professionals must carefully consider the unintended negative consequences associated with various assessment uses. Many of these unintended consequences stem from the misinterpretations described in the table above. It is critical to think ahead about what could go wrong, because something definitely will.

As state leaders prepare for the 2020-2021 school year, there is a crucial opportunity to attend to what assessments can and cannot achieve, and for what uses. Assessments are validated for

specific purposes and uses. No assessment can fulfill multiple and diverse purposes well. Therefore, it is incumbent upon educational leaders to limit assessments to only those uses that are supported by evidence and logic, and that serve the highest-priority purposes: that is, supporting instructional utility and promoting educational equity.

References

Landl, E. and Boyer, M. (2020). Summative assessments can wait! Dover, NH. Center for Assessment. Retrieved from: https://www.nciea.org/blog/school-disruption/summative-state-assessments-can-wait

Marion, S. and Snider, M. A. (2020). Remote learning provides an opportunity to rethink assessment (and learning). Dover, NH. Center for Assessment. Retrieved from: https://www.nciea.org/blog/school-disruption/remote-learning-provides-opportunity-rethink-assessment-and-learning

Perie, M., Marion, S.F., & Gong, B. (2009). Moving towards a comprehensive assessment system: A framework for considering interim assessments. *Educational Measurement: Issues and Practice, 28, 3*, 5-13. https://doi.org/10.1111/j.1745-3992.2009.00149.x

Shepard, L. (2005). Will Commercialization Enable or Destroy Formative Assessment? Paper presented at the ETS Invitational Conference, October 10-11, 2005, New York City.