



2021 No. 073

# Kentucky's Response to the COVID-19 Pandemic: Steps Toward Ensuring Continued Assessment Program Relevance and Validity of Student Scores Final Report

**Prepared for:** Kentucky Department of Education  
Office of Assessment and Accountability  
300 Sower Boulevard  
Frankfort, KY 40601

**Prepared under:** Contract #1900004339

**Authors:** Arthur A. Thacker

**Date:** June 15, 2021

# Kentucky’s Response to the COVID-19 Pandemic: Steps Toward Ensuring Continued Assessment Program Relevance and Validity of Student Scores

## Table of Contents

Introduction.....	1
Suspending Accountability for 2020-21 .....	2
Recycling Previously Administered Test Forms .....	3
Reducing the Number of Items on Tests.....	4
Postponing Standards-Setting for New Assessments.....	4
Extending the Testing Window.....	5
Monitoring Students’ “Instructional Mode” During the 2020-21 Academic Year.....	6
Revising the Quality of School Climate and Safety Survey to Address NTI and Instructional Modes .....	6
Establishing a Research Agenda to Evaluate the Impact of the Pandemic on Student Performance .....	7
Addressing changes in the tested student population pre- versus post-pandemic.....	7
Evaluating the impact of the pandemic on special populations (English Learners (EL), and Students with Disabilities (SWD) .....	8
Evaluating potential differences in the impact of the pandemic on student subpopulations (based on economics, race/ethnicity, gender) .....	9
Evaluating the impact of returning to in-person or hybrid instruction on student performance.....	10
Evaluating the effectiveness of school/district recovery efforts .....	10

# Kentucky's Response to the COVID-19 Pandemic: Steps Toward Ensuring Continued Assessment Program Relevance and Validity of Student Scores

## Introduction

Kentucky was forced to cancel the administration of statewide summative assessments due to the COVID-19 pandemic in the spring of 2020. The cancellation of the Kentucky Performance Rating for Educational Progress (K-PREP) also eliminated any potential for computing accountability indexes for schools for the 2019-20 academic year. Students were sent home and the education community scrambled to quickly provide non-traditional instruction (NTI) options for students. These efforts continued through the 2019-20 academic year and well into 2020-21.

Students returned to in-person school in 2021 in stages. Many returned to hybrid or reduced schedule versions of their previous instructional routines. Some students continue to receive only online instruction. Students return to in-person school varies by district across the state. Kentucky suspended accountability requirements but is administering the K-PREP for 2021.

In preparation for the 2021 K-PREP, the Kentucky Department of Education implemented several changes to the testing program. These changes were designed to reduce the burden on students and teachers required to test in this very atypical school year, while maintaining sufficient rigor for accurate student-level reporting. The 2021 assessment also represents the best data for evaluating the impact of the COVID pandemic.

Kentucky's approach to 2021 testing was to reduce burden where possible, but to maintain the ability to use the test data to "look back to pre-COVID and look forward to Kentucky's future." This brief report will examine several of the assessment decisions Kentucky made to accomplish this goal. They include:

1. suspending accountability for 2020-21
2. recycling previously administered test forms
3. reducing the number of items on tests
4. postponing standards setting for new assessments
5. extending the testing window
6. monitoring students' "instructional mode" during the 2020-21 academic year
7. revising the Quality of School Climate and Safety Survey to address NTI and instructional modes
8. establishing a research agenda to evaluate the impact of the pandemic on student performance.
  - a. addressing changes in the tested student population pre- versus post-pandemic
  - b. evaluating the impact of the pandemic on special populations (English Learners (EL), and Students with Disabilities (SWD))
  - c. evaluating potential differences in the impact of the pandemic on student subpopulations (based on economics, race/ethnicity, gender)
  - d. evaluating the impact of returning to in-person or hybrid instruction on student performance
  - e. evaluating the effectiveness of school/district recovery efforts

Each of these decisions will be discussed below. The rationale for the decision will be described, followed by the likely implications for 2021 assessment results, implications of the decision for making comparisons to prior student performance, and likely implications for statewide testing in the future. There are, of course, other decisions that may impact student test scores for 2021 and beyond, but these most directly impact the testing system. Other policies Kentucky put in place to address the pandemic may also be important but are too tenuously related to test scores to directly investigate (e.g., extending free- and reduced-lunch programs to accommodate students working from home).

### *Suspending Accountability for 2020-21*

Kentucky applied for, and received, permission to suspend federally mandated accountability for 2020-21. Assessments will be administered, and students will receive score reports, but no accountability indexes will be reported to schools or districts and local education agencies (LEA) will be cautioned regarding interpreting the results. The decision to suspend accountability for 2020-21 was made because there are significant obstacles to producing fair and valid accountability indexes due to sweeping changes in educational practices brought on by the pandemic.

One main purpose of the statewide assessment system is to generate student performance data for inclusion in school and district accountability indexes. Aggregated test scores are combined with other performance indicators (e.g., graduation rates, attendance) to create an overall accountability index for each LEA. These indexes are used to gauge school and district performance and to target support and interventions where they are most needed.

One of most important underlying assumptions for accountability is that the actions of educators in schools and districts have a substantial impact on student performance. It would not be fair to judge schools and districts on metrics they could not influence. Similarly, for accountability, we assume that all schools and districts have essentially the same capacity to influence student performance. If both these assumptions are met, we can support the premise that better educational practices (e.g., instruction, scheduling, curricula, support systems, learning environment) yield better student performance, and that K-PREP scores capture student performance.

For accountability, Kentucky also examines changes in scores from year to year. Positive changes increase the accountability index and negative changes reduce it. Statewide summative test scores are used to estimate year-to-year trends as well. Kentucky bases student-level performance growth on a metric that does not require a common scale across grades, but it is important that test scores are comparable from one year to the next for different cohorts of students (e.g., fourth grade students in 2018 compared to fourth grade students in 2019). Tests are equated across years to allow scores to be directly interpreted from one year to the next.

The 2020-21 academic year called many of these underlying assumptions supporting accountability into question. First, the link between student performance and teacher effectiveness has not been clearly established for NTI. This is especially problematic when we consider that teachers may have extremely varied comfort levels with the technology requirements for distance learning. Even if all students participated fully in NTI, the shift represents a substantial shift in the skills required to deliver effective instruction. LEAs could not reasonably have been expected to foresee the need for such sweeping changes in educational practice as those brought on by the pandemic. These factors call into question whether the statewide summative tests are appropriate for gauging the effectiveness of educational practices for 2020-21, and whether evaluating the effectiveness of NTI would be fair to LEAs.

Second, the effectiveness of NTI relies on many factors that are outside the control of the LEA. LEAs have little control over students access to reliable internet, a quiet, appropriate place to attend classes and complete assignments and competing priorities for students' time and attention (e.g., if they are responsible for younger siblings during class time). These factors could also impact student learning during NTI and performance on tests but have little to do with educator effectiveness.

Third, districts and schools may have approached NTI differently and the return from online to hybrid or in-person instruction happened at different times during the school year depending on the circumstances and policies of each school district. We do not yet know how mode of instruction impacts student learning or whether one mode is substantively better than another.

### ***Recycling Previously Administered Test Forms***

Because Kentucky did not conduct statewide testing in 2019-20, the most recent statewide assessment data is from 2018-19. Normally, test items are replenished on a schedule so that no item remains on the tests for too many years in a row. This is done to limit exposure effects, or students and teachers seeing an item repeated many times on the annual assessment. If an item becomes overly exposed, it can impact instruction and curriculum in negative ways. Teachers might, knowingly or not, begin to teach content directly related to a specific test item. This phenomenon can narrow the curriculum and limit the capacity of the test item to contribute to a fair estimate of student academic ability.

There was no exposure of test items in 2020 because no spring summative assessments were administered. This reduces the potential impact of exposure because teachers and students have more time to "forget" items and because there was no reinforcement for recurring items from 2019 to 2020. The overall effect of the skipped year is that we do not need to be quite so concerned about repeating test items.

The reasons Kentucky chose to use previously administered test items and forms for 2021 have more to do with the state of public schooling for the 2020-21 academic year than any direct assessment issue. First, using previously administered forms provides the best data for "looking back." Normally, Kentucky uses a sample of previously administered items to equate test forms to prior years. This allows multiple test forms across multiple years to be on the same reporting scale and to be directly comparable. This sample of repeated items, or the "equating set," is usually no more than about half the items on the tests. We expect item performance to be relatively stable across years, so this provides a generous sample for typical equating. Because the academic years from 2019 to 2021 were not typical, there are concerns that items may not perform as they have in the past. Using a larger sample of repeated items allow us to eliminate items that perform very differently than they did in the past from the equating set without losing precision. Note that we will still use those items on the test, just not as part of the equating set.

A second reason for recycling test forms from the past in 2021 is to avoid using the potentially atypical 2021 data to establish item parameters on new test items. Statistics and parameters for new items become part of the item data system. Those data are used to create new test forms and may be used to pre-equate items to shorten the time between testing and reporting. Because we do not fully trust item parameters from 2021 due to the atypical instructional year, the elimination of accountability computation, and potential changes to the tested student population, it may be better to wait for a more typical academic year to establish data on new test items.

### ***Reducing the Number of Items on Tests***

The 2020-21 academic year in most Kentucky districts has consisted primarily of NTI, with students attending classes online for a reduced amount of time compared to in-person instruction. Some districts had students come back to in-person or hybrid (some combination of NTI and in-person) instruction as early as February, but many districts did not resume in-person instruction until April, and some continued to operate on a hybrid schedule until the end of the school year. Whatever the instructional mode, the time students spent with educators during 2020-21 was substantially less than in a “normal” school year.

With in-person instructional time at such a premium, the Kentucky Department of Education decided to reduce the burden placed on educators and students wherever possible. For that reason, summative tests for 2021 were shortened to maximize instructional time while still generating valid and reliable data. Field test items (items that are being tried out but do not count toward a student’s score) were eliminated from the 2021 assessments. The elimination of the 2021 field test cycle may mean that additional items will need to be field tested on operational forms in the future, adding a stand-alone field test to replenish the item pool, or extending the time that current operational items may be used before retiring.

Some tests were substantially shortened even beyond removing field test items. For example, one of the writing prompts was removed from the writing assessment. This means that for 2021, the on-demand writing test included only one-third as many items as the 2019 test. Reduction in test length beyond the field-test items means that scores are based on less information than in previous years. Kentucky decided to shorten the tests enough to make a difference in burden, but not enough to jeopardize overall student-level test scores. Overall scores (e.g., fourth grade mathematics, fifth grade reading) are based on sufficient items to ensure reasonable reliabilities. Student-level sub-scores, however, are based on a sub-sample of the overall scores. Sub-scores represent subcategories of the overall scores. For example, mathematics tests may include sub-scores for algebraic thinking, geometry, probability/statistics, and number and operations. Sub-scores are typically based on a relatively few items, but the reduction in test length could make the number so few as to render the sub-scores unreliable. Sub-scores will not be reported in 2021.

### ***Postponing Standards-Setting for New Assessments***

Kentucky had planned to conduct standards-setting on new assessments expected to be administered in spring 2020 or 2021. These assessments are based on revised academic standards, or the tested grade level may have changed. For example, the social studies standards were revised, and a new test was created to appropriately measure students’ achievement of the knowledge, skills, and abilities described by those standards. There were newly administered assessments for the regular and alternate assessments.

Standards-setting is the process by which “cut scores” are established on an assessment. The process involves gathering the judgements of content experts about the items that students in each performance category are likely to answer correctly. There are numerous ways of capturing this information depending on the standards-setting method chosen, but the underlying reliance on expert judgement is common across most methods. Once the experts have made their judgements, preliminary cut scores are applied and the proportions of students in each performance category is computed. This “impact data” is shown to the experts to allow them to reflect on their judgements in comparison with the student performance distribution. The experts may revise their original judgements after reviewing the impact data if it does not match

their expectation. There may be several rounds of examining the items, examining the impact data, and making revisions before final cut scores are recommended.

Basing a standards-setting on 2020 data was obviously impossible since end-of-year assessments were not administered in 2020. There are two main concerns with setting standards based on data from 2021. First, item level data may not be representative of overall student performance due to fewer students testing in 2021 and an atypical academic year. We expect fewer students to test due to COVID concerns and we do not yet know how severe an issue that will be. If considerably fewer students test, we must be concerned that the students who do test may not be representative of the entire state. This concern is coupled with concerns that performance may not be typical due to changes in instructional mode and the dramatic reduction of in-person instructional time. These issues could impact item level data, making items appear more difficult or easier than they would be in a more typical year. Item level data is used to inform standards-setting panelists in several ways depending on the methodology chosen. Even ordering items by difficulty could be challenging if the data are untrustworthy.

The second major concern is the potential for non-representative impact data. We do not know the impact of COVID on student learning. Impact data allows panelists to consider the impact of the cut scores they recommend by examining the percentages of students who would score in each performance category. If student performance was substantially impacted, that data may not be optimal for providing panelists with a realistic idea of how the test will categorize Kentucky students. They might adjust cut scores based on faulty assumptions about typical student achievement. For these reasons, Kentucky decided to postpone standards setting until 2022 data are available.

Despite the decision to postpone standards-setting, Kentucky still needs to establish a way to provide students' scores and performance categories on the new or revised assessments administered in spring 2021. To accomplish this, Kentucky will conduct an equipercentile linking procedure. Essentially, Kentucky will set cut scores for 2021 to mimic the prior version of the new assessment, or its closest analog (e.g., in high school, a grade level assessment in mathematics might be considered roughly analogous to an end-of-course assessment in algebra). This allows Kentucky to give students and educators a reasonable estimate of performance using a relatively simple mathematical procedure rather than a standards-setting. It should be noted that these mathematically determined cut scores will only be used for 2021 and will not be used for accountability. Accountability computations in the future will rely on scores from these tests based on more traditional standards-setting.

### *Extending the Testing Window*

Testing windows refer to the period during which tests are available to be administered in schools. Typically, states try to limit the testing window to the shortest time that is practicable to allow schools to administer tests to all students. Short testing windows limit potential test and item exposure. The less time the tests are available, the lower the likelihood that test content will be shared among students or educators. Shorter testing windows also reduce the probability that major test security issues (e.g., items or forms posted to the internet) will occur.

There are many factors that can play a role in how long a testing window needs to be. These include test length, availability of test administration resources (e.g., computer labs, staff availability to administer one-on-one accommodated assessments, availability of appropriate testing space), and school schedules. In 2021, these factors were more complex than in a more typical year. Students may not be attending school every weekday, or they may not be attending

for the full day. Students may also be required to adhere to social distancing policies that expand requirements for testing space.

Another concern related to the testing window is that many schools had just begun reopening in spring 2021 and struggled with establishing new processes for keeping students safe, providing instruction, transporting students, and all the other processes that were necessarily adjusted as students returned to school. Adding a brief testing window, shortly after students returned to school, would likely have been a substantial burden for school staff members.

These issues provided the impetus for Kentucky to expand the testing window. The hold on accountability for 2021 helps reduce concerns about exposure, but Kentucky must still be on guard for more serious security breaches. Administering shorter test forms with items from prior administrations limits the severity of such a breach.

### ***Monitoring Students' "Instructional Mode" During the 2020-21 Academic Year***

Most Kentucky students spent much of the 2020-21 academic year in non-traditional instruction. The decision to return to in-person instruction was not implemented statewide, and districts approached the return to in-person instruction differently. The Kentucky Department of Education began collecting data on the proportion of students who had returned to in-person schooling in early spring 2021. The data was collected at multiple time points to help understand the process of students returning to school. The data include information regarding the mode of instruction the students received upon returning (e.g., full-time in-person, hybrid, non-traditional instruction), and the proportions of students who had returned to each school at multiple dates throughout the spring.

This information may be combined with student performance data to help Kentucky better understand the impact of COVID. It may be useful for understanding the effectiveness of non-traditional instruction and hybrid models compared to full-time in-person instruction. The data may also be useful for understanding the interactions among demographic variables and the return to schools. For example, if some students returned later than others to in-person schooling, were those students more likely to be economically disadvantaged, higher- or lower-performing, belong to a particular student subgroup? These data can help provide context for the 2021 assessment results and may point to follow-up research or indicate where supports may be most needed.

### ***Revising the Quality of School Climate and Safety Survey to Address NTI and Instructional Modes***

Kentucky field tested its Quality of School Climate and Safety in 2019 and planned to administer the survey to the full population of students in spring 2020. The survey was to be administered to students during the spring testing window but was cancelled due to the COVID pandemic. Survey administration was restarted in 2021, but it was necessary to revise some of the survey questions to address changes in how students experienced school climate and safety during the pandemic.

Some of the safety questions on the original survey referred to the physical space of the school building and grounds, for example. Other questions asked about interactions with classmates, faculty, and school staff that traditionally would occur in person. When students receive instruction from their homes these questions may be less relevant or need to be clarified or contextualized so that students understand what each question is asking.

Survey items were also added to address opportunity to learn, which may be influenced by mode of instruction. For students who received entirely non-traditional instruction online, the learning environment includes the home space where the student did work and attended classes, and it also includes the virtual space where classes and support took place. These virtual spaces may have included virtual classrooms (e.g., Google classroom), or meeting spaces designed primarily for business rather than for education (e.g., Zoom, Microsoft Teams, GoToMeeting). Similarly, support for online instruction may have taken numerous forms, from simple emails back and forth to real-time student group meetings via instant messaging or videoconferencing (e.g., Skype). These factors help make up the online learning environment, and they were experienced by students in Kentucky who participated in non-traditional or hybrid forms of instruction during 2020-21.

Opportunity to learn questions that were added to the survey addressed access to tools to support instruction, access to instruction, and instructional experience. These captured information about student experiences when they were receiving instruction outside the school building, as well as their perceptions of the quality of their learning during non-traditional instruction.

New survey items will be used in several studies related to the impact of COVID planned for the fall of 2021. The survey items can be linked to students' academic performance and demographic variables, making them an extremely valuable for capturing students experiences during the pandemic. These data can also be linked to school- and district-level data to help us learn more about what policies and practices hold promise for future distance learning options post-pandemic.

Commented [AT1]: Emily—please revise and expand this section as needed.

### ***Establishing a Research Agenda to Evaluate the Impact of the Pandemic on Student Performance***

Kentucky has been very deliberate and cautious in their modifications to the testing and accountability system in response to COVID. They have emphasized the safety of students and educators over the collection and reporting of data whenever feasible, while ensuring that Kentucky continues to meet federal requirements. Any changes to the testing and accountability system were made while considering the implications of that change for looking back to pre-COVID (e.g., comparing student performance post-COVID to pre-COVID) and for looking forward (e.g., monitoring recovery efforts from spring 2021 forward). A major part of the Kentucky's efforts has been establishing a robust and flexible research agenda to investigate how the landscape of educational testing was impacted by the pandemic, how the pandemic affected student learning, and what we learned during the pandemic that can help Kentucky students succeed in the future. The following sections briefly describe some of the research Kentucky is planning for the near-term.

### ***Addressing changes in the tested student population pre- versus post-pandemic***

One of the biggest challenges in estimating the impact of COVID is that the numbers and proportions of students testing pre-COVID may be considerably larger than post-COVID. First, not everyone agrees that the pandemic is over. Many parents/guardians may continue to be concerned for the safety of their students if they return to in-person schooling. Some are reluctant to be vaccinated and may have concerns about reported COVID variants. Younger students may not be eligible for vaccinations and may continue to have some risk of infection. Irrespective of the reasons, states expect that fewer students tested in spring 2021 than in 2019.

Fortunately, the federal government has allowed states to defer the participation requirement in the Every Student Succeeds Act (ESSA) that requires at least 95% of all public school students to test for spring 2021 testing. When states can comply with the participation requirements, it limits how much concern must be given to sampling. If the tests are essentially census samples, then we can be assured that inferences drawn regarding year-to-year performance are comparable. When that census sample is reduced, we must be much more concerned that the students who are tested represent the performance of the full population. This is especially challenging if the reduction in the tested sample is neither intentional nor random.

If the students who do not test in spring 2021 represent a higher-performing subgroup, statewide test score means and percent proficient are likely to decline, irrespective of the impact of the COVID pandemic. If the students who do not test are lower performing, the opposite is true. Understanding how the sample of tested students changed from 2019 to 2021 is the first step toward estimating the impact of the pandemic.

Kentucky has made several plans to compare the tested population in 2021 to the population in 2019. In some instances, the population of students can be directly compared. Students who were third graders in 2019 will mostly be in fifth grade in 2021. We can therefore match the students who tested in 2021 with the same students who tested in 2019. Then, we can compare the performance in 2019 of the tested versus non-tested population in 2021. This should help generate a reasonable estimate of the differences in performance for the tested versus not-tested population that can inform other decisions about how to characterize the impact of the pandemic.

Kentucky will also generate similar comparison samples based on these matched groups. Once we create the sample of students from 2019 in third grade who also tested in 2021 as fifth graders, we can use that sample to create a comparison sample of third graders in 2021. Using propensity score matching, we can use what we know about those students to select a comparable sample for 2021. Propensity matching accounts for demographic variables and other performance information (e.g., grades, scores on other assessments) to generate comparable samples of students. This should allow us to create a matched sample for all grades. This will also allow us to select this comparable sample for subjects that are not tested every year (e.g., if we generate the sample for students who took the math test, we can then look at those same students' performance on the science test).

### ***Evaluating the impact of the pandemic on special populations (English Learners (EL), and Students with Disabilities (SWD))***

When we consider the impact of the pandemic on special populations, we must consider both sampling and performance. There is speculation that the most vulnerable students may be the least likely to test in 2021 and may also be the most impacted by the pandemic. Students with significant health issues may be more susceptible to the corona virus and secondary infections or other health concerns. This could cause parents/guardians to be more cautious in their decisions to return to school than those with healthier children. Parents/guardians with limited English skills may not have access to the most updated information on school reopening policies and guidelines. Their students may therefore return to school as quickly as students of native speakers. A first step in understanding the impact of the pandemic on these students is to accurately gauge their participation in testing in 2021.

A significant challenge for estimating the impact of the pandemic on the performance of special populations is that they often take different assessments than the annual statewide summative

accountability tests. Those with significant cognitive disabilities may take the alternate assessment. English learners may take an English language proficiency test. These scores are incorporated into a school's or district's accountability index in a typical year, but they do not generate scores that are directly comparable to statewide assessment scores. This means that they must be treated separately when we estimate the pandemic's impact on performance. This is especially challenging because the number of students who take these examinations is very small in a typical year and may be even smaller in 2021.

Many of the procedures we describe above for estimating the impact of the pandemic on test scores may simply not work if the data are very sparse. Propensity score matching depends on a sufficient sample of students to make reasonable matches from many variables. The fewer available matches, the weaker the comparison sample will be. Weak comparison samples can undermine potential inferences that might be made about the impact of the pandemic on these populations. Kentucky will explore these comparisons but will focus first on descriptive statistics related to the tested versus untested populations before determining what kinds of inferences might be warranted regarding student performance.

#### ***Evaluating potential differences in the impact of the pandemic on student subpopulations (based on economics, race/ethnicity, gender)***

The other major advantage of examining the tested versus not-tested populations from 2019 is that it will allow Kentucky to estimate whether students from specific subgroups are more or less likely to opt out of testing. It will allow Kentucky to discover if socioeconomic plays a major role in who tests. It will allow Kentucky to gauge if testing in 2021 is associated with race/ethnicity, urbanicity, and geographic area of the state. This information will be very important as testing resumes in 2022 and beyond when Kentucky must make decisions about accountability related to trend analyses, statewide, by student subgroup, and for schools and districts. Knowing who is testing and who is not will be key to implementing fair and reasonable accountability policies as Kentucky testing continues.

An important next step, once we have matched student samples from 2021 to 2019, will be to compare the tested versus non-tested groups' performance in 2019 by subgroup. This will allow Kentucky to investigate the potential interactions among subgroup membership, participation in 2021 testing, and student performance. For example, we might find that lower performing students were more likely to opt out of testing in 2021, but that tendency might be even more prevalent among Hispanic students. This information will help Kentucky better monitor subgroup performance as the tested population becomes more stable and it will help inform any information on subgroup performance trends from 2021 and beyond.

Once we have established comparison samples, we can also compare subgroup performance for 2019 versus 2021 for tested students. These analyses must be done for both years to ensure fair comparisons across years. Kentucky will make these comparisons using multiple metrics to help audiences understand how long-standing achievement gaps may have changed during the pandemic. Kentucky will report, for the comparison sample, subgroup differences based on percent proficient, differences in mean scale scores, and effect size differences. Schools and districts will likely be most interested in percent proficient differences because they use that metric for accountability. Mean scale score differences provide a more precise measure that is directly related to percent proficient. Using mean score differences will help guard against some of the limitations of percent proficient, which can hide important information if the score means are not similar. Finally, effect size differences allow Kentucky's data analysts and

scientists a means of making direct comparisons across grades and subjects, even when different scales are used and different cut scores for proficient are set.

### ***Evaluating the impact of returning to in-person or hybrid instruction on student performance***

Kentucky has been collecting information on when students have returned to in-person instruction during the 2020-21 academic year and whether that return continues to include some online components. Kentucky will characterize these students' instructional mode as; a) completely in-person, b) hybrid (some in-person and some online instruction), and 3) entirely non-traditional (online) instruction. The data set will also indicate the time they spent in each mode prior to the administration of the assessment (as a function of the calendar date on which they modes changed).

These two factors (mode, and time) can be used as covariates in correlations predicting performance from 2019 to 2021. This will allow Kentucky to begin to evaluate the effectiveness of each instructional mode. This data may also be coupled with location (region or district) and student survey results to better understand how mode interacts with region and opportunity variables. Taken together, these data may point out areas where students and their schools may need additional support.

### ***Evaluating the effectiveness of school/district recovery efforts***

If we assume that the pandemic had a negative effect on student performance, then it will be vital to support schools and districts in their recovery efforts to ensure that students impacted by the pandemic recover as quickly as possible. Kentucky's efforts in 2021 have been rightly focused on ensuring that there is enough descriptive data regarding the tested student population to allow for fair comparisons after the 2021 assessment administration. It is not enough to evaluate which schools and districts suffered the most, the data must also support those schools' and districts' efforts to recover.

We expect schools and districts to vary considerably in terms of participation and performance. That variance creates challenges for evaluating change and monitoring trends as students come back to in-person instruction in larger numbers. Understanding what makes sense for monitoring student performance changes when students do not return at the same time or when their experiences during the pandemic may have been quite different will be a challenge. It will be especially challenging for students who did not test in 2021.

We expect Kentucky's plans for supporting school and district recovery efforts to change depending on data analyses from 2021. The more we know about the impact of the pandemic on participation, instructional mode, and student performance, the better we can target data collection and supports for schools and districts. Understanding what we know and what we do not know will help Kentucky evaluate what steps might be necessary before new accountability indices can be fairly and accurately computed.