

Accountability Interrupted: Guidance for Collecting, Evaluating and Reporting Data in School Year 2020-2021

Chris Domaleski, Michelle Boyer, and Carla Evans
Center for Assessment

Prepared for the Council of Chief State Schools Officers

Pre-publication Draft

Introduction

Over recent months we have experienced extraordinary disruptions in nearly every aspect of our lives. We cannot say with certainty whether we are experiencing a temporary interruption, or if we are at a turning point in the history of our educational intuitions and traditions; a turning point that requires us to rethink how we conduct the business of education in the longer term. Certainly, there is a great deal of speculation on such questions among politicians, the press, experts, and in our everyday conversations. But since we do not have the luxury of historical perspective, we are faced with the need to make near-term decisions based on limited information about what is happening in our schools now, and what they might look like in the coming year.

One area impacted by the disruptions is school accountability and reporting. Typically, state education agencies (SEA) annually produce public information with respect to a wide range of indicators for schools and student groups as required by the Every Student Succeeds Act (ESSA). However, the U.S. Department of Education (ED) waived school accountability and many reporting requirements for the 2019-2020 school year (hereafter: SY20) for all states, territories, and the Bureau of Indian Education in late March and early April 2020 (U.S. Department of Education, 2020).

While many federal requirements were lifted, states did not abandon their desire to collect and share useful information to help stakeholders better understand and support students and schools. State leaders may be asking questions such as:

- What information should be publicly reported for SY20?
- How can we determine if the information collected is useful for the intended purpose?
- What new information should be collected and reported?
- How can we support appropriate interpretation and use of what is reported?

This paper was written to provide guidance with respect to these questions. These are challenging issues and there is no easy path forward. Thoughtful leaders understand that a complete freeze on sharing information prior to SY21 is not a good alternative, but neither is proceeding in a 'business-as-usual' manner. To help state leaders thread the needle, we suggest

a framework to guide decision-making regarding related to data collection, reporting, and distribution.

Context

Federal Reporting Requirements

The Elementary and Secondary Education Act of 1965, as amended by the Every Student Succeeds Act, establishes requirements for State Education Agencies (SEAs) and Local Education Agencies (LEAs) to prepare and disseminate annual report cards that provide information on the performance of schools and student groups in an understandable and uniform format (U.S. Department of Education, 2019). States have discretion regarding when report cards are issued, but they are generally encouraged to provide the reports before the start of the next school year or as early in the next school year as possible. In general, states are required to include on the annual report card information about public schools related to:

- *Student achievement data based on state assessments*
- *State accountability system information*
- Data for the Civil Rights Data Collection¹
- Educator qualifications
- Per-pupil expenditures
- State performance on the National Assessment of Educational Progress (NAEP)
- Postsecondary enrollment rates for each high school

An SEA may include on its report card any other information it believes will best inform parents, students, and other members of the public about the progress of each elementary and secondary school (U.S. Department of Education, 2019). For example, an SEA may include information on the percentage of students: requiring remediation in postsecondary education; acquiring career and technical education certifications; who drop out of school; etc. SEAs can also choose to disaggregate student achievement information by additional groups than what the law requires such as by youth in the juvenile justice system or different categories of English learners (e.g., recently arrived English learners, long-term English learners, etc.).

ESSA State Plans Assessment Waivers

As noted, ED approved state waiver requests releasing states from certain requirements; namely, student achievement data based on state assessments and state accountability system information (italicized in the preceding list). More specifically, states are released from submitting the following report card information: accountability system description; assessment results; other academic indicator results; English language proficiency assessment results; school quality or student success indicator results; progress toward meeting long-term goals and measurements of interim progress; percentage of students assessed and not

¹ CRDC includes rates of in-school suspensions; out-of-school suspensions; expulsions; school-related arrests; referrals to law enforcement; chronic absenteeism (excused/unexcused); incidents of violence; number of students enrolled in preschool programs; number and percentage of students enrolled in accelerated coursework to earn postsecondary credit while still in high school.

assessed; number and percentage of student with the most significant cognitive disabilities taking an alternate assessment; and information showing how students in a LEA and each school achieved on the academic assessments compared to students in the state and LEA.

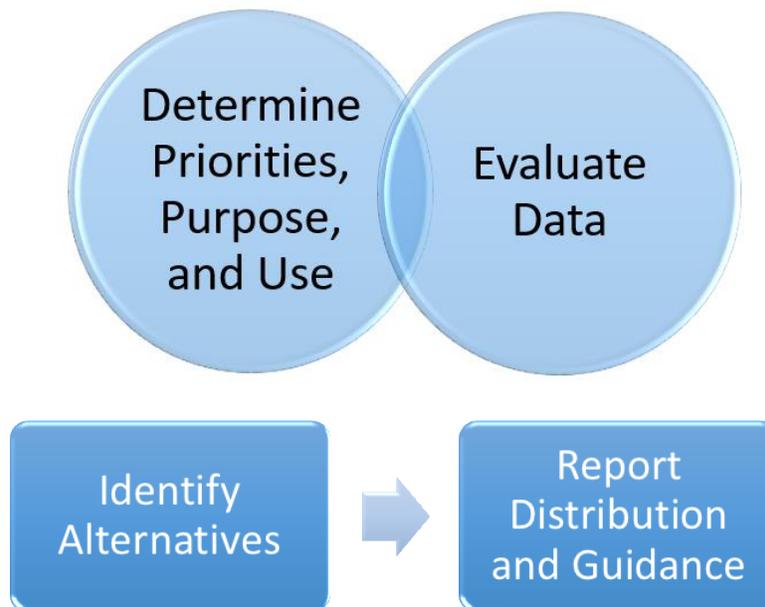
However, other report card requirements were not waived. These include high school graduation rates, data from the Civil Rights Data Collection, educator qualifications, per-pupil expenditures, state performance on NAEP, and postsecondary enrollment rates for each high school.

Framework for Supporting Decisions about Data Collection and Reporting

In this section we provide a framework that describes the prominent considerations to inform decision-making about data collection and reporting for SY20. We focus on data required for federal reporting, but we believe the framework can be applied more broadly. The framework is not intended to be entirely discrete or linear—as seen in Figure 1. While it may make sense for a state to start by determining their reporting priorities, purposes, and specifying use cases for the reports generated, it is equally valid for a state to begin by evaluating the current data context in light of federal reporting requirements. Or a state may consider both of these elements in an overlapping fashion. The key decisions made in the first two elements will then inform alternatives, which then directs the approach for distribution and guidance. Each element that makes up the framework is discussed in detail in this section.

Figure 1

Considerations to Inform State Decision-Making about Reporting for SY20



Determine Priorities, Purpose, and Use

In this “Accountability Interrupted” time period, providing useful information for decision-making in schools, districts, and states is an important heuristic for considering SEA data collection and reporting needs. Identifying and prioritizing the key use cases that must be supported is a critical first step to support decisions about what data to report if it is not federally required, or how to contextualize and report on data that is required, but incomplete or lacking in some way. To help states identify or determine their priorities and use cases, we suggest that SEAs consider the following questions:

- What decisions need to be made?
- Who needs the information to make decisions?
- What is highest priority?
- Is it feasible to collect the information?
- Are there potential unintended negative consequences associated with collecting or reporting the information?
- How will the data be used and interpreted?

Attending to these questions should help the state initially classify data reporting decisions into one the four categories depicted in Table 1.

Table 1

Determining Initial Data Collection and Reporting Actions

	Information is Available and Suitable for Proposed Use Case	Information is not Available or Questionable for Proposed Use Case
High Priority Use Case	Move forward	Further evaluation necessary
Low Priority Use Case	Move forward after other priorities are met	Suspend data collection and/or reporting for SY20

Naturally, if the state identifies a high priority use case and the disruptions related to the pandemic did not have an impact on the data typically used for this purpose, the state can move forward without delay. The challenge, however, is how to support the high priority needs when the data are not available or may have been impacted such that their use for this purpose is questionable. This leads to the next component of the framework.

Evaluate Data

To support the high priority use cases, it is important to evaluate the extent to which the necessary information can be defensibly used for their intended purposes. It is reasonable to start by taking stock of the condition of the legacy information the state typically collects.

One way to approach this task is evaluate the data elements with respect to the following factors.

- **Completeness:** *To what extent are elements of the data missing?*
- **Consistency:** *Were the data properties altered?*
- **Impact:** *Is it likely data values (e.g., performance) will substantially change?*
- **Practicality:** *Is it feasible and/or reasonable to collect and report the data?*

Completeness

As an initial check, one should evaluate the extent to which the data represent the full breadth and depth that was expected prior to the disruption. For example, with respect to assessment data, a complete file would include all planned assessments, for all districts and schools, for all participating students.

As a practical matter, the criterion for completeness is rarely 100%. It is typical for educational data to have missing cases for various reasons (e.g., students who were not present or declined to test, testing irregularities, etc.). Therefore, it is more technically appropriate to evaluate completeness with respect to the degree of deviation from standards during “normal” circumstances. It is important to note, however, that a file can appear to be mostly complete overall, but could contain pockets of missing data, such as by excluding cases disproportionately in a certain school or district. For this reason, checks for completeness should include multiple disaggregations (e.g., by school, student group, program, etc.).

Consistency

This attribute refers to the extent to which circumstances prompted a change in the way the data were defined, calculated, or collected. For example, a district may have changed the graduation requirements, such that all or some credit requirements were relaxed. As another example, schools or districts may have adopted a pass/fail policy for determining course credit and reporting student performance in lieu of grades. Under such circumstances, not only will the individual metrics be altered, but aggregations based on these elements (e.g., graduation rate or grade point average) will also be altered.

Impact

If the data are not complete or the properties have been altered, it stands to reason that results, such as performance summaries, will change. We refer to this as impact. For example, proficiency rates on the state test or percent of students earning course credit will likely change if cases are missing or the elements are calculated based on different rules.

However, it's important to note that the values can change based on other circumstances, even if the elements are complete and calculated based on the same procedures. For example, assessment performance will likely decline if students did not have an adequate opportunity to learn. As another example, results from school climate surveys completed by teachers or students could be very different based on experiences during the disruption. If the distribution is likely to be substantially altered, this raises questions about whether and how the data should be reported in the usual manner.

Practicality

Finally, making decisions about whether to collect and report data should include a “reasonableness check.” For example, will collecting the data cause an undue burden and deflect from higher priorities? Is it likely the data could be misunderstood, misinterpreted, or misused in a manner that could not be adequately addressed by providing guidance and support? If conditions suggest collecting and reporting data is not a good use of time and cannot be associated with a valuable purpose in light of the pandemic, it stands to reason the data should be excluded from plans for collection and/or reporting. This is true regardless of the outcome of the prior inquiries in this section.

It may be helpful to summarize the results of the evaluation with respect to the proceeding criteria. For example, the state may decide to bin the data elements into categories such as:

- **Green:** It is likely feasible and appropriate to proceed with reporting.
- **Yellow:** More information or analyses are necessary to determine whether or how the elements should be reported and used.
- **Red:** It is not feasible or appropriate to proceed with reporting the legacy element.

	Completeness	Consistency	Impact	Practicality	Category
Element 1	✓	✓	✓	✓	Green
Element 2	X	✓	X	✓	Yellow
Element 3	X	X	X	X	Red

To be clear, failing to meet any criterion likely precludes assignment to the green category. Moreover, failing to meet the practicality criterion is singularly sufficient to assign an element to the red category. For elements in the red category, the state may decide to delay data collection and reporting to a subsequent cycle or launch a new data collection and reporting initiative for the most urgent priorities. Otherwise, all elements coded yellow may then be further examined to determine the most appropriate alternative for reporting.

Identify Alternatives

When an evaluation of legacy data reveals gaps in the adequacy of information to support the state's reporting priorities, there are two general alternatives: modify the current approach or initiate a new data collection.

Modify Current Approach

Modifications refer to changes in the state's standard procedures for calculating or reporting the data. For example, one alternative is to change the business rules used to produce a reported indicator. Consider an element such as chronic absenteeism. The business rules will address features such as the number of days absent to classify the student as chronically absent and the enrollment period required to include a student in the calculation. It is possible to change both aspects to report a modified indicator for chronic absenteeism. This could be accomplished by shifting the end of the enrollment period to a date prior to when in-person attendance was suspended and identifying a threshold that maintains the same proportion of absences regarded as chronic used with the legacy indicator (e.g., 10%).

There are a variety of statistical adjustments that can be considered as well. These include approaches such as imputation to address missing data or altering the weights of a data file to achieve some target for representativeness. We mention these statistical approaches for the sake of completeness, but we are skeptical that such approaches are very promising under the circumstances. The perceived benefits and limitations of such approaches should be carefully evaluated by policymakers and technical experts, such as the state's Technical Advisory Committee (TAC).

Collect New Data

When legacy data, either intact or modified, are not available to support the state's high-value use cases, it may be advisable to consider a new data collection and reporting initiative. While there are a number of reasons to consider this alternative, in this section we focus on those specifically related to disruptions due to the pandemic. We classify these with respect to two broad categories: understanding impact and identifying areas of support.

Understanding Impact. This category refers to data collection to take stock of the previous and ongoing impact of the pandemic on districts and schools. For example, the state may have provided flexibility for meeting course credit or graduation requirements and it is important to understand the ways in which districts responded to that flexibility. Among other benefits, this will give the state context for supporting interpretation and use of indicators such as graduation rate.

It is also important to understand the range of practices deployed to support teaching and learning in a distributed and remote manner. What approaches did districts and schools use and what data may have been collected to gauge the efficacy of these models (e.g., student parent feedback, participation or engagement)? While these data are certainly not suitable for

high stakes accountability, the state may find it beneficial to partner with districts to help identify and share promising practices.

Identifying Areas of Support. A second reason to collect new information is to identify support needs. This refers to efforts to pinpoint ways the state can partner with districts and schools to address challenges associated with the pandemic. In what ways were districts able to meet the needs of students who may be food insecure? What are the primary needs to continue support services for students with disabilities? Did districts and schools have adequate technology, training, and other instructional resources? These data may be collected in a variety of ways, such as via surveys or virtual meetings with district leaders. Again, the point is not to collect this information for any high stakes purposes. Rather, such initiatives can help states more effectively identify and support high priority needs.

Another use case for new data collection pertains to helping the state develop plans moving forward. For example, the state may be issuing guidance for the timing and manner in which schools will open and operate in the fall. Understanding the capacity and needs of districts and schools to implement the guidance is vital.

Report Distribution and Guidance

Best practices for reporting are in many ways the same as they were pre-pandemic. Effective reporting follows directly from its defined purpose, intended use, and audience, where each is clearly articulated prior to the details of the design and distribution. When reporting follows these principles, it is less likely to be misunderstood, misinterpreted or misused, and more likely to be actionable, efficient, and effective. With this in mind, we suggest state leaders address the following areas in developing their plans for reporting and distribution: presentation and support.

Presentation

Zenisky, Mazzeo, and Pitoniak (2016) emphasize the importance of assessment reports based on: 1) the defined needs of the intended audience; 2) attaching clear and accurate meaning to the information reported; and 3) paying careful attention to how the information is displayed. This requires consideration of the different user groups and what level and type of information each requires. For example, general public reporting may be most effective and efficient when it includes high level summaries only; whereas educators and parents may require more detail about score meaning. Policy, technical or research-oriented consumers may require yet another view of the results that includes fine-grained details. While this guidance pertains to assessment reports, we believe it can be extended to a wider-range of reporting practices.

With respect to public presentations of performance data in SY20, we offer two specific cautions. First, we advise against using characteristics or metrics associated with the state's school accountability system given the purpose of reporting in SY20 is not to assess, rank, or classify school performance. For example, many state accountability systems will aggregate various elements and assign an index score for one or multiple indicators, such as a 'college

readiness index’ or ‘academic performance score.’ Even if the data exist to produce the index in the traditional manner, it is not advisable to do so given the likelihood that it will be interpreted inappropriately. It is preferable to simply report the information in their ‘native metric’ (e.g., percent of students who earned a qualifying credit or credential by category).

Our second recommendation is to consider presenting information with additional narrative information and visual cues that serve to distinguish any element that is not considered comparable to the same element in previous reports. For example, assume the state reports chronic absenteeism rates using a different business rule as described in a previous section. Such information might be presented with prominent explanatory text clarifying the differences. Moreover, the state may avoid graphic or tabular presentations of the current year rate next to those reported in previous years. The state may also adopt a distinct naming convention (e.g., modified or partial-year rate) and a unique presentation style for the indicator. The objective is to produce a report that minimizes the likelihood that users will make inferences, such as trend comparisons, that are not supported.

Distribution

A reporting plan should also address who will get what information and for what purpose. In general, there are three categories of distribution: 1) reports that are explicitly public, 2) reports that are neither protected nor public, and 3) reports that are secure.

The first category describes information available to stakeholders, most commonly via the state’s web-based reporting platform. This information is typically higher level and requires considerable support to help users interpret and use the data appropriately. Naturally, these data require a high level of quality assurance, which typically involves careful evaluation and methodical review before release. The second category refers to data distributed in a more limited manner, such as to district leaders. These data may contain more detail than public reports and may be accompanied by distinct guidance or documentation (e.g., written for a more technical reader). However, this category of information may be available publicly by request. For this reason, states should be prepared to provide appropriate guidance and support for public release as needed. The third category refers to data that are secure and are shielded from public release. For most states, this is likely a narrow reporting category, such as reports that contain personal identifiable information.

Support

In the case of SY20 reporting, we urge states to consider additional supports to bolster accurate interpretation of assessment results under the very unique circumstances in which the data were collected. This is especially true if the data were modified in any manner prior to reporting. It also applies to the need to provide contextual information on topics such as student engagement, teacher and student access and use of technology supports, and variable policies across districts and schools (e.g., rules for assigning course credit).

In such cases, a state should offer supports that clearly articulate the appropriate interpretations and uses of the data. Supports may include resources such as supplemental documents, training sessions tailored to the intended audiences such as teachers, or it may be embedded in reporting with attention to the clarity and consistency of the message.

Finally, we urge states to keep detailed records and documentation of all procedures for calculating, evaluating, and distributing the information. This will help ensure procedures are complete and consistently applied.

Summary: Do's and Don't's for Reporting

In summary, we recommend the following actions and considerations related to SY20 reporting:

- Establish use cases and the audience for each
- Determine the level of information required
- Design for clear communication to avoid misinterpretation
- Consider alternative formats or data visualizations
- Consider inclusion of related contextual information
- Where the potential for misunderstanding or misuse is high, offer supporting training or documentation to the intended audience

We recommend that states mitigate opportunities for misuse by:

- Avoiding the use of incomplete, or COVID-19 influenced achievement data to draw strong conclusions overall, by student groups, or about trends
- Avoiding the use of indicators for high-stakes purposes, especially when those indicators were affected by limited ability to collect accurately (e.g., chronic absenteeism, suspensions, graduation) to examine trends
- Over- or under-interpreting data with missingness or expected effects due to disruptions

Final Thoughts

It may be tempting to think of SY20 as a 'one-off' – an extraordinary year that prevents any serious efforts to collect and distribute data. However, we suggest another perspective. While the challenges of reporting are certainly substantial, so too is the need for good information to inform school support initiatives. Effective reporting practices will require resolve and resourcefulness, but these efforts are critical. Presently, it's too soon to know if the disruptions will persist in 2020-2021, but early signs suggest some deviation from 'typical' practices are likely. If so, the need for innovative solutions to produce, share, and support useful information will also persist.

References

- U.S. Department of Education. (2019). Opportunities and responsibilities for state and local report cards under the Elementary and Secondary Education Act of 1965, as amended by the Every Student Succeeds Act: Non-regulatory Informational Document. Washington, DC: U.S. Department of Education.
- U.S. Department of Education. (2020). ESSA State Plan Assessment Waivers. Retrieved from <https://oese.ed.gov/offices/office-of-formula-grants/school-support-and-accountability/essa-state-plans-assessment-waivers/>
- Zenisky, A., Mazzeo, J., & Pitoniak, M. (2016). Toward improving the reporting of test score results. In C.S. Wells & M. Faulkner-Bond (Eds), *Educational Measurement: from foundations to future* (pp335-355). New York: The Guilford Press.