



THE
COLORADO
EDUCATION
INITIATIVE

**MEASURES OF STUDENT LEARNING
IN EDUCATOR EVALUATION:**

GUIDING QUESTIONS FOR MSL SYSTEMS



Introduction

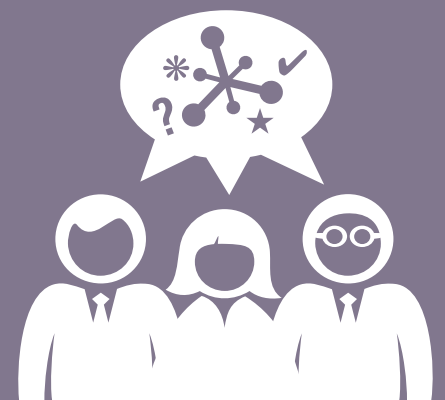
In 2010, Colorado Senate Bill 10-191 (S.B. 10-191) established new evaluation requirements for teachers and principals in all Colorado school districts, with the twin goals of supporting educators' professional development and positively affecting student growth. S.B. 10-191 requires that at least 50 percent of an educator's annual evaluation be based on multiple measures of student learning (MSL) in relation to the Colorado Academic Standards.

By thinking through the nuance of their MSL systems, school districts can ensure those systems provide educators with fair, accurate data about their performance and useful information about how to improve results. Developing an MSL system is complicated work that is new to many districts; as such, a commitment to continuous improvement is imperative. To help districts in this effort, The Colorado Education Initiative (CEI) engaged the Telluride Research Group (TRG) to invite school districts to share information about their MSL systems for inclusion in a statewide study of trends. TRG ultimately collected information from 53 Colorado districts representing a diversity of sizes and settings, from small rural to large urban, and including almost 60 percent of Colorado's teachers and students.

TRG created an anonymous MSL profile for each district, populated with its information. A panel of experts from CEI and the Colorado Department of Education (CDE) — including school district staff; former school and district leaders; and representatives with expertise in the state MSL system, psychometrics, and statistics — reviewed the anonymous profiles to understand the state's MSL landscape and give districts individualized questions to consider for ongoing implementation.

This document outlines overarching themes that emerged across districts and presents guiding questions for all districts to consider as they develop and improve their MSL systems in the coming years. Guiding questions are organized into four themes:

1. **Overarching system design:** Here, districts consider whether their MSL system reflects a balanced assessment system, ensures comparability and fairness across educators and measures, and results in any unintended consequences.
2. **Measure selection:** In this section, districts consider the selection and design of the individual measures that comprise their MSL system. In particular, they are asked to reflect upon their decisions regarding individual versus collective attribution, the selection of weights and measures, and target setting.
3. **Data use and assessments:** Next, districts examine their decisions regarding the use of pre- and post-assessments, student learning objectives (SLOs), state summative assessments, and data from the School Performance Framework (SPF) or District Performance Framework (DPF).
4. **MSL process and stakeholder engagement:** Finally, districts consider the MSL process and stakeholder engagement by reflecting on educator input, transparency and timing, and assessment literacy.



1) OVERARCHING SYSTEM DESIGN

Balanced Assessment System

Decisions about which assessments and measures to use and value within a district are best made on the basis of district goals and priorities, and the expectations for what students need to know and be able to do as outlined in the Colorado Academic Standards. MSL systems that are founded directly on a district's overarching assessment framework connect evaluation measures to the overarching priorities of the district and are better able to motivate instructional changes and give educators meaningful information about teaching and learning. **CDE's Measures of Student Learning guidance** includes a six-step process that helps ensure the measures selected as part of an evaluation system thoughtfully reflect district values.

How has your district used your comprehensive assessment system to design an MSL system that reflects district goals and priorities and the Colorado Academic Standards?

- a. Is your district's assessment system balanced? Does it reflect your district's goals and priorities, and the expectations for what students should know and be able to do as outlined in the Colorado Academic Standards?
 - In what ways is your MSL system a reflection of your assessment system?
 - How have you ensured that the MSL system is not the driving force in determining which assessments are important to your teachers and district leaders?
- b. What evidence do you have that the assessments included in each educator's evaluation are a balanced reflection of your district's assessment system and reflect the range and scope of key activities undertaken in the classroom? Moreover, what evidence do you have to suggest the assessments provide a balanced picture of student performance for each educator?
- c. Have you ensured that each educators' measures of student learning are:
 - Aligned to the Colorado Academic Standards?
 - Comprehensive (i.e., they cover the majority of the standards and content)?
 - Related to classroom instruction and/or practice?
 - Content-relevant (i.e., reflective of what each educator teaches or contributes to in a meaningful way)?
- d. Have you been thoughtful in using assessments for their intended purposes?



MSL Comparability and Fairness

The **rules** promulgated by the State Board of Education for S.B. 10-191 implementation ensure that measures used as part of an MSL system “are comparable among Teachers of similar content areas and grades.” This focus on comparability is an important step in ensuring fairness for educators — defined loosely as a system where all educators from various content areas and grade levels, serving various student populations, have an equal opportunity to earn each rating.

Are you confident that your district's MSL system appropriately accounts for issues of fairness and comparability?

- a. Does the design of your system encourage comparability? That is, are educators being assessed in a manner at least roughly equivalent to one another?
- b. How does the design of your system ensure *fairness* for all educators in all content areas?
- c. What evidence do you have that it is equally challenging for all teachers — regardless of content or grade level — to earn each rating?
- d. Is the proportion of collective versus individual attribution relatively consistent across educator types? If not, what rationale can your district provide for any variance in the system to ensure stakeholder support?
- e. Have you thoughtfully considered which measures can be appropriately applied to new educators? For new educators, how are you approaching the use of multiple-year measures like SPF and DPF?

Unintended Consequences

Decisions regarding the inclusion (or exclusion) of measures in an MSL system come with a host of intended and unintended consequences. In reflecting on an MSL system, consider whether it creates unintended consequences.

How has your district considered potential unintended consequences as you made decisions about system design and communicated with stakeholders?

- a. What does your system encourage teachers and/or principals to do or focus on? In what ways might your system inadvertently prioritize certain content or curriculum over others?
- b. How might the measures used in your system encourage or discourage educator retention or recruitment in certain grades or subjects, or certain schools or programs (e.g., by encouraging a belief that some schools or subjects make it easier to earn an effective rating)?
- c. Are there any other ways in which your system might send teachers or principals the wrong messages, or encourage them to do things that you do not necessarily want them to do (e.g., by incentivizing teachers to encourage learning more at certain times of the school year because of how that learning will be reflected on assessments used in their evaluations)?
- d. Have you considered the unintended consequences that may be associated with the selection of growth [e.g., Student Growth Percentiles (SGPs) versus achievement (e.g., percent proficient) measures]? (See the [State Summative Assessments](#) section for more information.)
- e. What are your district's business rules for student inclusion (e.g., which students are attributed to which educators)? Are certain student groups disproportionately impacted by these rules? Are there students who are unassigned to any educator? Are these rules consistent with your district's values?

2) MEASURE SELECTION

Individual and Collective Attribution

The S.B. 10-191 **rules** also require that teachers' evaluations include both an individually attributed and collectively attributed measure.

Have you ensured that the divide between individual and collective attribution reflects your district's values?

- a. How might the division of individual and collective attribution motivate educators to do or spend their time on certain activities? Are those the activities you want educators to focus on?
- b. What evidence do you have to suggest that the relative weights of the individual and collective measures do not artificially adjust overall educator ratings (e.g., by ensuring all teachers receive a proficient rating on the growth side, and providing an associated "boost" to teachers who are not rated at least proficient on professional practice)?

Selecting Weights and Measures

Step 2 of **CDE's Measures of Student Learning** guidance provides directions for selecting and weighting measures.

How have you considered district and local priorities in selecting weights and measures for your MSL system?

- a. What is the minimum number of measures (i.e., pieces of the pie) you feel you need to achieve a comprehensive body of evidence?
- b. How have you balanced the need to have enough measures to give a comprehensive view of each educator's work with students without slicing the pie into too many small pieces?
- c. Do any of your measures overlap? That is, are you *inadvertently* using the same assessment or measure more than once [e.g., SPF growth ratings and school reading median growth percentiles (MGPs), both of which are based on TCAP reading growth percentiles]?
 - If you are using an assessment or measure more than once, does the total weight reflect the value of that assessment within the philosophy of your district?

Target Setting

For a measure to be included in the MSL system, districts should thoughtfully approach target setting — and consider students' baseline performance — to evaluate whether student performance is much lower than expected, lower than expected, expected, or higher than expected (as outlined in **CDE's Measures of Student Learning** guidance).

Are your targets reasonable and defensible?

- a. What baseline data can you use to assist in setting targets?
- b. When considering available baseline data, are the targets rigorous but achievable?
- c. Are the measurement and statistical methods used to set targets and evaluate performance appropriate for your context and data or analytical capacity?
 - Does your district have the technical or analytical capacity to execute planned analyses and/or growth calculations?
 - Can you communicate the targets and analyses to educators in accessible, understandable ways?
 - Have you considered the number of students (i.e., n-size) in your district — both overall and assigned to each educator — to determine whether there are methodological limitations based on n-size?
- d. Does your district have a process to vet and validate targets once they have been set? If not, is there a plan in place to develop such a process?

3) DATA USE AND ASSESSMENTS

Pre- and Post-Assessments

Measures that are based on growth between a *pre-assessment* (an assessment given at the start of the year, semester, or unit to measure student proficiency in specific content and skills) and a post-assessment (another assessment given at the end of the year, semester, or unit measuring proficiency in the same content and skills) present special measurement challenges.

Has your district considered the additional difficulties in measuring growth from a pre-assessment?

- a. It is not always appropriate to [use the same instrument](#) as a pre- and post-assessment. In these circumstances, what is your district's plan to support educators in developing robust pre-assessments that measure the planned content and appropriate standards?
- b. How does your district plan to calculate growth between the pre- and post-assessment?
 - In making this determination, have you weighed the analytical capacity of your district and potential data limitations (e.g., the number of students, or n-size)?

SLO Process and Supports

Many districts are also considering the use of Student Learning Objectives (**SLOs**), where teachers identify assessments and set meaningful goals for student learning over time rather than using centrally determined assessments and targets for all teachers. SLOs are attractive to districts because they are applicable across grades and content areas, although they require transparency and training to be implemented in a high-quality way.

How has your district ensured that educators are prepared to select meaningful measures and targets?

- a. Who is making the decisions about the assessments used in an educator's SLO?
- b. Does your district have a plan to help educators select rigorous yet attainable targets for their SLOs? (See the [Target Setting](#) and [Assessment Literacy](#) sections for more information.)
 - What supports will be available to educators on an ongoing basis?
 - Does your district have a process to vet and validate SLO targets once they have been set? If not, is there a plan in place to develop such a process?
- c. How has your district ensured that the SLO process provides educators with meaningful and actionable feedback on instruction?



State Summative Assessments

Although the S.B. 10-191 [rules](#) require the use of state summative assessment data, when available, districts must make several decisions regarding their use. In its report, *[Using Student Growth Percentiles for Educator Evaluations at the Teacher Level](#)*, the National Center for the Improvement of Educational Assessment released technical guidance regarding the use of SGPs in educator evaluations. (An [executive summary](#) is also available.)

How has your district addressed the technical issues associated with using state summative assessments in educator evaluation systems?

- a. Has your district consulted [guidance from CDE](#) regarding the timing and use of state summative assessments in educator evaluations?
- b. What are your plans for transitioning to new statewide assessments (e.g., PARCC)? Have you consulted CDE's guidance related to the [accountability transition](#)?
- c. Has your district considered whether your measures capture growth or achievement?
 - By focusing only on TCAP growth, are you inadvertently excluding:
 1. Other state assessments (e.g., social studies or science, ACT, CoAIT)?
 2. Grades (e.g., third grade)?
 3. Students (e.g., new or transfer students from out of state, students missing prior year(s) of data, students who do not take state assessments)?
- d. If you are relying on achievement measures, have you considered possible unintended consequences from the use of achievement or status measures (e.g., impacting the recruitment and retention of teachers in low-achieving schools)? (See the [Unintended Consequences](#) section for more information.)

Use of SPF or DPF Data

Many districts are exploring the use of data from the School Performance Framework (SPF) or District Performance Framework (DPF) in their MSL systems because they are readily available and easily satisfy the legislative requirements to include Colorado state summative assessment results and Colorado growth data, when available.

If your district is considering the use of SPF or DPF data, have you consulted CDE's guidance document, [Using Colorado School/District Performance Frameworks in an Educator's Body of Evidence for Evaluation](#)?

- a. Have you considered the multiple ways to use the data from the SPF or DPF, as outlined in the guidance from CDE (e.g., a straight rating or the change in percent of points earned)?
- b. In selecting your measures related to the SPF or DPF, have you considered whether it captures student growth (e.g., student growth percentile and/or SGP) or student achievement (e.g., percent proficient or advanced)? While either of these choices is defensible, it is important to be purposeful about this decision and to weigh the potential [unintended consequences](#).

4) MSL PROCESS AND STAKEHOLDER ENGAGEMENT

Educator Input

District processes that adequately incorporate educators' voices and perspectives will benefit from educators' expertise and buy-in. Preliminary results from **CEI's Integration Project** suggest that teachers who participate in the selection of the MSLs used in their evaluations are significantly more likely to respond positively regarding the use and validity of MSLs than those who did not report participation in MSL selection.

In what ways did your district solicit and incorporate educators' voices and perspectives in the MSL decision-making process?

- a. How has your district included educators across grades and content areas?
- b. Has the role of educators been collaborative? Have they contributed to the design of the MSL system? If not, how will your district engage stakeholders and ensure buy-in?

Transparency and Timing

Transparency and the provision of timely information to educators are critical to building trust in the system.

How has your district ensured transparency with regard to the MSL system?

- a. Are educators given information regarding the MSL system, measures, assessments, and targets early enough to allow them to understand expectations and plan curriculum efficiently?
- b. Do educators have the necessary information to understand the measures used in their evaluation? Is there a communication plan in place?
- c. Is there transparency in the target-setting process?
- d. Has your district taken the appropriate safeguards to protect student and educator data privacy and personally identifiable information, in compliance with all applicable federal, state, and local laws, including the Federal Educational Rights and Privacy Act (FERPA)? Appropriate safeguards should include district policies, procedures, data systems security, and employee training. CDE has developed [guidance](#) for school districts on data privacy and security.

Assessment Literacy

An MSL system and its associated processes should ideally contribute to increasing assessment literacy among educators.

Do you have a plan for increasing assessment literacy among your educators?

- a. Have you considered the current levels of assessment literacy within your district?
- b. What plans do you have in place to ensure that limited assessment literacy does not impact educators and their ability and willingness to engage with the MSL system and the measures used in their evaluations?
- c. If your district is using SLOs, do educators have the assessment literacy they need to set appropriate targets? If not, how does your district plan to address this need?

