



THE
COLORADO
EDUCATION
INITIATIVE

Measures of Student Learning in Educator Evaluation:

Guiding Questions for MSL Systems



October 2016

TABLE OF CONTENTS

INTRODUCTION	1
Background	
Guiding Questions	
OVERARCHING SYSTEM DESIGN	2
Balanced Assessment System	
MSL Comparability and Fairness	
Unintended Consequences	
MEASURE SELECTION	4
Individual and Collective Attribution	
Selecting Weights and Measures	
Target Setting	
DATA USE AND ASSESSMENTS	6
Assessment Literacy	
Pre- and Post-Assessments	
SLO Process and Supports	
State Summative Assessments	
Use of SPF or DPF Data	
MSL PROCESS AND STAKEHOLDER ENGAGEMENT	9
Educator Input	
Transparency and Timing	
Continuous Improvement and System Evaluation	

INTRODUCTION

Background

Since the 2013-14 school year, Colorado districts have been creating and implementing Measures of Student Learning (MSL) systems that comprise the 50% of an educator's evaluation that is based on student academic growth. To help districts in this effort, [The Colorado Education Initiative \(CEI\)](#) has engaged in a multi-year study focused on MSL system trends across the state beginning in the 2013-14 school year. During the 2015-16 school year, CEI contracted with [Slope Research \(Slope\)](#) to gather information about MSL system design and implementation from school districts across Colorado. This most recent study also had a focus on Student Learning Objectives (SLOs). Slope collected information from 56 districts representing 31 percent of all Colorado districts and 46 percent of Colorado's teachers and students.

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.

Guiding Questions

Guiding questions are organized into four themes:

- 1. Overarching system design:** Here, districts consider whether their MSL system reflects a balanced assessment system grounded in the standards and district curricula, 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 reflect on the importance of assessment literacy in implementing their MSL systems and examine their decisions regarding the use of pre- and post- assessments, 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 continuous improvement.



OVERARCHING SYSTEM DESIGN

Balanced Assessment System

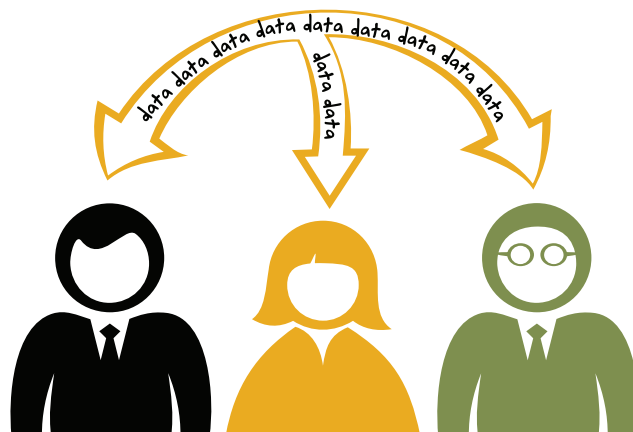
Decisions about which assessments and measures to use and value within a district are best made on the basis of district goals and priorities and the expectations for what students need to know and be able to do as outlined in the [Colorado Academic Standards](#). MSL systems that are founded directly on a district's overarching assessment framework connect evaluation measures to the 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 grounded your MSL system in the Colorado Academic Standards?

- How have you ensured that the Colorado Academic Standards are the driving force in determining which assessments are important to your teachers and district leaders?
- Does your district's assessment system reflect the expectations for what students should know and be able to do as outlined in the Colorado Academic Standards? Was it created by first critically examining your instructional priorities and aligning assessment systems to the most important elements? Does it include a variety of summative, interim, and formative assessments at each grade level and content area?
- What evidence do you have that the assessments included in each educator's evaluation reflect your district's priorities and the range and scope of key activities undertaken in the classroom?

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

- [Is your district's assessment system balanced?](#) Does it reflect your district's goals and priorities?
 - o How have you ensured that the MSL system aligned with your district's broader assessment system, but is not singularly driving decisions about what is considered important to your teachers and district leaders?
 - o What evidence do you have to suggest the assessments used in your MSLs provide a comprehensive picture of student performance for each educator and give teachers meaningful information about student learning?
 - o How is each assessment related to the relevant grade-level and content standards and adopted curriculum?
 - o In what ways do the collection of assessments used in your system work together to provide a holistic picture of student learning?
 - o Is the assessment schedule mapped meaningfully to the cycle of instruction for each grade/subject?
- Have you ensured that each educators' measures of student learning are:
 - o Aligned to the your district's values?
 - o Aligned to the priorities and needs set forth in your district's [Unified Improvement Plan](#)?
 - o Reflective of the unique context of your district and students?
- Have you been thoughtful in [using assessments for their intended purposes](#)?



MSL Comparability and Fairness

The State Board of Education adopted rules for [S.B. 10-191](#) that require measures used as part of an MSL system to be 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.

What evidence do you have that it is equally challenging for all teachers — regardless of content or grade level — to earn each rating?

- Is the proportion of collective versus individual attribution relatively consistent across educator types? If not, what rationale does your district provide for differences in collective and individual measures across different types of educators?
- Current-year assessment data can only be used in MSLs if results are available at least two weeks prior to the last day of school. This limits the use of most state assessment data to data from the prior school year (as well as some other assessment results that may not be available in time). Have you thoughtfully considered how this impacts measure selection with regard to fairness and comparability?
 - How does this impact the selection of measures for new educators in particular, for whom assessments from the prior year are not available?

Unintended Consequences

Decisions regarding the inclusion (or exclusion) of measures in an MSL system come with a host of intended and unintended consequences, which may impact student learning, educator satisfaction, and school culture/climate. In reflecting on your MSL system, consider whether it creates unintended consequences.

How has your district considered potential unintended consequences with regard to both system design and communication with stakeholders?

- 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?
 - 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 it may be easier for some schools or subjects to earn an effective rating)?
 - 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 “teach-to-the-test” by focusing on certain content over the full curriculum because of how that learning will be reflected on assessments used in their evaluations)?
- 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.)
- What are your district’s business rules for student inclusion in measures of student learning (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? What message might this inadvertently communicate about the value of those students’ learning?
 - Do your inclusion rules provide fair and meaningful sets of data points for all educators?
 - Are your inclusion rules easy for educators to understand?
- How might your district get feedback on potential unintended consequences from teachers, building leaders, and district leadership?
- How might your district seek an outside perspective on your system in order to more effectively identify potential unintended consequences?

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 of student learning.

Have you ensured that your district decisions about the relative weight of individual and collective measures are defensible and reflect your district's values?

- How might the division of individual and collective attribution motivate educators to prioritize certain activities? Are those the activities you want educators to focus on?
- Has your district ensured that collective measure(s) are meaningful for all educators?
- What evidence do you have to suggest that the relative weights of the individual and collective measures do not artificially adjust the overall MSL /Standard VI rating either higher or lower (e.g., by heavily weighting a collective measure where every educator is assured a particular rating)?
 - Have you also considered how inflated or deflated ratings on MSLs may impact the accuracy of overall educator ratings?

Selecting Weights and Measures

Selecting weights and measures for inclusion in an MSL system is critically important in that it demonstrates your priorities and values. These decisions also provide an opportunity to communicate expectations about student learning and teaching. Step 2 of [CDE's Measures of Student Learning guidance](#) provides directions for selecting and weighting measures.

How has your district ensured that the selection of measures, and determinations about their associated weights, were made purposefully and in alignment with your goals and values?

- Do the relative weights of the MSL measures used for each educator reflect the instructional priorities for that grade and content area?
 - What is the minimum number of measures (i.e., pieces of the pie) you feel you need to achieve a comprehensive body of evidence?
 - 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?
- Have you considered potential duplicity in measure selection? 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 CMAS 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?
- How have you accounted for issues arising from cycles of instruction and assessment schedule(s) when selecting specific measures?
 - Do assessment timelines match cycles of instruction in order to enable the provision of meaningful information to teachers?
 - Will the measures selected have results that are available two weeks prior to the last day of school?
- Do you have protocols or feedback loops in place to determine whether your MSL results are consistent with other measures of educational success in your district? For instance, you could examine the correlation between MSLs and Professional Practices or how the results of assessments used in your MSLs correlate with other assessment results.

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 academic growth 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?

- What baseline data can you use to assist in setting targets?
 - How are you ensuring that the source of baseline data is appropriate and comparable across like content and grade level groups?
 - When compared to available baseline data, are the targets set for each educator ambitious but achievable?
 - Do your targets ensure that students will make significant learning gains, sufficient to progress to proficiency or other performance standards established through the Colorado Academic Standards, local standards, or other adopted learning goals (e.g., through a 504 plan or formal IEP)?
- Are the measurement and statistical methods used to set targets and evaluate performance appropriate for your context, the available data, and the analytical capacity of your district?
 - Does your district have the technical skills or analytical capacity to execute planned analyses and/or growth calculations? If not, can you identify other approaches that are less methodologically rigorous but still provide meaningful information to educators about student and teacher performance?
 - Have you considered consulting other districts or outside providers (e.g., institutes of higher education or technical assistance providers) for assistance in creating or evaluating your targets?
 - 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?
- Does your district have a process to evaluate the defensibility and rigor of targets once they have been set? If not, is there a plan in place to develop such a process?

DATA USE AND ASSESSMENTS

Assessment Literacy

An MSL system and its associated processes should ideally contribute to increasing [assessment and data literacy among educators](#). A lack of assessment and data literacy can lead to decreased buy-in and inconsistent implementation of your MSL system.

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

- Have you considered the current levels of assessment and data literacy within your district?
- What plans do you have in place to ensure that limited assessment literacy does not impact educators' ability and willingness to engage with the MSL system?
- If your district is using SLOs, do both evaluators/coaches and those being evaluated have the assessment and data literacy necessary to set appropriate targets and evaluate evidence of student learning? If not, how does your district plan to address this need?
- Does your MSL system enable teachers to implement cycles of inquiry and data-driven instructional practices in meaningful ways? Does participating in the MSL system provide teachers with a meaningful experience that can actually boost their overall assessment and data literacy?

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 to a post-assessment?

- How does your district plan to calculate growth between the pre- and post-assessment?
 - In making this determination, have you consulted [best practices for measuring student learning over time](#)?
 - If you plan to calculate growth centrally, have you considered the analytical capacity of your district and potential data limitations (e.g., the number of students, or n-size)?
- It is often inappropriate to use the same instrument as a pre- and post-assessment because most students can be expected to learn at least some content between them. In these circumstances, what is your district's plan to support educators in developing robust pre-assessments that measure what students should be expected to know at the beginning of instruction?

SLO Process and Supports

Many districts are using Student Learning Objectives (SLOs), where teachers set meaningful goals for learning over time based on the baseline performance of students and identify assessments to measure those goals (teachers may do so individually or use school- or district-determined assessments and targets). SLOs are attractive to districts because they are applicable across grades and content areas, although they require transparency and training/coaching to be implemented in a high-quality way.

Has your district designed the SLO system to ensure credibility with educators?

- How do your SLO system and processes reflect district priorities and values?
- How has your district connected the various components of the SLO process to everyday instructional practice?
- What training and coaching strategies is your district using to ensure that all educators have consistent experiences in implementing and being evaluated on SLOs?
- Does your district have a process through which teachers share their SLOs with one another to increase transparency and help ensure comparable levels of rigor?
- Does your district collect sufficient SLO data to evaluate the effectiveness of the SLO process overall?
- How has your district ensured that the SLO process provides educators with meaningful and actionable feedback on instruction?
- How are you gathering feedback from educators on the overall process?

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

- Has your district ensured that you have a [balanced assessment system](#), aligned to standards and adopted curricula?
- 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.)
 - o What coaching will be available to educators on an ongoing basis?
 - o Does your district have a process to vet and evaluate the defensibility and rigor of SLO targets once they have been set? If not, is there a plan in place to develop such a process?

What systems does your district have in place to make the SLO process as easy as possible for teachers and principals to implement?

- Have you considered distributive leadership models where the responsibility for training and support is not solely the responsibility of one group (e.g., principals)?
- What data collection tools have you provided to teachers and principals to help them better understand and manage the data involved in their SLOs?
- Does your district provide SLO templates to facilitate the creation of high-quality SLOs?
- Does your district provide examples to help teachers develop high-quality SLOs?
- Have you considered how to align your district's SLO process with other data-driven instructional practices, such as professional learning communities, data teams, or other processes through which educators analyze student work?



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?

- Has your district consulted [guidance from CDE](#) regarding the timing and use of state summative assessments in educator evaluations?
- What are your plans for transitioning to new statewide assessments (e.g., PARCC or SAT)?
- Has your district considered whether your measures capture growth or achievement? By focusing only on CMAS growth, are you inadvertently excluding:
 - Other state assessments (e.g., social studies or science, ACT, DLM, ACCESS)?
 - Grades (e.g., K-third grade)?
 - 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)?
 - How is your MSL system accounting for these students and assessments?
- 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 use 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.

Has your district consulted [CDE's guidance document, Using Colorado School/District Performance Frameworks in an Educator's Body of Evidence for Evaluation?](#)

- 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)? Some options include:
 - Teachers selecting specific areas of the SPF that align to their content area.
 - Schools selecting areas that align to the school or district UIP.
- 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 are defensible, it is important to be purposeful about this decision and to weigh the potential [unintended consequences](#).
- Have you considered whether other collective measures, such as UIP goals or team-based SLOs, are more aligned to the goals and values of your MSL system?

MSL PROCESS AND STAKEHOLDER ENGAGEMENT

Educator Input

District processes that adequately incorporate educators' voices and perspectives will benefit from educators' expertise and buy-in. [Data gathered 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 has your district solicited and incorporated educators' voices and perspectives in the MSL decision-making process?

- At a minimum, has your district solicited input from educators on the system (e.g., feedback survey)?
- What avenues of communication have you utilized in communicating about the MSL system to educators?
 - How have you highlighted the role of educators in MSL system creation when communicating about your process with your staff and the broader community?
- How have educators contributed in meaningful ways to the design and ongoing refinement of the MSL system in your district?
 - Have you purposefully included educators across grades and content areas?
 - If you have not yet engaged your educators in these ways, how might you include them moving forward, and how has your district worked with various stakeholders to ensure buy-in?

Transparency and Timing

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

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

- Are educators given information regarding the MSL system, measures, assessments, and targets early enough to allow them to understand expectations and plan instruction efficiently?
- Do educators have the necessary information to understand the measures used in their evaluation? Is there a communication plan in place?
- How have you made connections for educators between the MSL system and other district programs and initiatives?
- Is there transparency in the target-setting process?
- 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.](#)
- Does your district have processes for informing other stakeholders (e.g., parents) about your MSL system?
- How is your district ensuring that your processes and decisions are properly documented in order to maintain institutional knowledge and communicate consistently with stakeholders?

Continuous Improvement and System Evaluation

While it is easy to get bogged down or overwhelmed with technical considerations, including the assurance of reliability and validity in your MSL system, it is also important that your district focuses on continuous improvement and overall system evaluation. There are processes and strategies that all districts can adopt to ensure the credibility and defensibility of the MSL system.

How has your district planned for continuous MSL system improvement and adopted formal feedback mechanisms for stakeholders?

- What data and information do you collect to allow you to evaluate your MSL system and its alignment to overall educator effectiveness data, student outcomes, and the effectiveness of other district initiatives? How are you defining success with regard to your MSL system?
- Is your district engaging a diverse group of stakeholders in system evaluation and continuous improvement processes?
- How is your district communicating to your stakeholders about system changes made through continuous improvement process?

About CEI

The Colorado Education Initiative is an independent nonprofit working in partnership with the Colorado Department of Education, educators, schools, districts, and other public education stakeholders to unlock the unique potential of every student by incubating innovation, shining a spotlight on success, and investing in sustainable change that improves outcomes for all students. CEI envisions that every student in Colorado is prepared and unafraid to succeed in school, work, and life, and ready to take on the challenges of today, tomorrow, and beyond.

