



## Guidance: Measures of Student Learning in Teacher Evaluation (version 2.1)

---

*The purpose of this document is to highlight possible approaches for districts and BOCES to consider when constructing their approach to select measures of student learning for use in educator evaluations. CDE will be collecting on-going feedback to improve this guidance. [Please click here](#) to submit feedback to CDE. This guidance will be revised annually with refined versions released each summer in order to reflect new learning and emerging best practices.*

In an effort to improve the quality of education provided to all students in the state, Colorado has: adopted new academic standards that represent what students should know and be able to do at each level of their schooling; implemented school and district accountability strategies that are tied to unified improvement planning; and, adopted standards for educators who will be evaluated annually. Each of these efforts has the shared purpose of improving student learning and raising student achievement levels. It is important to recognize the interdependence of each of these strategies so that they can be implemented as parts of a cohesive and aligned system. It is also important to ensure that these strategies address how all educators in the system, individually and collectively, can contribute to the desired outcomes for Colorado students.

The focus of this guidance is on the student academic growth requirements outlined in the Senate Bill 10-191, the Great Teachers and Leaders Act. Senate Bill 10-191 requires that fifty percent of an educator's evaluation in Colorado be based on educator impact on student learning determined by using multiple measures in relationship to the Colorado Academic Standards. In Colorado, the term "academic growth" is closely associated with results from the Colorado Growth Model (CGM) as reported in the School and District Performance Frameworks. The phrase "measures of student learning" is employed throughout this document to ensure that districts understand that evaluating student learning for educator evaluations is not confined to results from the CGM, but is inclusive of results from multiple types of measures that districts may use in educator evaluation and to support instructional goals.

Measures of student learning may include student "academic growth" results as well as results from other state assessments, district assessments, and teacher-developed assessments. Multiple types of assessments may also be used to generate results from goal or target setting processes.

There are four basic requirements outlined in State Board Rules to be considered as districts design systems to incorporate the results from multiple measures of student learning into their educator evaluation systems. The first two requirements are specific to attribution (who has contributed to the results). Each educator is required to have at least one measure of student learning that is individually attributed, meaning that results are attributed to an individual educator, and another that is collective, meaning that the student results on the measure are attributed to more than one educator. The third requirement is that results from statewide summative assessments must be included, when available and appropriately connected to the subject, grade, or course. The final requirement is that, for subjects with statewide summative results in two consecutive years, results from the CGM must be included. Example: a fifth grade teacher who is responsible for reading instruction must include, in some way, results from the reading state summative. It is important to note that the third and fourth requirements may be applied either individually, or collectively as illustrated in Table 1 below:



Table 1: State Requirements

Requirement 1: Individual Attribution	Requirement 2: Collective Attribution
<p>Student results on a measure are attributed to one licensed person.</p> <p>(Example: The class performance on a history final may be used as a part of a target setting process with the results attributed to the history teacher who taught the students who took the assessment.)</p>	<p>Student results on a measure are attributed to more than one licensed person.</p> <p>(Example: The TCAP reading results in an elementary school may be attributed to all the teachers in the school as reflected in the targets section of the Unified Improvement Plan (UIP); or the results agreed upon in an SLO process for all the teachers in that grade.)</p>
<b>Requirement 3: Statewide Summative Assessment Results, when available</b>	
<p>Include when available and appropriately connected to the subject, grade and course (could be used for individual and/or collective attribution)</p>	
<b>Requirement 4: Results from Colorado Growth Model (CGM), when available</b>	
<p>Include for subjects and grade levels with annual statewide summative assessment results in two consecutive years. CGM results may also be attributed to teams of educators contributing to the growth of students in a district, school, or subject area (could be used for individual and/or collective attribution). The Department specifically recommends the use of Median Growth Percentiles for a given grade level, school, or specific content area. This data can be obtained by using SchoolView to access the school and district growth summary reports, the Colorado Growth Model Visualization Tool, Data Center, and the Data Lab.</p> <p>CDE is developing guidance for districts on considerations for using CGM results in educator evaluation.</p>	

The sections in this document consist of recommended steps for identifying and determining the measures of student learning that are included in a district’s educator evaluation system. Taken together, these steps detail a sample process that may be used by districts to determine measures of student learning in order to generate a performance rating for the measures of student learning quality standard for teachers (quality standard five) and principals (quality standard six). The steps are as follows:

- Step 1:** Begin with the Colorado Academic Standards to identify critical learning goals for students.
- Step 2:** Identify available assessments aligned to those goals that are being used in your district to evaluate student learning throughout the year.
- Step 3:** Group available assessments according to teacher types.
- Step 4:** Select and preliminarily weight assessments for use in educator evaluations.
- Step 5:** Determine how the results from the selected student learning measures will be scaled for expected growth.
- Step 6:** Combine weighted scores into a “measures of student learning” rating.

*A Note to Readers:*

As districts follow the steps outlined in this guidance, they will find that, for many assessments, districts will need to work with their teachers to establish student learning objectives (SLOs). Student learning objectives are a participatory method of setting measurable goals or objectives for a specific assignment or class in a manner aligned with the subject matter taught and in a manner that allows for the evaluation of the baseline performance of students and the measureable gain in student performance during the course of instruction (1 CCR 301-87 -1.23).

Student learning objectives are constructed through a process that enables educators to utilize academic standards to establish learning outcomes for individual or groups of students, monitor students' progress toward these outcomes and evaluate the degree to which students achieve these outcomes using relevant, meaningful measures (see CTAC, 2011; Goe & Holdheide, 2011; Marion & Buckley, 2011; Goe, 2012). CDE is currently developing resources to support districts in using SLOs in educator evaluations (see Textbox 1 below).

*Textbox 1:*

A process for creating student learning objectives enables educators to utilize academic standards to establish learning outcomes for individual or groups of students, monitor students' progress toward these outcomes and evaluate the degree to which students achieve these outcomes using relevant, meaningful measures (see CTAC, 2011; Goe & Holdheide, 2011; Marion & Buckley, 2011; Goe, 2012). The active involvement of teachers in supporting student learning is the defining feature of a process for creating SLOs. The design of the process reflects good teaching practices such as setting clear learning targets, differentiating instruction for students, monitoring students' progress toward these targets, using student learning data to adjust instruction, reflecting on professional practice, and evaluating the extent to which students have met the targets. In other words, a process for creating SLOs is an instructional improvement process driven by teachers in all grades and subjects. The essential components of a process for creating SLOs include:

- Using the Colorado Academic Standards to establish goals for student learning
- Determining what assessment instrument(s) will be used to measure student learning in relation to the learning goals at the end of the instructional period, specifying scoring approaches, and evaluating the quality of the assessment instrument(s)
- Collecting and analyzing baseline data about student learning in relationship to the learning goal(s) set for students with similar initial performance and to inform target and scale setting for groups of students
- Monitoring student learning (formative practice) throughout the instructional period
- Determining attainment of student performance targets
- Reflecting on, and refining the process for creating learning goals or targets

Because the process for creating learning goals asks teachers to start with the academic standards, to set targets based on what they know about their students, and to reflect continuously on instructional practices in order to facilitate student mastery, it connects authentically with educator evaluation systems. To support districts, CDE is currently developing a sample process as a resource (is it a process or a resource) for using student learning objective results as measures of student learning for evaluation purposes.

To ensure you are reviewing the most updated version of this document, [please click here to download from the CDE website](#). Please contact the [CDE Educator Effectiveness team](#) if you have any questions or to request support.



---

## Step 1: Begin with the Colorado Academic Standards to identify critical learning goals for students

---

Colorado has adopted new academic standards that describe what students should know and be able to do at the end of their schooling. In order to make sure that students are prepared for college and the workforce, districts can use the standards to support teachers in the identification of the learning goals in each course and grade level.

Defining learning goals, which in the simplest terms, represent a subset of the knowledge and skills that students are expected to master by taking a particular course (or courses) in school, can help districts to narrow the types of measures that are selected for use in educator evaluation. This step of defining learning goals is critical to provide clarity on the expected outcomes for students, and to understand the different ways in which students should be able to apply and demonstrate this knowledge.

The identification of learning goals starts with a simple question that can be asked at the school and classroom levels: “What do we expect our students to know and be able to do as a result of the instruction received in each (content/course/grade) that is offered in our district?” The goal of beginning with this question is to get everyone in the district thinking concretely about what it is they want to accomplish with their instruction and to think about what aspirations they have for their students. Making explicit connections to the state standards will help teachers realize and appreciate that their aspirations are consistent with state and district expectations. Only after the learning goals for courses are made explicit and are connected to standards, can the process of identifying and selecting assessments for teacher evaluations begin.



---

## Step 2: Identify available assessments that are being used in your district to evaluate student learning throughout the year

---

In Step 1, districts determined what they expected their students to know and be able to do. In Step 2, districts identify the different types of assessments that are commonly used across multiple schools/grades/courses to measure what students should know and be able to do. The purpose of this step is to help districts identify:

- Which assessments are being used to measure student learning across the district
- Grades and subjects where no formal instrument or assessment exists to evaluate student learning

In addition to the results from assessments that may be required as presented earlier in Table 1, the majority of assessments used by the districts will likely consist of commonly used or teacher-developed assessments. Teachers use a variety of informal and formal assessments to evaluate student learning in their classrooms every day. They use the information generated from these types of assessments to inform instruction, determine mastery of standards, check on the progress made by students and determine a student's level of mastery by the end of the instructional period.

Districts and schools can use the sample assessment inventory template included in the appendix, the inventory of performance data associated with [Unified Improvement Planning](#) or develop their own inventories to identify assessments that are in place across their districts.

Once the inventory has been completed, districts should use the following questions to identify the assessments that will be considered in educator evaluation:

- Which assessments are aligned to the Colorado Academic Standards?
- Which assessments best measure student progress toward district/school/course learning goals, described in Step 1?
- Which assessments are deemed by educators to provide valuable information to help inform their instructional decisions?

To further assist districts as they answer the above questions, CDE has provided an Assessment Review Tool. The Assessment Review Tool allows users to evaluate the quality of assessments based on the following quality criteria: provision of relevant tasks to students, alignment to the Colorado Academic Standards, fairness, provision of clear scoring criteria and guidelines, and accessibility of content for all learners. Districts and BOCES should also check the technical quality and rigor of the various assessments that are used to generate results that are used in educator evaluation.

Districts are advised to keep the assessment selection process simple by selecting the assessments that are aligned with school and district goals, generate results that educators use to make the greatest impact on student learning, and most importantly, are aligned with the learning goals specified in Step 1.

The identification of common assessments can then be applied for different teacher types (discussed in the next section) and to help guide policy decisions on which assessments are most appropriate for use in teacher evaluations.



## Tools/resources to assist with Step 2:

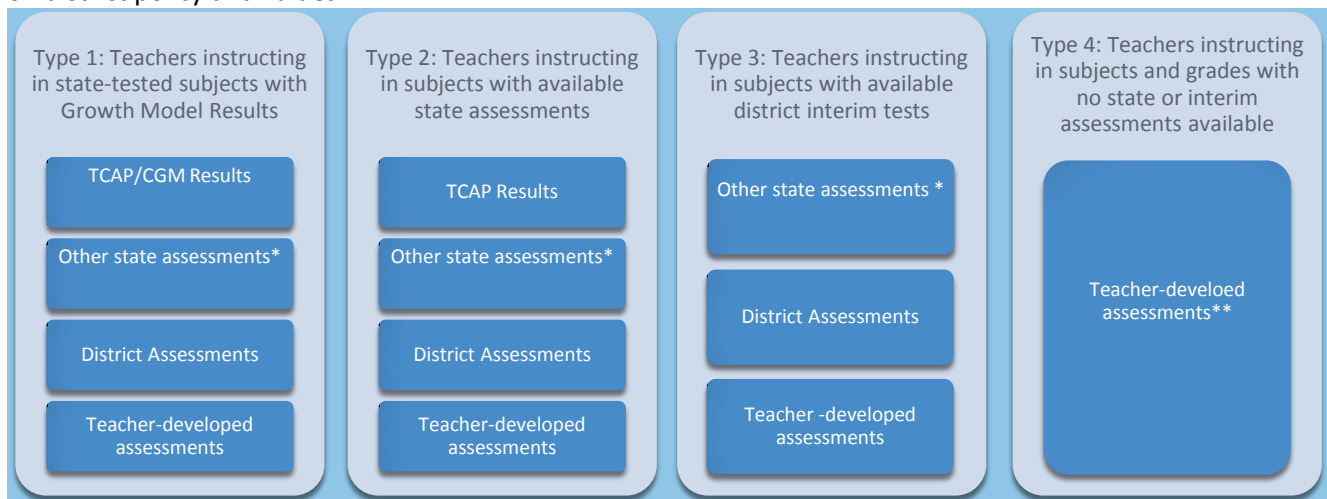
- Assessment Inventories: A sample assessment inventory is included in the Appendix. The use of these templates can help highlight limitations in the type of approaches and methods that can be used to evaluate student learning for a given content area and grade.
- A sample [Assessment Review Tool](#) used by the Content Collaboratives is included in the Appendix.
- [CDE Resource Bank](#): The CDE Resource Bank includes assessments in all grades and content areas that may provide districts with a starting point for assessing student learning and may be used for educator effectiveness purposes.

## Step 3: Group available assessments according to teacher type

In the previous step outlined, it is recommended that districts conduct an assessment inventory to identify which assessments are commonly used to evaluate student learning across grades and content areas. In Step 3, districts classify teacher types according to the types of assessments appropriate to each identified group. Information from the assessment inventory can be used by districts to identify the common types of assessments available to teachers instructing in different content areas and grades. This approach to classifying educators according to commonly available assessments is used to help set district policy for requiring certain types of assessments for specific educator types and determining which measures should be attributed individually or collectively. This categorization can also help make expectations clear to all teachers how different types of assessments will be considered in their evaluations. Figure 1 presents one example of how teacher types can be defined according to the types of assessment results available for types of teachers.

*Figure 1.* Teacher types based on types of assessments available for each group

*Note:* Results from any of the measures summarized below may be applied either individually OR collectively depending on district policy and values.



\* When available and appropriately connected to the subject, grade, or course.

\*\* Teachers in subjects or grades without state or district assessments available may contribute to the content being measured by state or district assessments and have the results included in their evaluations when appropriate.

In the above example, a district may elect to use all available TCAP results for Type 1 teachers, and also engage those teachers in creating student learning objectives that would incorporate results from the district assessments and teacher-developed assessments for evaluating expected targets set for students (identifying expected targets as a part of an SLO process will be discussed in the SLO guidance).

When considering at what level (e.g., individual or collective) to attribute results from measures identified in Figure 1, districts may decide to use selected results (e.g., results from just math and reading) for “collective-attribution” across all teacher types. For collective attribution, schools are encouraged to use the targets and/or interim measures set in their school’s Unified Improvement Plan since staff should already be familiar with the UIP targets and should be working together toward meeting those targets. Be aware that “double-dipping” of measures, for example counting math TCAP as a collective grade level measure and then counting it again as an individually-attributed measure may positively or negatively inform the overall performance rating. There are additional design considerations for collective attribution, five are listed below:



1. Identify which measures could be used to encourage partnerships or teams where teachers have an opportunity to impact growth. For example, a district may want to develop a team goal to encourage a group of teachers instructing in the same grade to emphasize the same set of learning goals in a specific content area and align those to targets and/or interim measures in the school's Unified Improvement Plan.
2. Identify which assessments could be used to help foster and support a district's focus on a specific priority area. For example, a district with a growing population of English Language Learners (ELLs) may want to use the growth results from the ACCESS assessment in the evaluation of all teachers. Using ACCESS as a "collective attribution" measure may incentivize all teachers to work collaboratively to support the growth of all English Language Learners.
3. Align measures collectively attributed to teachers with the measures used in the evaluation of principals. For example, if all teachers in a school are working together to raise fifth grade reading proficiency, the results should also be included as a measure of student learning for the principal of the school.
4. Identify appropriate results included in the School Performance Framework (SPF) for use as "collective-attribution" in the evaluation. The SPF consists of additional results besides growth, including academic achievement information and post-secondary and workforce readiness information for high schools. CDE has developed a resource to provide districts with sample approaches and considerations on the use of School and District Performance Frameworks in Educator Evaluations.
5. Identify minimum N sizes (number of data points for consideration) for individually and collectively attributing results from any measure used in educator evaluation. The Department recommends that use of results from the Colorado Growth Model only be used when there are at least 20 individual growth percentiles in the data set. Districts may explore different approaches for increasing the N size. CDE is developing guidance for districts on considerations for using CGM results in educator evaluation.

[\*\(See also – the State Council for Educator Effectiveness' Student Growth Work Group Report\)\*](#)

Other approaches that districts may consider for classifying teacher types may include grouping by: grade level, content area, and by elementary/middle/high schools. The decision to configure these groupings should be done to establish clarity with teachers on policies regarding whether the results from each measure will be attributed to them individually or collectively, and which results will apply to them.





## Step 4: Select assessments for use in educator evaluations

After taking an inventory of available assessments and determining which assessments apply to different teacher types, the next step entails narrowing down the selection of assessments in order to select those that meet quality criteria specified in the sample Assessment Review Tool included in the Appendix. In other words, any measure of student learning that is to be used in educator evaluations should be closely related to the standards being taught, curriculum, scope and sequence, and expected outcomes for a given class/course. Districts are also encouraged to consider the use of district assessments that are identified as targets and/or interim measures in the Unified Improvement Plans (UIP) as progress monitoring tools during the school year. A district decision to use targets specified in the UIP should be based on a close examination of whether those assessments are tightly aligned with course expectations and whether a good rationale can be established to use results from those assessments individually or collectively.

For example, results from a district math test may not serve as an appropriate measure for individual attribution for a social studies teacher since the test content may not have a clear relationship to the course expectations taught by the social studies teacher. To continue with this example, the results from the same math test may be considered for use as a “collective attribution” measure for the same social studies teacher if a clear argument can be made by the district that all teachers are required to incorporate some level of math practices across content areas and those practices are captured by the math test being considered. Districts are advised to select assessments that are aligned with school and district goals that generate results that educators use to inform their instruction, and most importantly, are aligned with the learning goals specified in Step 1. State Board rules require that, when available, results from statewide assessment be included in educator evaluations. See Textbox 2 for important information about the timing of these results and sample approaches for their use.

### *Textbox 2: An important note about the timing and release of results*

The timing of results from state summative assessments is an important consideration for use in educator evaluations. For example, current year state summative results are not available before the end of the current school year. Since it is required that results from statewide summative assessments be included in educator evaluations, it is suggested districts select an approach for including the results that makes sense to their educators. One approach is to use prior year statewide summative results in current year calculations. Results from other measures, such as locally-developed interim assessments, may generate results that are available at the end of each instructional period and that are directly related to the group of students that the educator taught in the current year. This means that the measures of student learning portion of an educator’s evaluation will likely consist of both prior year and current year data. It is important to understand this in order to weight each measure so that there is relevance for the educator and so that results from the prior year are not weighted such that a higher rating negates local measures, or that a low rating prohibits an educator from overcoming it with local measures.

Other approaches to address the timing of statewide summative assessment results are to wait to finalize a rating until the summative data can be incorporated into the evaluation or to use rolling averages. CDE will be exploring a range of approaches with districts to provide further guidance on how best to address the timing of assessment results and their incorporation into educators’ evaluations.



---

## Step 5: Select assessments for use in educator evaluations

---

In the previous steps, districts would have: identified student learning goals; conducted assessment inventories to find assessments that teachers can use to measure student learning; classified educators into groups to determine the set of common assessments available for different teacher types; identified which assessments and instruments will be reported at the teacher or other aggregated levels; and, narrowed down the specific assessments and instruments that apply to individual teachers. In this step, the results from selected measures are to be scaled for expected growth. For example, any results that are used in evaluations need to be evaluated and scaled according to the learning goals or targets that are associated with them. The district will have to determine what constitutes results that are much less than expected, less than expected, expected, and more than expected based on where students or groups of students began, in order to be included as a measure of student learning. However, the general standard is that students at least make a year's worth of growth in a year's time (or for the Colorado Growth Model, they make "typical growth").<sup>1</sup>

Districts are encouraged to collaborate with stakeholder groups to construct the scoring rules, determine the weights assigned for the results specified in the system, and identify the method used to combine the weighted results. Ideally, decisions about the scoring rules, weights assigned, and resulting outcomes from the method applied to combine results should be informed and refined by impact data.

Because the 2013-14 school year is the first year that these requirements are to be implemented, districts should consider the following and plan to make amendments based on lessons learned in subsequent years:

1. Consider the use of Colorado Growth Model results as reported on the School Performance Frameworks at each school and apply these results as "collective-attribution" measures for all teacher types.
  - a. The Department specifically recommends the use of Median Growth Percentiles for a given grade level, school, or specific content area. This data can be obtained by using SchoolView to access the school and district growth summary reports, the Colorado Growth Model Visualization Tool, Data Center, and the Data Lab.
2. Encourage the use of results from statewide summative assessments, district assessments, and teacher-developed assessments in developing student learning objectives as included measures for all teacher types (see Figure 1).

An approach for evaluating results for the upcoming school year is as follows:

### **Including Colorado Growth Model results collectively in educator evaluation:**

1. Depending on district size and school size, districts will want to choose an approach to using growth model results. For school level collective attribution districts may choose to use the median growth percentile as reported on the School Performance Frameworks for each available content area (reading, math, and writing). Districts may also choose to use the Median Growth Percentiles for disaggregated groups of students within a school that are also included in the SPF. Educators may have one MGP result included, or all three. Note: MGPs are not to be combined into a composite MGP for multiple content areas; they are to be treated as separate measures. Table 2 presents the ratings and scores associated with the MGP ranges defined in the SPFs. Districts may also access MGP data by using SchoolView to access the school and district growth summary reports, the Colorado Growth Model Visualization Tool, Data Center, and the Data Lab.

---

<sup>1</sup> CDE understands that this specific goal is very difficult to translate into a particular measure or target, we include it here, simply to underscore the idea that it reflects around expectations of student learning over a school year.



- Note that the SPF can include state summative growth results for content areas assessed in consecutive years depending on the size of the school (reading, writing, and math). Schools may also have growth results from the ACCESS assessment. If a school does not have any growth scores reported on the SPF, the district may want to initially assign the district score from the district performance framework (DPF) for each content area to each teacher in the school for the 2013-2014 year.

Table 2: Determining a rating using results from the Colorado Growth Model. (CDE’s State Model will use a 0-3 point scale as illustrated in the second row of the table below.)

Ratings for Growth Model Results	Much less than expected	Less than expected	Expected (Typical)	More than expected
Score Assigned	0	1	2	3
Ranges from the SPF when including Median Growth Percentiles	1 <sup>st</sup> to 34 <sup>th</sup> percentile	35 <sup>th</sup> to 49 <sup>th</sup> percentile	50 <sup>th</sup> to 64 <sup>th</sup> percentile	65 <sup>th</sup> to 99 <sup>th</sup> percentile
Example	The School MGP for the students on the Reading TCAP was between 1 and 34	The School MGP for the students on the Reading TCAP was between 35 and 49	The School MGP for the students on the Reading TCAP was between 50 and 64	The School MGP for the students on the Reading TCAP was between 65 and 99
NOTE: The cut points shown in this example are for use when evaluating Median Growth Percentiles.				

**Including results from other measures**

- In selecting multiple measures for use in educator evaluation, districts can work with their educators to set targets for student learning. (CDE’s State Model will use a 0-3 point scale for differentiating targets illustrated in row two of Table 3 below.) At the end of the course, or evaluation cycle, districts will have to evaluate the degree to which the targets set were met.
- Districts may establish processes for educators to use the results on the selected measures to determine expected targets for different groups of students in their classroom(s) at the beginning of the class/course/grade. (Guidance on setting expected targets will be addressed in the, soon to come, student learning objectives process document.)
- Student performance will be evaluated relative to the expected targets set for each of the measures included. Based on the rubric criteria in the sample below, teachers can earn a possible rating of 0-3 on each of the measures of student learning that are included in their evaluation.



Table 3: Rubric for rating the results on selected measures or targets

Target Evaluation Rating Rubric				
Ratings for Results on selected measures or targets.	Much less than expected student performance	Less than expected student performance	Expected student performance	More than expected student performance
Score Assigned	0	1	2	3
SAMPLE Criteria*	Less than 64% of students defined in the SLO meet expected targets set	65-74% of students defined in the SLO meet expected targets set	75-84% of students defined in the SLO meet expected targets set	Greater than 85% of students defined in the SLO meet expected targets set

*\*The sample criteria in Table 3 is to illustrate how targets may be set based on the learning targets and local context within a district, school, or classroom.*

CDE has identified a rating scale that has four categories, much less than expected, less than expected, expected, and more than expected for use in the Model System. There are points associated with each rating as shown in row three of Table 3. When the points from each measure are combined, a composite score is established for each educator based on a combination of scores from each measure that is included in their body of evidence. Combining scores will be described in the next section.



## Step 6: Combine weighted scores into a “measures of student learning rating”

By assigning weights to each score associated with the multiple measures in educator evaluations, districts are signaling which results or measures in the system are deemed to have more value than others, are better aligned with learning goals, are more appropriate for measuring educator impact, or may signal that all results should be weighted equally. After each of the measures of student learning are scaled (e.g., on a 0-3 scale), the next step would entail assigning weights to each and applying an approach to calculate a total score earned by teachers on measures of student learning. Districts may wish to preliminarily weight the results from each measure as it is selected at the beginning of the school year. Districts are encouraged to continuously evaluate the impact of weighting decisions and make revisions as needed.

For example in terms of weighting considerations, districts may want to assign more weight to:

- Outcomes from measures deemed to be of higher technical quality;
- Outcomes reflecting collective efforts from a team of teachers (note that the statute and rules do not specify a minimum weight for either individual or collective attribution measures but do suggest that each must have a “measurable influence”); or,
- Outcomes from measures deemed by district stakeholders to have higher value for teachers.

Although districts can decide how to weight the scores from each of the multiple measures, districts may want to keep things simple by selecting weighting percentages that sum up to 100 percent. Multiplying the scores earned by the assigned weight yields the weighted score for each measure. The composite score in this example represents a compensatory approach, which was selected as a design choice to ensure that each measure included in an educator’s body of evidence can have a measureable influence on the student learning score. Table 4 provides an illustration of how districts *may* consider distributing the weights assigned to each score for their teachers, and how a single index score is computed.

Table 4: Weighting and combining scores example (refer to Tables 2 and 3 for possible scores)

Measures/Results from Colorado Growth Model and Student Learning Objectives (SLO)	Score Earned	Weight Assigned	Weighted Score
TCAP Reading MGP (collective school)	2 (typical)	.15	.3
TCAP Writing MGP (collective school)	2 (typical)	.15	.3
SLO 1 Results (collective grade level reading)	2	.35	.70
SLO 2 Results (individual teacher)	1	.35	.35
<b>Sum of Weights</b>		<b>1</b>	<b>1.65</b>

In Table 4, the assumption is made that the district has agreed to attribute Colorado Growth Model results from reading and writing (total of 6 points possible) to all teachers in the school. Further, Table 4 illustrates that all teachers will have two additional measures based on targets yielding 2 scores (total of 6 points possible) for attainment of expected targets. The first column is the measure that is included. The second column reflects the rating earned -- much less than expected (0 points), less than expected (1 point), expected (2 points), and more than expected (3 points) -- by a hypothetical teacher with all these measures relevant to her goals. To assign weights to scores, a district can allocate smaller or higher percentages to each rating and ensure that the weights assigned across all measures sum up to 1 or a 100 percent as shown in the third column. In this example, the district has decided that each of the results from their SLO targets and the set of combined TCAP growth results should have about the same weight. The third column shows that each SLO result has a weight of .35 and the set of combined TCAP growth scores



has a total weight of .30. The fourth column in the table shows the weighted scores. These are computed by multiplying the score earned for each measure (column 2) by the assigned weight (column 3). Table 4 illustrates how multiple measures are combined for an educator. In Column 4, it is determined that the raw score from measures of student learning for this educator is 1.65. Districts will combine the weighted scores in order to determine a “*measures of student learning*” rating for each teacher by adding the scores and applying them to a new scale as shown in Table 5 below. Apply the weighted score from Table 4 to the measures of student learning scale in Table 5 to determine the educator’s rating on the measures of student learning quality standard. From Table 5, we find that the teacher who earned a total score of 1.65 earned an *expected student learning score*, meaning that during the year, students in the teacher’s classes generally met expectations as a result of her instruction.

In the State Model evaluation system each educator also earns a professional practices rating based on their performance on the state model rubric (combining the other 5 Quality Standards). In order to be able to combine the ratings for the professional practice and student learning, the same scale must be used. Because of this, the State Model converts both the professional practice rating and the student learning rating into a score between 0 and 540. The number 540 is based on the number of elements included in the professional practices quality standards for teachers. In order to combine the professional practice score with the measures of student learning score so that each score represents 50% of an educator’s evaluation the measures of student learning score will be translated into an index score that can be translated to a 540 scale as described below. More information about how to combine the professional practices and the student learning score can be found in [“Determining a Final Educator Effectiveness Rating for Teachers”](#)

*Calculating a measures of student learning score*

The sum of all weighted scores in Table 4 (above) represents the composite student learning score earned by a teacher. Table 5 translates the composite score range into qualitative judgments about student learning for a given teacher. The cut points in Table 5 for raw composite scores are based on scores of 0 for *Much less than expected*, 1 for *Less than expected*, 2 for *Expected*, and 3 for *More than expected*. When numbers in the four ranges in this table are combined and rounded to the nearest whole number, they are placed in the four categories as shown. The fractions are produced when multiple assessment teacher evaluation scores are weighted and averaged together. For example, if a teacher has earned a 1, 2, and 3 from three different assessments, and if they are weighted 20%, 30%, and 50% respectively for the 100% of the measures of student learning side of the pie, then we can compute the weighted average as:

$$(1 * .20) + (2 * .30) + (3 * .50) = 2.3, \text{ which is in the range for } \textit{Expected student learning}.$$

Table 5. Cut points for composite measures of student learning scores

Composite Rating	Much less than expected	Less than expected	Expected	More than expected
Total RAW Composite Score ranges (0-3)	0.0 to 0.49	0.50 to 1.49	1.50 to 2.49	2.50 to 3.0

In Figure 2 the raw composite score of 2.3 is converted to a ‘measures of student learning’ score between 0 and 540. The measure of student learning score will be added to an educator’s professional practices score in order to determine an overall effectiveness rating.

Figure 2: Illustration of calculating a student learning score.

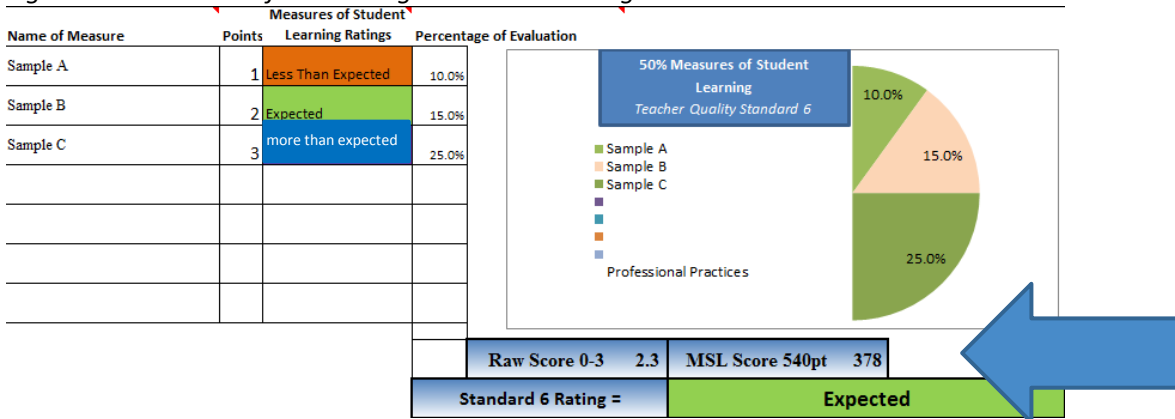


Table 6 describes the method for converting the measures of student learning raw composite score into a measure of student learning score. Note, all CDE tools will do this math for users.

Table 6. Rules for converting a Measure of Student Learning raw score to the 540 point scale

Measures of Student Learning Raw Composite Score	Formula for computing a Measures of Learning Score
much less than expected growth ( $0 < \text{score} < .5$ )	$(\text{score} - .0) * 270$
less than expected growth ( $.5 \leq \text{score} < 1.5$ )	$(\text{score} - .5) * 135 + 135$
expected growth ( $1.5 \leq \text{score} < 2.5$ )	$(\text{score} - 1.5) * 135 + 270$
more than expected growth ( $2.5 \leq \text{score} \leq 3.0$ )	$(\text{score} - 2.5) * 270 + 405$

Using our example of 2.3 above as the weighted average of three assessment scores, we can use Table 6 to convert 2.3 to the 540 scale with the *expected growth* formula:

$(2.3 - 1.5) * 135 + 270 = 378$ , which would be our final Measure of Student Learning rating for this teacher.

*Weighting considerations*

Note that an overly high weight or percentage attributed to collective attribution measures may decrease the ability at the school or district level to recognize high-performing teachers (who may be held back by the average) and to identify struggling teachers (who may be “propped up” by the average). Therefore, it is imperative that districts understand the importance of finding the right balance between weighting the measures that reflect teacher-level and collective attribution results.

To help districts visualize the impact that weighting has on the overall student learning score in the educator evaluation, CDE has developed in Microsoft Excel a Measures of Student Learning: Weighting and Scoring tool (link below) for use by personnel evaluation committees and educators. Districts can use the tool to explore the impact varied weights assigned to different measures can have on an overall rating and use the tool to get feedback on weighting decisions.

Although this guidance includes approaches that are based on compensatory approaches that allow for strong performance on some measures to compensate for weaker performance for others, districts may consider other approaches. For example, a district may want to use a conjunctive approach that requires a minimum threshold to be earned across all measures before assigning a “passing” or “meeting” rating for teachers on student outcomes. The value statement articulated by a district electing to use a conjunctive approach is that they believe teachers should demonstrate a minimum level of proficiency on each measure being considered prior to earning a rating of “meeting



expected student learning.” Using a conjunctive methodology would indicate high confidence in each measure’s technical validity and appropriateness for attributing the results to an individual teacher.

### Tools/resources for completing Step 6:

- [Measures of Student Learning Tool](#): This Microsoft Excel tool is one sample approach designed to help Colorado educators input the measures that will be used in their evaluations, see the impact of the weighting decisions for each measure, input the desired learning targets that are expected as a result of their instruction, and synthesize the evidence from multiple measures into one score that will be used in educator evaluation. It includes the requirements included in S.B. 10-191, the rationale for decisions made, and creates sample graphics for various groups of teachers.





---

## Summary

---

The purpose of this document is to illuminate a step by step process for districts to consider as they determine their approaches and design considerations for evaluating measures of student learning for use in educator evaluations. As districts develop their approach, they will want to keep the following principles in mind:

- The process begins with districts identifying key learning goals in each content area and grade. Identifying these learning goals serves as the foundation for selecting the set of appropriate assessments and measures to use in teacher evaluations and is consistent with the intended vision that all measures used in an educator evaluation system provides meaningful and useful information for educators.
- The guiding principles and values for selecting assessments and selecting an approach and method for combining measures are made transparent and clear to all stakeholders.
- Educator evaluator systems employed by districts are continuously monitored and improved based on impact data.

There are five additional documents for all districts to review that relate to the design of the educator evaluator system:

1. [User's Guide for Evaluating Colorado's Educators with the State Model Rubric](#) (50 percent professional practice)
2. [Determining a Final Effectiveness Rating for Teachers](#) (combining the 50 percent professional practice with the 50 percent measures of student learning)
3. Student learning objectives process, a Colorado approach (coming soon)
4. [Use of the SPF/DPF in Educator Evaluations](#)
5. Guidance for districts on considerations for using Colorado Growth Model results in educator evaluation. (coming soon)