Measures of Student Learning:
Local Examples and Analysis from the Field

AKNOWLEDGEMENTS

The contributions of the various districts/BOCES that provided example measures of student learning and/or participated in providing feedback are gratefully acknowledged. These educators willingly shared their time and expertise to examine and think critically about sample measures. Their insights are captured and summarized here in order to benefit educators throughout the state of Colorado. District and BOCES members whose contributions led to the development of this document include:

Adams 12 Five Star Schools
Adams County School District 14
Adams County School District 50
Aurora Public Schools
Center Consolidated School District 26JT
Cherry Creek School District
Colorado Springs School District 11
Denver Public Schools
Douglas County School District
East Central BOCES
Falcon School District 49
Fountain-Fort Carson School District 8
Harrison School District 2
Jefferson County Public Schools
Littleton Public Schools
Mesa County Valley School District 51
Mountain BOCES
North Conejos School District
Park County School District Re-2
Salida School District R32J
Santa Fe Trails BOCES
Sargent School District
St. Vrain Valley Schools
Steamboat Springs School District
Stratton School District R-4
Thompson School District
Weld County School District Re-8
Widefield School District 3

Introduction

The purpose of this document is to highlight examples of specific measures used by districts/BOCES for the Measures of Student Learning (MSL) portion of educator evaluations. Per Senate Bill 10-191, the MSL portion shall comprise fifty percent of educator evaluations. Within this half of the educator evaluations, districts/BOCES have the flexibility to determine what measures will be used and how these measures will be weighted as long as certain state requirements are met. These requirements along with other helpful information can be found in the Measures of Student Learning Guidance.

The Educator Effectiveness Unit at CDE is providing this document in response to feedback from the field that example measures of student learning would be beneficial as districts/BOCES continue to revisit and modify their MSL systems. *This document does not reflect an endorsement of any particular measure by the EE Unit.*

Measure Collection and Feedback

The specific measures of student learning shared in this document were collected from participants in the January 2016 Educator Effectiveness (EE) Leaders' Retreat. Participants were asked to submit at least one measure that is utilized within their district and/or school to be shared and reviewed by the rest of the participants. A template form was provided to participating districts/BOCES to complete when making their MSL submission (see Addendum A).

Retreat participants were divided into three groups (for the purpose of this document they will be referred to as Group 1, Group 2, and Group 3) and were asked to discuss each measure and provide feedback on the measures. Each group reviewed ten measures; however, not every measure was reviewed by every group. A total of seventeen measures were reviewed. Feedback consisted of comments on what the participants liked, what they had questions about and suggestions for improvement. A blank feedback form is shown in Addendum B.

After participants had the opportunity to review each measure, they were asked to vote individually on the *likelihood they would use* each measure in their district/BOCES. Each participant was given nine opportunities to vote and could use all of their votes on one measure or spread them out over multiple measures. Participants were asked to use their district/BOCES values as a guide when voting for measures.

Document Use

In this document, measures are organized into four types based on the assessments used:
- School Performance Measures, pages 5-7;
- Vendor-Based Assessments, pages 8-13;
- Student Learning Objectives (SLO), pages 14-22; and
- Specialized Service Professionals (SSP), pages 23-26.

The feedback from participants has been restructured into three categories: Questions to Consider, Strengths, and Limitations. Specific feedback was consolidated and summarized by CDE staff for the sake of brevity and clarity.

In this document, the tabulation of participants' votes is shown in the Voting Reflections section after each measure. As indicated previously, each participant was given nine opportunities to vote for the *likelihood they would use* each measure in their district/BOCES. These votes were then tallied and each measure was placed into one of four categories based on the total number of votes that was received by each group. The following four categories were established for simplicity and to demonstrate participants' interest in each measure:

1) Low Interest (indicates a measure received 0-4 votes);
2) Medium Interest (indicates a measure received 5-9 votes);
3) High Interest (indicates a measure received 10-14 votes); and
4) Very High Interest (indicates a measure received 15 or more votes)

It is important to note that these categories are not intended to be an indicator of the quality of each measure, but is a reflection of the values individual participants have identified as guiding their district/BOCES decisions when determining the use of measures for their MSL systems.

School Performance Measures

In this section, two examples of School Performance Measures are provided. Generally, these measures combine multiple performance indicators to arrive at a school-level aggregate of the outcomes of all students in the school. Examples of these measures include, but are not limited to, the School Performance Framework (SPF) and the District Performance Framework (DPF) which are both calculated by CDE. Frequently, these measures are used as collective measures as they are interpreted to be the result of the combined contributions of most or all educators in a school and/or district.

School Performance Measure – Example 1

Measure Name	School Performance Framework (SPF)
Educator Type	Teacher
Attribution (Individual or Collective)	Collective
Weight (Out of 100%)	40%

Description: SPF

Measure Type: SPF/DPF

<table>
<tr><th colspan="4">Success Criteria</th></tr>
<tr><th>Much Less Than Expected</th><th>Less Than Expected</th><th>Expected</th><th>More Than Expected</th></tr>
<tr><td>0-37.4%</td><td>37.5-62.4%</td><td>62.5 - 87.4%</td><td>87.5 - 100%</td></tr>
</table>

Questions to Consider

- Is there consideration for the fact that this data is a year in arrears?
- Are all teachers using the SPF measure regardless of hire date? What is being used for new teachers?
- Is 40% attribution too high? What are the unintended consequences (e.g., masking)?
- Due to variability in school performance, is there a concern regarding equity with a measure greater than 10%?
- Are success criteria based on overall percent of points earned?
- Would using portions of the SPF (achievement, growth, growth gaps) be better than using the overall SPF points?
- Is this sustainable from year to year?

Strengths and Limitations

Strengths:
- Focuses all instructional staff on improvement of school accountability system (so they will all strive toward school improvement).
- It is simple, clear and meets the guidelines for a collective measure.
- It is based on points earned and not plan type (e.g., turnaround, performance)

Limitations:
- This data is not meaningful when the accountability clock is frozen and data is two years old.
- High weight could benefit ineffective or partially effective teachers (but would instill a collective effort amongst peers).
- This does not allow for targeting of specific sections of the SPF (i.e., achievement, growth, growth gaps).

Voting Reflections

This measure was reviewed by all three groups at the EE leaders' retreat. Based on the sum of votes from each group, one group showed medium interest in this measure and two groups showed low interest in this measure (**Group 1:** Low Interest; **Group 2:** Medium Interest; **Group 3:** Low Interest).

School Performance Measure – Example 2

Measure Name	School Composite Score
Educator Type	Teacher
Attribution (Individual or Collective)	Collective
Weight (Out of 100%)	49%

Description: The School Composite Score is the collective attribution component of Standard VI for our district for both teachers and Specialized Service Professionals (SSP). It will serve as Standard VII for principals. There are a total of 100 possible points, with 12. 5 additional points available as part of the impact factor, for a maximum of 112.5 possible points. It is made up of points from the District DPF (10 points possible) the School's SPF (40 points possible), and points earned for the Median Growth Percentile (MGP) of Students of Color and the MGP of Below Proficient Students on state standardized assessments (25 points possible for each). Schools also get a maximum of 12.5 additional points based on the mobility of their population and the Free and Reduced-price Lunch (FRL) rate relative to the other schools in the district. This measure was developed for use with TCAP scores and will be modified to be used with PARCC data.

Measure Type: District Custom

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
0 to 42.6 points	42.7 to 60.3 points	60.4 to 73.8 points	73.9 to 112.5 points

Questions to Consider

- Why such a high percentage of weight for this measure? What are the unintended consequences (e.g., masking)?
- How does this align with SPF as a measure of school effectiveness?
- What do you do with SPF during a frozen clock when it is 2 years out of date?
- From where does the 12.5 bonus points come (why not something easier like 10)?
- How will this change in future years?
- Are the additional points for mobility + FRL all or nothing or a percentage? Does this send the wrong message? (i.e., a "handicap" for highly impacted schools)?
- What is the impact on teachers who are not tested/included in the SPF? How is this monitored?

Strengths and Limitations

Strengths:
- The inclusion of many factors.
- Factors focus on equity. The playing field is being leveled for various factors (e.g., FRL & mobility).
- This MSL can be used for both teacher and principal evaluations.
- The point system is clear and simple.

Limitations:
- Need for strong communication plan. Do teachers and principals really understand on what they are being evaluated and how points are determined?
- May be difficult to use PARCC data if opt-out is high.
- May not be applicable for new teachers who did not contribute to the data.

Voting Reflections

This measure was reviewed by all three groups at the EE leaders' retreat. Based on the sum of votes from each group, two groups showed medium interest in this measure and one group showed very high interest in this measure (**Group 1**: Medium Interest; **Group 2:** Very High Interest; **Group 3:** Medium Interest).

Vendor-Based Assessments

In this section, five examples of Vendor-Based Assessment measures are provided. Vendor-based assessments are assessments that are constructed by third parties outside of a district. They are often used as common assessments across grades and schools. Examples may include NWEA MAPS, STAR, DIBELS, and Galileo.

Vendor-Based Assessments – Example 1

Measure Name	Galileo Pre Math Test-grade level
Educator Type	Teacher
Attribution (Individual or Collective)	Individual
Weight (Out of 100%)	30%

Description: Galileo is a grade level standards test, measuring all standards for Math

Measure Type: District Custom/District Pre Test

Content Area/Other: Math grade level standards

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
Less than 55% of total students show growth	55%-69% of total students show growth	70%-84% of total students show growth	Greater than 84% of total students show growth

Questions to Consider

- What defines "growth?" Any, 1 year, or other?
- Is a post-test given? How else would growth be determined?
- How can this test be used to progress-monitor students? Is there any benchmark assessment given along the way?
- Is this measure better suited for specific types of schools (i.e. low performing versus high performing)?
- To what "standards" is this measure aligned? Colorado Academic Standards?

Strengths and Limitations

Strengths:
- This measure is used in all grade levels.
- It is a common assessment based on standards.
- Very clear.

Limitations:
- Not clear on how "growth" is defined.
- It is unclear to what standards this is referring.

Voting Reflections

This measure was reviewed by one group at the EE leaders' retreat. Based on the sum of votes for this group, they showed low interest in this measure (**Group 1**: Low Interest).

Vendor-Based Assessments – Example 2

Measure Name	MAPS Reading
Educator Type	Teacher
Attribution (Individual or Collective)	Individual
Weight (Out of 100%)	40%

Description: None

Measure Type: Vendor-based

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
<40% of students will make a 16 point growth from fall to spring.	40-69% of students will make a 16 point growth from fall to spring.	70-90% of students will make a 16 point growth from fall to spring (one year's growth)	>90% of students will make a 16 point growth from fall to spring.

Questions to Consider

- Is Depth-of-Knowledge addressed?
- What specific skills are being tested/measured?
- Is this referring to RIT score? Is 16 points attainable/typical growth?
- How does this accommodate for special populations (SPED, GT, ELL)?
- How is instruction being monitored in the interim (between the two tests)? Progress monitoring?

Strengths and Limitations

Strengths:
- MAPS is universally recognized and data growth targets are established.
- Data is already available (so it is not something that is added on top of other things).
- It is clear and simple.

Limitations:
- This does not specify grade.
- This does not align goals to the baseline performance of individual students.

Voting Reflections

This measure was reviewed by one group at the EE leaders' retreat. Based on the sum of votes for this group, they showed a very high interest in this measure (**Group 3**: Very High Interest).

Vendor-Based Assessments – Example 3

Measure Name	NWEA Growth
Educator Type	Teacher
Attribution (Individual or Collective)	
Weight (Out of 100%)	Varies

Description: Typical growth plus 10%

Measure Type: School Custom

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
30%	60%	80%	95%

Questions to Consider

- Is this used in conjunction with criterion-referenced assessments? Or other performance assessments?
- What does "typical growth plus 10%" mean? Is this about percent of students who met their Fall to Spring growth targets?
- High School NWEA RIT targets defined by CDE are all within the standard error of the baseline score (meaning they could indicate no growth). How are you accounting for this?
- Is this a RIT score increase or a percentile change?
- How was the percent of students chosen for the success criteria? What was the baseline?
- Why are success criteria not written as ranges? This would make it more clear (for example, is "Much Less Than Expected" 30-59%? If so, where is 0-29% captured?).

Strengths and Limitations

Strengths:
- The assessment spans multiple grade levels.
- This is a nationally-normed assessment given every year and has prior year data on students.
- The percent for more than expected (95%) is a high bar.
- It is a growth-based measure.

Limitations:
- A more involved description is needed.
- Do teachers understand the measurement and attainable growth targets?
- There is no reward for teachers who have a high percent of students who score more than one year growth.
- The targets may be really high. The national norm is that 50% meet target; so on average a school is beating the odds to have anything above 50. Meaning on average all your schools will be at "Less Than Expected".

Voting Reflections

This measure was reviewed by two groups at the EE leaders' retreat. Based on the sum of votes from each group, both groups showed low interest in this measure (**Group 1**: Low Interest; **Group 2:** Low Interest).

Vendor-Based Assessments – Example 4

Measure Name	Math/Regrouping
Educator Type	Teacher
Attribution (Individual or Collective)	Individual
Weight (Out of 100%)	50%

Description: 80% of students will demonstrate mastery of regrouping as measured by a task assessment. 3rd graders will demonstrate mastery of regrouping on 3 digit addition tasks and 4th graders will demonstrate mastery on multi-digit regrouping tasks. Mastery is determined as scoring 8 out of 10 on calculations involving regrouping.

Measure Type: Vendor-based

Content Area/Other: Math

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
0-39% of 3rd grade students and 4th grade students will demonstrate mastery of regrouping appropriate for their grade level as measured by the vendor-based assessment.	40-74% of 3rd grade students and 4th grade students will demonstrate mastery of regrouping appropriate for their grade level as measured by the vendor-based assessment.	75-84% of 3rd grade students and 4th grade students will demonstrate mastery of regrouping appropriate for their grade level as measured by the vendor-based assessment.	85% or more of 3rd grade students and 4th grade students will demonstrate mastery of regrouping appropriate for their grade level as measured by the vendor-based assessment.

Questions to Consider

- How were ranges determined?
- Is this for an entire year? What is the time frame?
- Are performance targets differentiated? Tiered?
- What about other standards?
- How can this be expanded to more than one assessment of learning? Or to represent year-long achievement/growth?

Strengths and Limitations

Strengths:
- Teachers set success criteria so there is buy-in from educators for whom this measure applies.
- Simple and easy to understand.
- Specific and measurable.

Limitations:
- Weight may be too high for how narrowly focused the skill is among the whole standard for this grade level/subject.

Voting Reflections

This measure was reviewed by one group at the EE leaders' retreat. Based on the sum of votes for this group, they showed very high interest in this measure (**Group 3**: Very High Interest).

Vendor-Based Assessments – Example 5

Measure Name	Scantron Performance Peer Group Growth
Educator Type	Teacher
Attribution (Individual or Collective)	Individual
Weight (Out of 100%)	15%

Description: Students are placed in peer group based on the beginning of year Scantron Performance Series quartiles. Their growth from fall to winter is determined in both reading and math. They are then compared to the median of their peer group.

Measure Type: Vendor-based

Content Area/Other: Reading and Math

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
Less than 35% of student scale score differences are at or above the median of their academic peers	At least 35% of student scale score differences are at or above the median of their academic peers	At least 49% of student scale score differences are at or above the median of their academic peers	At least 62% of student scale score differences are at or above the median of their academic peers

Questions to Consider

- Is the students' performance also compared to grade level standards?
- How do you account for error? How do you know students were placed in the right peer group at the start?
- Is median peer group within the school, district or nation?
- Which standards are being taught?
- Are you looking for growth or growth toward competency?

Strengths and Limitations

Strengths:
- Description is clear and easy to understand.
- Fall to winter comparison is clearly stated.
- It can be used to identify intervention groups.

Limitations:
- Success criteria are not easily understandable (median of their peer group?).
- Ranges are not clearly defined (0-35%, 36-48%, etc.)

Voting Reflections

This measure was reviewed by one group at the EE leaders' retreat. Based on the sum of votes for this group, they showed medium interest in this measure (**Group 1**: Medium Interest).

Student Learning Objectives (SLO)

In this section, seven examples of Student Learning Objective (SLO) measures are provided. Defining features of the SLO process include active collaboration between the evaluator and educator in setting clear success criteria, differentiating student goals, and evaluating the extent to which student have met targets. Student Learning Objective measures are commonly attributed individually but may also be attributed in a collective manner, especially for teachers working in small teams (e.g., grade level or PLCs). It is important to note, however, that other measures such as state assessments and vendor-based assessments, may also be classified as Student Learning Objective measures if the evaluator and educator are defining goals collaboratively.

Student Learning Objectives (SLO) – Example 1

Measure Name	SLO - Citing Textual Evidence
Educator Type	Teacher
Attribution (Individual or Collective)	Individual
Weight (Out of 100%)	33%

Description: Students will find appropriate quotes to justify their answers, correctly write and cite their quotes, and explain how the quote connects to their response.

Measure Type: Teacher Custom/IB Rubric

Content Area/Other: 7th grade language arts

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
0-40% of my students will increase their score on the IB rubric by 2 points, or will move to an advanced level (6-7), or stay the same at 7-8 level.	41 - 74% of my students will increase their score on the IB rubric by 2 points, or will move to an advanced level (6-7), or stay the same at 7-8 level.	75 - 90% of my students will increase their score on the IB rubric by 2 points, or will move to an advanced level (6-7), or stay the same at 7-8 level.	91-100% of my students will increase their score on the IB rubric by 2 points, or will move to an advanced level (6-7), or stay the same at 7-8 level.

Questions to Consider

- How were the cut scores determined? What was the justification?
- Are these aligned to IB expectations? Core content standards?
- Is this a one-time assessment or an average over time? Is this broad enough or just a stepping stone to a bigger learning outcome?
- What criteria are used to determine if a quote is "appropriate"?
- How many performances are students doing to determine success? All in the same context or different contexts?
- What is the timeframe of the learning cycle?
- Is this aligned with the Colorado Academic Standards?

Strengths and Limitations

Strengths:
- It is an authentic assessment that can be used as a formative assessment that enables the educator to make real-time adjustments in instruction.
- Students in the "more than expected" category can continue to advance.
- Success criteria are given in specific ranges and levels.
- Good details, easily understandable.
- Fits/aligns to classroom instructional practice.
- There are multiple ways to meet success criteria (increase 2 pts, move to advanced, stay at high level).

Limitations:
- Non-IB educators need clarification on point system and levels.
- There is a disconnect between the description and success criteria with reference to "IB". More information is needed.

Voting Reflections

This measure was reviewed by two groups at the EE leaders' retreat. Based on the sum of votes from each group, one group showed medium interest in this measure and the other group showed high interest in this measure (**Group 1**: Medium Interest; **Group 2:** High Interest).

Student Learning Objectives (SLO) – Example 2

Measure Name	Student Learning Objective
Educator Type	Teacher
Attribution (Individual or Collective)	Individual
Weight (Out of 100%)	49%

Description: Students will independently write a proficient 5th grade informative text to examine a topic and convey ideas and information clearly.

Measure Type: School Custom; 5th grade informational/expository writing rubric

Content Area/Other: Writing

Additional Info: Pre-assessment: Student will write an informative text about Silver Creek Elementary(SCE). Students were given one hour to complete this assessment (November 2015). Students were assessed on the SCE-created informational/Expository Rubric. There are seven categories scored, each worth 1-4 points. The total rubric is worth 28 points. The following overall scale was established: 0-10 rubric points = 1 overall; 11-18 rubric points = 2 overall; 19-23 rubric points = 3 overall; 24-28 rubric points = 4 overall.

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
0-10% of students scoring a 3+ on Writing Composition	11-59% of students scoring a 3+ on Writing Composition	60-80% of students scoring a 3+ on Writing Composition	81-100% of students scoring a 3+ on Writing Composition

Questions to Consider

- What is the rationale for the weight of 49%?
- What is the timeline for giving the assessment? Is there more than one assessment given?
- Is this SLO for the year or a shorter time period?
- Would it make sense to narrow the "Less Than Expected" range (11-59%) and widen the "Much Less Than Expected "range to make them more comparable (0-10%) and increase the rigor of the success criteria?
- Out of how many points total is the "3+" scale for the rubric? Can more information be provided?

Strengths and Limitations

Strengths:
- It is an actual authentic assessment that is being used in the classroom
- Very clear, objective is a major goal for 5th grade.
- Robust size of learning objective.
- Connected to standards, well-structured and provides baseline data.

Limitations:
- Success criteria in the "Less Than Expected" category contain the majority of students.
- No consideration given for student growth in success criteria.
- The alignment of performance targets is too low compared to baseline data.

Voting Reflections

This measure was reviewed by two groups at the EE leaders' retreat. Based on the sum of votes from each group, both groups showed very high interest in this measure (**Group 1**: Very High Interest; **Group 2:** Very High Interest).

Student Learning Objectives (SLO) – Example 3

Measure Name	English 11 Synthesis Writing SLO
Educator Type	Teacher
Attribution (Individual or Collective)	Collective
Weight (Out of 100%)	Varies

Description: Students will construct effective synthesis essays that draw upon a variety of complex texts in order to develop persuasive theses; success criteria are specified in the SLO. At the beginning of the instructional interval, we possess a clear understanding of our students' strengths and weaknesses in synthesis writing, including data points related to each of the nine success criteria. Following these baseline evidence sources, the following formative assessments will allow us to track our students' progress toward their learning goals.

Measure Type: Teacher Custom; Student Learning Objectives customized for grade level teachers operating in a PLC

Content Area/Other: 11th grade English (College Prep English 11)

Additional Info: We have designed our Student Learning Objective in order to build upon multiple, triangulated sources of baseline evidence. Our first two data points consist of PLAN/ASPIRE English and Reading scores. Additionally, the Beginning-of-Year Student Survey (9-45 points possible) and Pretest (Formative Essay A; 18-90 points possible) are used to calculate a Baseline Score Composite (27-135 points possible) and to assign students to Baseline Performance Groups

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
67 and below =unsatisfactory	68-94 =partially proficient	95-121=proficient	122 and higher = advanced

Questions to Consider

- What is the target for the teacher (% movement and/or how much improvement)? Seems to only identify student's total point success measure.
- What if students perform lower at the end?
- How is "complex texts" defined (by grade level or individual student level)?
- How do the description and the success criteria come together? They appear to be disconnected.
- How was the post-test process outlined/established?
- What percent of students should fall into each success criteria category?

Strengths and Limitations

Strengths:
- Synthesis writing promotes higher level DOK.
- Use of multiple sources of baseline data to set performance targets.
- Supported through a PLC.
- Potential for inter-rater reliability through the PLC.

Limitations:
- Goals for specific student groups are not differentiated.
- There is no specification for the number of students in each success criteria category.
- May be overly complicated.

Voting Reflections

This measure was reviewed by two groups at the EE leaders' retreat. Based on the sum of votes from each group, one group showed low interest in this measure and the other group showed medium interest in this measure (**Group 2**: Low Interest; **Group 3:** Medium Interest).

Student Learning Objectives (SLO) – Example 4

Measure Name	Individual Educator Goal Example
Educator Type	Teacher
Attribution (Individual or Collective)	Individual
Weight (Out of 100%)	30%

Description:
Students will receive the following scores:

• Ratios and Rates Text scores:

      90% and higher correct, students exceeded the goal
      80 - 89% correct, students meet the goal
      60 - 79% correct, students partially meet the goal
      59% and below correct, students did not meet the goal

• The "How Much to Pay Task" and "Quiz Question" will be scored using rubrics which includes the accuracy of response, steps followed, and quality of explanation. Students will receive a rating of 1 (lowest) to 4 (highest) with a rating of '3' indicating they have met the growth goal.
• The MAP Achievement Status and Growth Report shows the comparison between Projected Growth and Observed Growth for each student. Students with a "yes" in the Met Projected Growth indicator meet the goal.
Using these data points and professional judgment, I will combine the assessments and tasks as follows:

Did not meet: Student performance was below expected on the assessments and tasks. (For example, a student with less than 59% on the test, 1 on the rubric for both the task and quiz question, and a "no" in the MAP Met Projected Growth
Partially meets: Student performance was below expected on some, but not all of the assessments and tasks. (For example, a student with less than 59% on the test, and a 1 on the rubric for the Quiz Question, but a 2 or 3 on the task and "yes" on MAP Met Projected Growth indicator).
Meets: Overall the body of evidence shows that the student has met the goal for each of the assessments and tasks.
Exceeds: Student performance was above expected on some of the assessments and tasks. (For example, a student with 90% or more on the test, and a 4 on the rubric for quiz question, and "yes" on MAP Met Projected Growth indicator, but a 3 on the rubric for the How Much to Pay Task).
Measure Type: Local Custom - Combination of Assessments
Multiple sources of baseline and end of interval data are encouraged. In this example, the teachers used vendor-produced and teacher custom tasks and assessments

Success Criteria

Much Less Than Expected	<------ 0 to 30 points ----->	More Than Expected
	90% or more of students meet/exceed the target = 30 points	
	80-89% of students meet/exceed the target = 25 points	
	60-79% of students meet/exceed the target = 20 points	
	40-59% of students meet/exceed the target = 15 points	
	20-39% of students meet/exceed the target = 10 points	
	10-19% of students meet/exceed the target = 5 points	
	0-9% of students meet/exceed the target = 0 points	

Questions to Consider	Strengths and Limitations
- Does each math teacher use the same measure? - How might this be simplified to create a greater understanding of how scores are found? - How does this convert to a 4pt scale for those using the State Model System? - Does the complexity of the measure hurt the progress of this SLO? - Is this anchored to objective measurement? - Can professional judgement be anchored to the objective measure (i.e., MAPS)?	**Strengths:** - Very rigorous, detailed and provides a clear rationale. - Teacher created for buy-in. - Uses multiple indicators for growth. - Great format to set expectations-no guessing games. **Limitations:** - Hard to understand at first glance. - High need for training and communication. - Needs simplification and/or help synthesizing how to navigate.

Voting Reflections

This measure was reviewed by two groups at the EE leaders' retreat. Based on the sum of votes from each group, one group showed high Interest in this measure and the other group showed very high interest in this measure (**Group 1**: High Interest; **Group 3:** Very High Interest).

Student Learning Objectives (SLO) – Example 5

Measure Name	Student Learning Outcome (SLO)
Educator Type	Teacher
Attribution (Individual or Collective)	Individual
Weight (Out of 100%)	30%

Description: Educators choose a single district proficiency scale (extended learning target) and follow an outlined SLO process.

Measure Type: Teacher Custom

Content Area/Other: Elementary: Reading or math. Secondary: Content area

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
Less than 25% of students reach expected student performance (as defined by educator in SLO template)	Student performance 10% lower than expected but above 25% (as defined by educator in SLO template)	Achieved expected student performance as set by the educator (defined in SLO template)	Exceeded expected student performance by 5% (as defined by educator in SLO template)

Questions to Consider

- Is this growth or achievement based?
- How do results from this measure help a teacher modify instruction?
- Is the "More than Expected" rigorous enough?
- Why is it only Math or Reading for elementary but all/any content for secondary?

Strengths and Limitations

Strengths:
- Measure can be attributed to a single individual.

Limitations:
- Success criteria unclear. Need to define what is meant by "expected" student performance.

Voting Reflections

This measure was reviewed by one group at the EE leaders' retreat. Based on the sum of votes for this group, they showed high interest in this measure (**Group 3** : High Interest).

Student Learning Objectives (SLO) – Example 6

Measure Name	4th Grade Math Expressions
Educator Type	Teacher
Attribution (Individual or Collective)	Individual
Weight (Out of 100%)	50%

Description: On the Math Expressions comprehensive Unit 2 posttest, 70% of students will achieve growth from the pretest of 50 percentage points or more, or will score 75% or higher.

Measure Type: Teacher Custom

Additional Info: Percent of students who earn a score of 70% or higher on the Math Expressions unit 2 exam. Percent of students who showed growth from the pretest to the posttest of 50 points or more. The median pretest score was 12%. 1/3 of the students were unable to apply previous knowledge to earn a single point. The concepts in Unit 2 are priority math standards for 4th grade math, requiring deep conceptual understanding to achieve mastery.

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
0 to 50%	51% to 70%	71% to 80%	81% to 100%

Questions to Consider

- How many SLOs/units are there throughout the year?
- Is the success criteria rigorous enough?
- What is the instructional time frame for this?
- Would it be appropriate to use the pre-test to set individual growth goals for student?

Strengths and Limitations

Strengths:
- The measure is descriptive.
- The assessment is already in place.
- It is specific to grade, content, & unit (possible model for other units).

Limitations:
- Possibly too much weight is being given to one unit.

Voting Reflections

This measure was reviewed by one group at the EE leaders' retreat. Based on the sum of votes for this group, they showed very high interest in this measure (**Group 3** : Very High Interest).

Student Learning Objectives (SLO) – Example 7

Measure Name	Student Learning Objective
Educator Type	Teacher
Attribution (Individual or Collective)	Individual
Weight (Out of 100%)	40%

Description: An SLO is a measurable, long-term academic goal informed by available data that a teacher or even a teacher team sets at the beginning of the year for all students. The teacher(s) and students work toward the SLO growth targets throughout the year and use interim, benchmark and formative assessments to assess progress toward the goal. At the end of the year, the teacher(s) meet with their evaluator to discuss the attainment of the SLO and determine the teacher's impact on student learning.

Measure Type: Teacher Custom

Content Area/Other: It is aligned to the content of the specific teacher.

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
Percentage of students meeting their expected target is less than 63%.	Percentage of students meeting their expected target is at or above 63%- below 75%.	Percentage of students meeting their expected target is at or above 75%- below 91%.	Percentage of students meeting their expected target is at or above 91%.

Questions to Consider

- What downside is there to pre-set success criteria?
- Are there common assessments?
- Is there an approval process at the start of the year?
- What are common assessments that may be used?

Strengths and Limitations

Strengths:
- Success criteria is pre-set.
- It is based on growth (differentiated for each student).
- Covers all content areas.

Limitations:
- Need to identify assessment(s) and measures and also the targets.

Voting Reflections

This measure was reviewed by one group at the EE leaders' retreat. Based on the sum of votes for this group, they showed medium interest in this measure (**Group 3**: Medium Interest).

Specialized Service Professionals

In this section, three examples of measures for specialized service professionals (SSP) are provided. The nine categories of SSPs are: audiologists, occupational therapists, physical therapists, school counselors, school nurses, school orientation and mobility specialists, school psychologists, school social workers and speech language pathologists. Measures for SSPs, referred to as Measures of Student Outcomes (MSO), are not limited to academic measures but may include measures focused on increasing access to learning and general student well-being.

It is important to note that these measures do not specify if the measure is attributed collectively or individually as, per S.B. 191, this is not a requirement for specialized service professionals.

Specialized Service Professionals (SSP) – Example 1

Measure Name	Occupational Therapist SLO
Educator Type	SSP
Attribution (Individual or Collective)	Not Applicable
Weight (Out of 100%)	60%

Description: Elementary OT Writing - Rubric evaluates (letter formation, size, spacing, word formation, and sentence format) November to March

Measure Type: Educator Custom

Success Criteria

Much Less Than Expected	Less Than Expected	Expected	More Than Expected
Less than 55% of students grew a minimum of 33%	56% to 74% of students grew a minimum of 33%	At least 75% of students grew a minimum of 33%	All students grew at least 33%

Questions to Consider

- What is the baseline measure and how is growth defined?
- What is the N size? Is it substantial enough for data collection?
- Is there additional progress-monitoring throughout this timeframe?
- From where does the 33% come? What exactly does this mean?
- Are pre- and post- the same prompt?
- Is the severity of the disability taken into account?

Strengths and Limitations

Strengths:
- Description is time-specific (Nov-March).
- Criteria is clear and makes sense.
- Meaningful and student specific.
- Very applicable to the SSP position.
- The baseline does not start at zero.

Limitations:
- It is not clear what is meant by a student grew by 33%.
- This does not account for student growth before November or after March.
- Success criteria does not account for students who grew substantially more than 33%.

Voting Reflections

This measure was reviewed by all three groups at the EE leaders' retreat. Based on the sum of votes from each group, one group showed medium interest in this measure, one group showed low interest in this measure and one group showed high interest in this measure (**Group 1**: Medium Interest; **Group 2:** Low Interest; **Group 3:** High Interest).

Specialized Service Professionals (SSP) – Example 2

Measure Name	IEP Objectives Obtained
Educator Type	SSP
Attribution (Individual or Collective)	Not Applicable
Weight (Out of 100%)	60%

Description: Percentage of objectives obtained on IEP goals related to Standard 3 would be self-monitored using dated annuals and triennials. Objectives would be from students on the SSP's case load.

Measure Type: District Custom

Content Area/Other: Aligns to multiple content areas and includes both behaviorally and academically based goals.

Success Criteria			
Much Less Than Expected	**Less Than Expected**	**Expected**	**More Than Expected**
0-24.9%	25-49.9%	50-74.9%	75% or above

Questions to Consider

- What criteria is being used to determine percent in levels?
- What is the instructional period?
- Is the success criteria talking about the percentage of students or the percentage of goals met?
- What number of students is this referring to?
- How is this information being tracked?
- Is it really 'ok' for expected for only 50% of goals to be met?
- Is there a base number of IEP goals for each student and is it consistent?

Strengths and Limitations

Strengths:
- Designed for the SSP to become an expert on his/her IEP case load. Self-monitoring promotes buy-in.
- Meaningful, student specific and well-structured.
- This aligns both behavior and academic outcomes.

Limitations:
- More clarity is needed on the success criteria (i.e., 50-75% of students will obtain 50-70% of their objectives).
- More specific information is needed regarding the success criteria.
- Need clarification regarding Standard 3 and the relationship to IEP goals.

Voting Reflections

This measure was reviewed by two groups at the EE leaders' retreat. Based on the sum of votes from each group, one group showed very high interest in this measure and the other group showed low interest in this measure (**Group 2**: Very High Interest; **Group 3**: Low Interest).

Specialized Service Professionals (SSP) – Example 3

Measure Name	Domain Team Progress on Goals
Educator Type	SSP
Attribution (Individual or Collective)	Not Applicable
Weight (Out of 100%)	15%

Description: Percentage of IEP goals met at level 3 (progress made, goal to be met on time) or level 4 (goal met)

Measure Type: Other; Based on most recent progress report results pulled from Infinite Campus - no later than March 30th of each school year.

Content Area: Speech Pathologist, Occupational Therapist, School Psych, Family Social Worker

Success Criteria			
Much Less Than Expected	**Less Than Expected**	**Expected**	**More Than Expected**
<65% of goals met at level 3 or 4	65% - 76% of goals met at level 3 or 4	77% - 85% of goals met at level 3 or 4	> than 85% of goals met at level 3 or 4

Questions to Consider

- How do you know the IEP goals are rigorous? Are they academic goals or behavioral?
- How do you set a reasonable data cutoff date and make sure the SSP provides services through the end of year?
- How do you manage reporting and progress towards goals prior to IEP meetings?
- What are the monitoring systems along the way?

Strengths and Limitations

Strengths:
- Cut-off date in time to compile data.
- Rigorous criteria.
- Team approach is good.

Limitations:
- Could differentiate more criteria for level 3 versus level 4.
- Teachers may lose motivation to teach after the cutoff date.

Voting Reflections

This measure was reviewed by one group at the EE leaders' retreat. Based on the sum of votes for this group, they showed low interest in this measure (**Group 1**: Low Interest).

Addendum A: Measure Template

MEASURE

Educator Type	
Measure Name	
Weight (out of 100% MSL)	
Attribution (collective or individual)	
Description:	
Measure Type:	
Content Area/Other	

Success Criteria			
Much Less Than Expected	**Less Than Expected**	**Expected**	**More than Expected**

Notes:

Addendum B: Feedback Form

Feedback

Measure:		Group Color	
I Like:	• • • •		
Questions:	• • • •		
Suggestions:	• • •		
Vote:			