WAYNE RESA
Leading... Learning for All

SERVICE
LEADERSHIP
COLLABORATION
EXCELLENCE

Measuring Student Growth:

A Practical Guide to Educator Evaluation

Table of Contents

Participants

The development of this ***Student Growth Guidance Document*** has been a collaborative effort involving many educators from across Wayne County, Michigan. These educators have been dedicated to identifying fair, transparent and appropriate methods for measuring student growth throughout the educator evaluation process. Teachers, administrators, central office leaders and ISD staff worked together to understand the research related to student growth models and the best ways with which to implement those models in today's educational environment.

The guidance suggested in this document is based upon a year and a half of study, analysis, debate and thoughtful reflection. This guidance document was not designed with the intention of being read cover to cover. Rather, each section could be read as a stand-alone to further your understanding of student growth. Targeted professional learning will be an important component as you implement this process. The intent of this guidance is to provide several methods whereby a district may be able to measure student growth for purposes of conducting evaluations. The list of participants below reflects the dedicated educators that contributed to this work:

NAME	TITLE	DISTRICT/ORGANIZATION
Brian Amorose	Teacher	Lincoln Park Public Schools
Blaine Armstrong	Principal	Flat Rock Public Schools
Mary Ruth Bird	District Data Coordinator	South Redford Public Schools
Dr. Sandra Brock	Director of Instructional Programs and Services	Northville Public Schools
Chris Buehner	Teacher	South Redford Public Schools
Dr. Patricia Drake	Special Education Data Consultant	Wayne RESA
Mike Flannery	Teacher	Henry Ford Public School Academy
Tom Martin	Principal	Woodhaven Brownstown Public Schools
Angelyn Maxon	Principal	Riverview Public Schools
Dr. Joseph Musial	Assessment Consultant	Wayne RESA
Linda Lazar	Title I Resource Teacher	Dearborn Public Schools
Lena Nemeth	Assessment Consultant	Wayne RESA
Brittany O'Brien	Teacher	Riverview Public Schools
Joe Orban	Director of Data Analysis and Assessment	Wayne Westland Community Schools
Stacy Peterson	Curriculum Director	Woodhaven Brownstown Public Schools
Dr. Lori Roy	Teacher	Livonia Public Schools
Dr. Paul Salah	Associate Superintendent, Educational Services	Wayne RESA
Dr. Sybil St. Clair	Executive Director of Research, Evaluation, Assessment and Accountability	Detroit Public Schools
Stephen Taylor	Coordinator of Student Services	Livonia Public Schools
Cindy Taraskiewicz	Assessment Consultant	Wayne RESA
Mari Treece	Manager/Educational Services	Wayne RESA
Dr. Ellen Vorenkamp	Assessment Consultant	Wayne RESA
Michelle Wagner	Teacher (Instructional Coach)	Van Buren Public Schools
Amy West	Teacher	Allen Park Public Schools

Graphic Design/Document Layout: Kate de Fuccio, Wayne RESA

WAYNE RESA
Leading... Learning for All

SERVICE
LEADERSHIP
COLLABORATION
EXCELLENCE

EDUCATION SERVICES

33500 Van Born Road
Wayne, Michigan 48184
wwww.resa.net

Paul Salah, Ed.D.
Associate Superintendent
(734) 334 1587
(734) 334 1729 FAX

July, 2016

Dear Educator:

Measuring student growth for purposes of educator evaluation is, in my summation, the most challenging aspect of assigning effectiveness labels to educators. Our country has grappled with the following question for several years: How does student growth align with an educator's effectiveness?

Wayne County educators decided that continuing to wait for an answer to this question was fruitless and potentially damaging to the education profession. Yes, damaging is a strong word, and I feel appropriate given the current climate of the education community. The focus of using student growth should be upon the improvement of teaching and learning and thus, logical, fair measures must be implemented. Selecting random cuts based upon proficiency or guesswork is not only inappropriate but also harmful. Harmful because until we solve the student growth quandary, people from many walks of life will not be focused upon teaching and learning, which is the single most important consideration for helping children achieve at high levels. Thus, as a Wayne County, we decided to be proactive and create an approach that determines effectiveness in a fair, thoughtful and transparent way.

This project began during the Winter of 2015 with a small group of dedicated educators grappling with the research, orchestrating a plan, and making a commitment to developing solutions rather than waiting for answers.

We read…

As an internal Wayne RESA team, a group of seven people began by delving into the research. We studied works by Stiggins, Popham and Darling Hammond. We studied the recommendations of Michigan Council for Educator Effectiveness along with works like the Widget Effect and Standard Setting by Cizek and Bunch. We explored the work of other states related to Student Learning Objectives, Formative Assessment, Assessment Choice and overall systems of high quality student growth.

THE WAYNE COUNTY REGIONAL EDUCATIONAL SERVICE AGENCY

Board of Education • James S. Beri • Kenneth E. Berlinn • Mary E. Blackmon • Lynda S. Jackson • James Petrie • Randy A. Liepa, Ph.D., Superintendent

We developed a team…

After some internal study amongst the Wayne RESA group, we invited fourteen school districts and Public School Academies from across Wayne County to come together around a common purpose—developing guidance regarding student growth. Our goal was to challenge the paradigms of the research, continue the learning and foster the voices of teachers, principals and central office administrators toward a common end—fair, transparent methods for measuring student growth. We also met with a subcommittee of Superintendents in order to help facilitate the thinking and development of this process.

After learning…

The team divided into sub-groups with a focus upon key areas related to student growth. As a result of continued debate, thinking and dialogue, a comprehensive Guidance Document designed to provide districts with choice was created. The Guidance Document that follows is designed to give districts options related to Student Growth.

In order to do this work well, districts must commit to intentional implementation, which includes growing capacity and understanding. The Guidance Document in and of itself is not the final answer. Rather, the thoughtful reflection and implementation that occurs after the fact will be essential to any district's success.

I want to thank each and every person that participated in this work. I truly valued the journey we embarked upon and am hopeful that the education community will benefit.

Sincerely,

Dr. Paul Salah
Associate Superintendent, Educational Services
Wayne RESA

List of Tables

List of Figures

SECTION 1
Growth Models

What are the different types of growth models?

Growth models measure the amount of academic progress students make between two points in time. There are numerous types of growth models but most tend to fall into six general categories:

1. **Improvement Model**
2. **Performance Index Model**
3. **Simple Growth Model**
4. **Growth to Proficiency Model**
5. **Student Learning Objective Model**
6. **Student Growth Percentile Model**

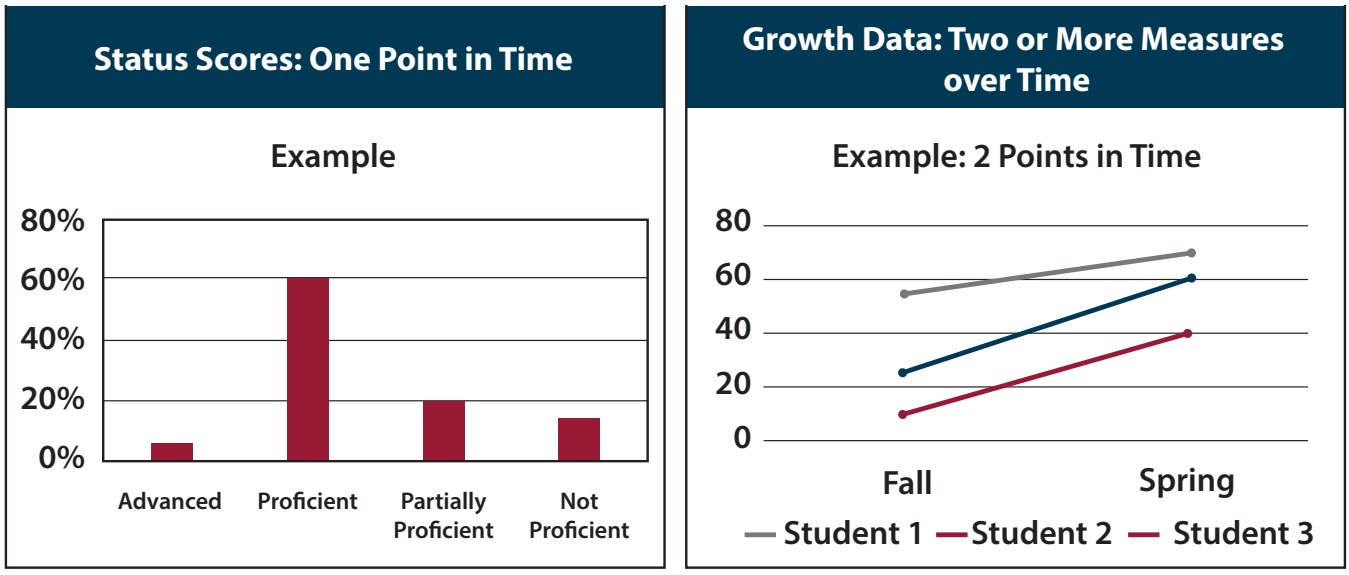
Each of these categories encompasses several variations depending on the model's purpose and available data. Because growth models are relatively new in education, and different models continue to be developed, these six categories may not capture all models.

Growth Based on Cohorts of Students vs. Longitudinal Data

It is important to make a distinction between status scores and growth data ***(Figure 1.1).*** **A status score is a measure taken at one point in time.** Each group of students is their own cohort. There are growth models that use status scores, comparing the performances of cohorts of students.

Growth data that are collected for the same cohort or group of students, looking at their progress in a longitudinal manner typically **involves assessment of students at two or more points in time.** There are growth models based on these longitudinal measurements. Often, these models also involve the use of targets as a basis for making comparisons and judgments of sufficient progress.

FIGURE 1.1 WHAT IS THE DIFFERENCE BETWEEN STATUS SCORES AND GROWTH DATA?



MODEL 1 *The Improvement Model*

The Improvement Model compares the scores of one cohort, or class of students in a particular grade to the scores of a subsequent cohort of students in the same grade. This model is based on achievement status—for example, students scoring at proficient or above. The difference in scores over time however would be considered growth for the educator's performance.

For example, if fifty-five percent of last year's fourth graders scored at or above proficient and sixty percent of this year's fourth graders reached proficiency, then, using the Improvement Model, this educator showed five percentage points in growth when considering fourth grade scores ***(Table 1.1).***

It does not measure growth among individual students or even the same cohort of students. The model actually compares two totally distinct groups of students, or in this example, last year's fourth graders to this year's fourth graders. The benefit of the Improvement Model is that it is fairly easy to implement and understand. While it does not track how individual students progress, it provides some indication of whether more students in a particular grade level are achieving proficiency from year to year. However, the change in the percent of students reaching proficiency may be due to different characteristics of the students in each cohort rather than a change in educator effectiveness, which can be perceived as a significant limitation of the Improvement Model. For example, the difference between last year's fourth graders' performance and this year's fourth graders could have been due to an increase in class sizes due to closing a nearby school in the district or a significant increase in special populations because of factors outside of the school district's control.

TABLE 1.1 EXAMPLE OF SUMMARY CUT SCORES FROM INDIVIDUAL PARTICIPANTS

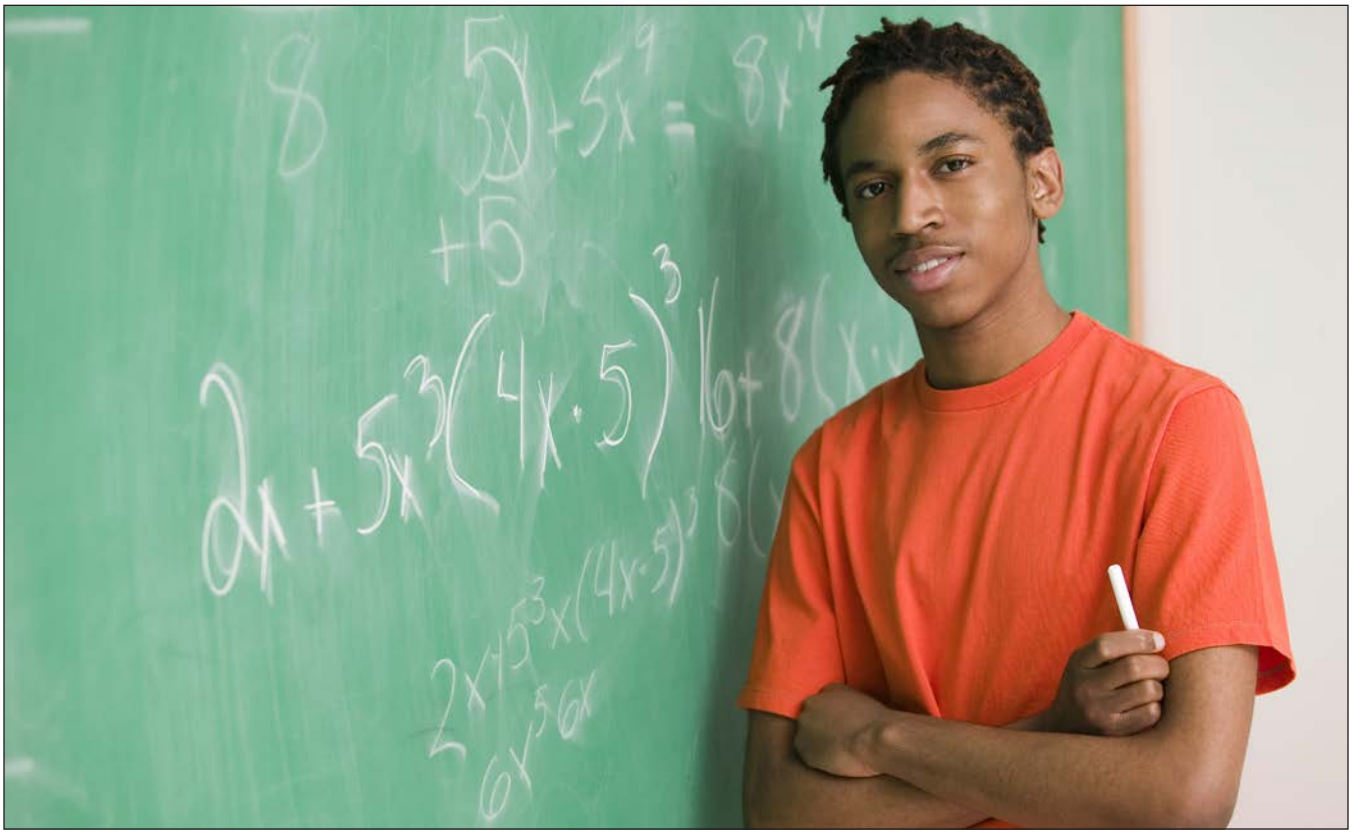
Achievement Level	Last Year's Grade 4	This Year's Grade 4	Change or "Growth"
Percent Proficient	55%	60%	+5%

MODEL 2 *The Performance Index Model*

Most Performance Index Models are status type models that give credit to educators for getting more students out of the lowest achievement levels even if they haven't reached proficiency. Just as with Status Models, Performance Index Models can be used as an Improvement Model. And just as with Improvement Models they do not necessarily measure the academic growth of individual students, but account for change in students' performance from one year to the next. There is, however, one important distinction: This model can be used to recognize change for students who are not proficient. As the example below shows, the educator received partial credit for the students scoring at the basic level but not below basic level.

In statistics, an index combines several indicators into one. Grade Point Average (GPA) is an index that we are all familiar with. It covers several indicators—grades students earn in various courses—and it is weighted in favor of the highest grades, an "A" is worth four points, a "B" is three points, a "C" is two points, and so on. To calculate the GPA, it is a matter of elementary math: Add up the grade points, divide by the number of courses, and the result is the GPA. The GPA shows how close students come to earning A's across their classes with straight A's earning a perfect 4.0 GPA.

Performance Index Models follow this same general principle. Think of it as the GPA for a school or educator where the goal is to determine how close the school or educator comes to getting all students to proficiency. It does so by **measuring student performance based on the percent of students scoring in each achievement level. More points are then awarded for students scoring at the highest levels**, just as students earn more points for higher grades. **The points are averaged and the result is the index.**



HOW A PERFORMANCE INDEX MODEL WORKS

The following example demonstrates the use of classroom data using the Index Model for educator evaluation. The index awards points for achievement as follows: ***(Table 1.2)***

TABLE 1.2 HOW A PERFORMANCE MODEL WORKS

Achievement Level	Points Possible
Students at Proficient and above	100 points
Students at Basic	50 points
Students at Below Basic	0 points

A perfect score of one hundred points means that all students reached proficiency. The educator would earn seventy-two points as shown in the table below ***(Table 1.3).*** Using an Improvement Model, this same educator would earn only fifty-five points for the percent of students who reached proficient.

TABLE 1.3 PERFORMANCE INDEX MODEL

Achievement Level	This Year's Grade	Computation	Points Awarded
Proficient +	55%	.55 x 100 points	55 points
Basic	25%	.25 x 50 points	13 points
Below Basic	20%	.20 x 0 points	4 points
Index Score for School or Teacher			**72 points**

When comparing performance from year to year, a Performance Index will include changes that occurred at the low end of student achievement and can also be designed to include changes among students scoring at proficient or better. For more information on how to determine the ratings to associate with points, refer to the Standard Setting section of this document.

Performance Index Models help schools and educators to concentrate on the growth of all students at all levels. There is credit applied for those students who are not yet proficient. Another advantage is that it can be used in situations where there is only one year of data. The index calculation helps to make the points awarded comparable across grade levels, content areas, and years.

Most Performance Index systems do not require sophisticated data systems. Keep in mind, however, that these models generally do not measure the growth of individual students from year to year. They also do not capture change within each achievement level. For example, if a team set a cut score of two hundred for "basic" and three hundred for "proficient," educators would not get credit for students whose scores improved from two hundred to two hundred ninety-eight. They would get credit for students who improved from two hundred ninety-nine to three hundred one. Establishing more achievement levels would help to capture these changes, making the model a more accurate measure of growth.

MODEL 3 *Simple Growth Model*

The Simple Growth Model is easy to calculate. **This model is most appropriate when working with assessments that provide scaled scores. A scaled score is the total number of correct questions (raw score) that have been converted into a consistent and standardized scale. The model is used to determine the difference in scaled scores from one point in time to the next.** But unlike the Improvement and most Performance Index Models, which compare successive cohorts at the same grade level (fourth graders in our hypothetical classroom), **Simple Growth Models actually document change in the scores of individual students as they progress from a baseline or pre-measure to the outcome or post-measure of learning.** The growth is calculated for each student who took both measures and the change or gain scores are then averaged for the educator's class or school.

HOW A SIMPLE GROWTH MODEL WORKS

This hypothetical educator has five fourth graders who took the fourth grade pre-assessment at the beginning of the year. The changes in scores are calculated in the table below for each student and a class average is reported in Table 1.4.

One drawback of this model is that only those students who took both assessments are included in the educator's growth calculation. Another limitation is that the points themselves provide no information. A fifty-point gain may or may not mean the student has met a set target or is on track to meet it in the future. For Simple Growth Models to be useful, experts, educators, and in many cases, policymakers must make informed judgments about how much growth is enough.

TABLE 1.4 SIMPLE GROWTH MODEL

Student	Pre-Test Score	Post-Test Score	Change
Student A	350	400	+50
Student B	370	415	+45
Student C	380	415	+35
Student D	325	390	+65
Student E	316	370	+60
Class or School Average	347	398	+51

MODEL 4 *Growth to Proficiency Model*

While Simple Growth Models measure individual student growth, they do not indicate if students are developing the skills they need to meet state standards of proficiency. **Growth to Proficiency Models—also known as Growth to Standards or On-Track—are designed to show whether students are on track to meet standards for proficient and above.** Although there are several variations, the **key ingredient across all Growth to Proficiency Models is that educators get credit when a student's progress keeps them on pace to reach an established benchmark—usually proficient—at a set time in the future,** typically within three to four years or by the end of high school (Davidson and Davidson 2004).

The advantages to this model are

- that schools and educators are recognized for producing gains in student performance even if their students score below proficient and
- there is more incentive to focus on all students below the proficiency level, not just the "bubble kids."

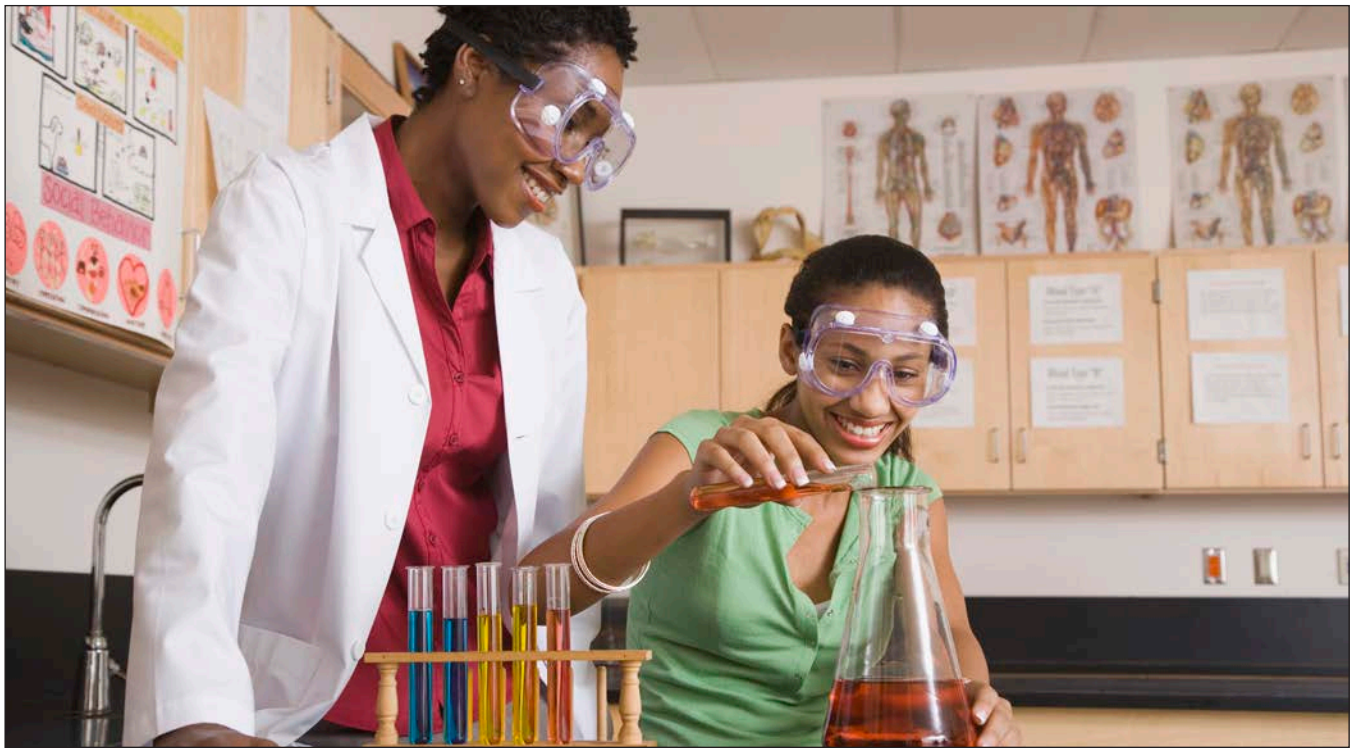
However, without targets, the model itself cannot determine which students are on track for proficiency. No matter what model is chosen, Growth or Status, it is up to the district to determine the process for setting goals in order to determine how much students should know and when they should know it. Then the model can be designed to determine which students are meeting those targets.

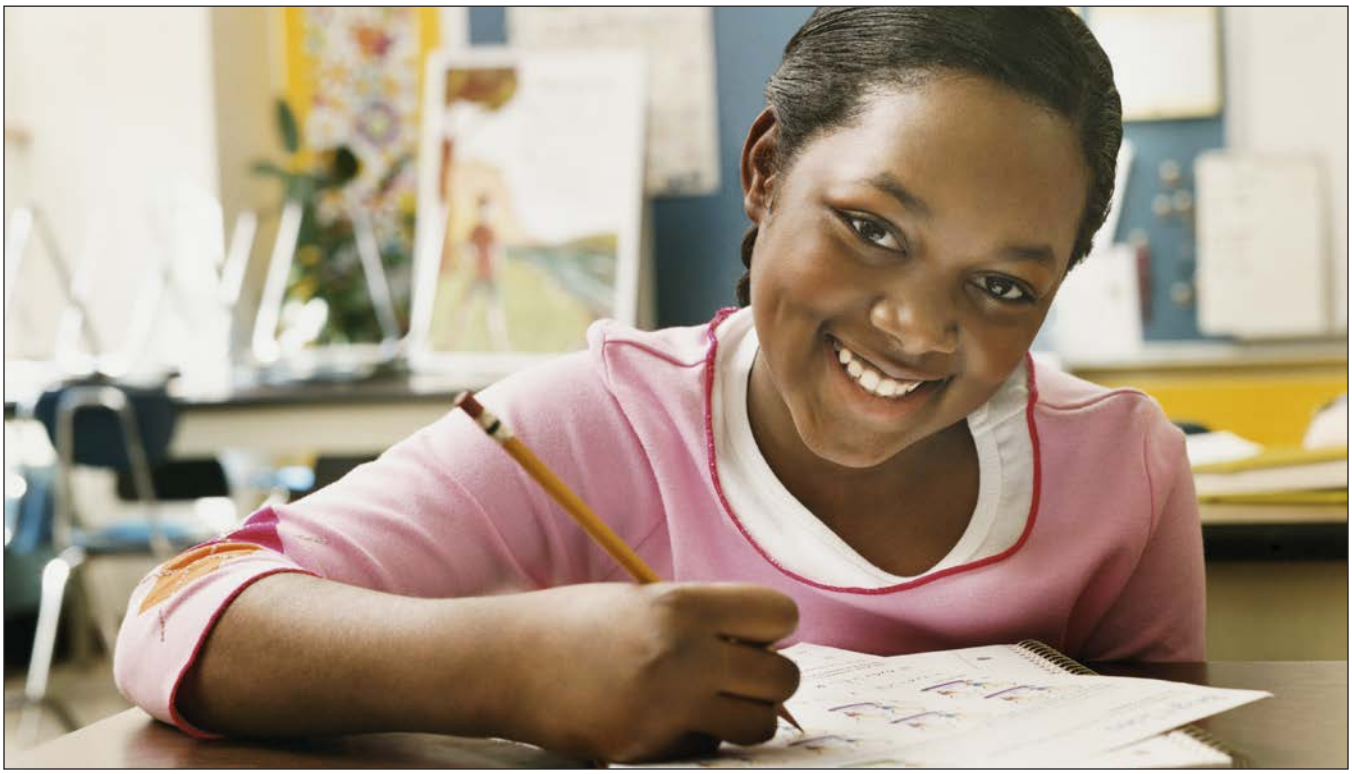
HOW A GROWTH TO PROFICIENCY MODEL WORKS

Our hypothetical classroom has five students whose growth targets were established at the end of fourth grade based on meeting proficiency in seventh grade ***(Table 1.5).*** Growth targets are based on the yearly growth needed to reach the seventh grade proficient score.

TABLE 1.5 HOW A GROWTH TO PROFICIENCY MODEL WORKS

5th Grade Proficient Score	7th Grade Proficient Score
400	600





In the example below, three of five students met the proficient target and therefore do not have to meet a growth target. Two students did not meet the proficiency target: One met his growth target while the other student did not meet hers. This means four out of five students met criteria, or eighty percent, which exceeds the seventy-five percent goal for this year. Therefore, the growth data indicate that the teacher's performance was satisfactory.

TABLE 1.6 PROFICIENCY TO GROWTH MODEL

Student	This Year's Grade 4 Score	This Year's Grade 4 Score	Change	Was the student's score Proficient?	What is the student's growth target?	Did Non-Proficient students hit growth target?	Did the student meet criteria?
Student A	350	400	+50	**Yes**	—	—	**Yes**
Student B	370	415	+45	**Yes**	—	—	**Yes**
Student C	380	415	+35	**Yes**	—	—	**Yes**
Student D	325	390	+65	**NO**	59	**Yes**	**Yes**
Student E	310	370	+60	**NO**	64	**NO**	**NO**
Class Goal	Goal=75% of students will meet criteria. 80% of students met goal.					+51	**Yes**

MODEL 5 *The Student Learning Objectives Model*

A Student Learning Objective (SLO) is a measure of an educator's impact upon student learning within a given interval of instruction. A SLO is a measurable, long-term academic goal informed by available data that an educator or educator team sets at the beginning of the instructional interval for all students or subgroups of students. The teacher and students work toward the SLO growth targets throughout the instructional interval and use interim, benchmark, and formative assessment data to assess progress toward the goal. At the end of the interval of instruction, the teacher meets with a principal or building team to discuss attainment of the SLO and determine the teacher's impact upon student learning.

HOW THE STUDENT LEARNING OBJECTIVES MODEL WORKS

Educators must understand assessment data and identify student achievement trends to set rigorous, yet realistic student growth targets that align with state standards, district priorities, and course objectives. These growth targets should include specific indicators of growth that demonstrate learning between two points in time.

In this hypothetical classroom, the teacher has assessed her students with a pre-test that has 100 points possible. The teacher then reviews where the students started and develops a growth target for each student, considering what will be reasonable, appropriate, and rigorous. The students are given another assessment as a post-test then a calculation of the difference is conducted. The difference becomes the growth score that is compared to the growth target. The teacher then determines the percent of students who met their target and compares that to the class target.

TABLE 1.7 SIMPLE GROWTH MODEL

Student	Pre-Test Score (Out of 100)	Post-Test Score (Out of 100)	Growth	Target Growth	Met Growth Target?
Student A	20	48	+28	+20	**Yes**
Student B	24	49	+25	+20	**Yes**
Student C	28	40	+12	+20	**NO**
Student D	45	55	+10	+15	**NO**
Student E	46	46	+0	+15	**NO**

The use of Student Learning Objectives has many benefits. The SLO process promotes learning, reflective teaching practices, the retention of teachers, and it aligns with many quality administrative and school improvement high impact initiatives.

Though the use of Student Learning Objectives is a promising practice, it is not without its challenges. SLOs can be a powerful solution if implemented with care and purpose, but they are not an easy solution. There can be a misconception that SLOs are the quick and easy fix to the challenge of assessing student growth, but in reality, much time and effort is required to execute the SLO process in a credible manner. Time and effort are necessary for planning, communicating, training, and monitoring SLO implementation to make the hoped for improvement in teacher effectiveness and student learning.

MODEL 6 *Student Growth Percentile Model*

The Student Growth Percentile (SGP) Model uses the calculation of student growth percentile scores to describe a student's growth compared to other students with similar test score histories (their academic peers). Although the calculations for SGPs are complex, percentiles are a familiar method of measuring students in comparison to their peers. Percentiles range from 0 to 99 and indicate how many scores in the comparison group are below that score. For example, if a student receives an SGP of 85, it means she demonstrated more growth than 85% of the students in the same grade and subject who had similar prior test scores.

HOW THE STUDENT GROWTH PERCENTILE WORKS

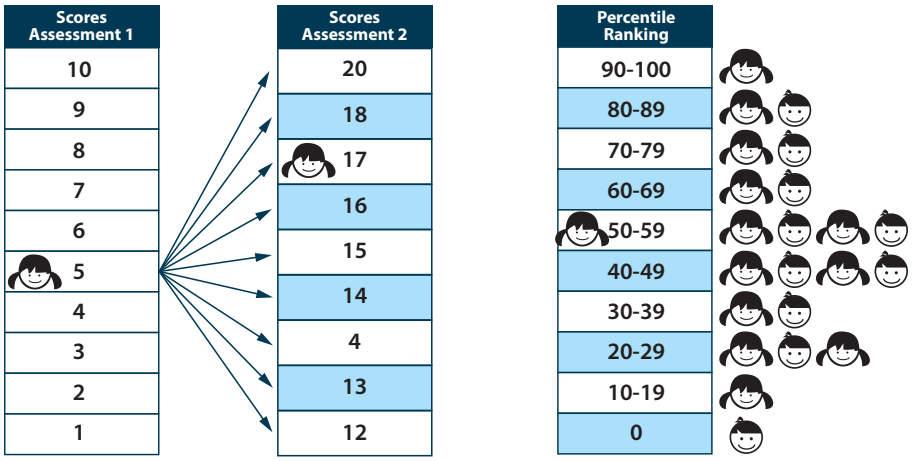
The student growth percentile score is typically calculated by the assessment developer. For example, assessments such as the NWEA and STAR provide a growth percentile calculation. The interpretation of the student growth percentile score is subject to the interpretation of the state or district. In the case of state assessment data, some states will set the cut point for establishing whether or not a teacher's performance is effective. With district purchased assessments, the district would be wise to engage in a process to review the data and define a context for establishing the percentiles that will be demarcated as effective versus ineffective. It can also demonstrate a student's growth and academic progress, even if she is not yet meeting the standard.

SGPs show how a student's achievement at the end of the year compares with that of other students who started the year at the same level. For example, if a student scored 5 on last year's test, their score at the end of this year would be compared with the scores of all the other students who scored 5 last year. The student's SGP would be the percentile rank (from 1 to 99) within this group of similar peers. If the student's SGP is 50, it means that his/her growth in test scores is right in the middle: half of the similar students who scored 5 last year scored higher than she did this year, and half of them scored lower.

The middle SGP score provides a simple indicator of how well the typical student in a class performed relative to similar students. This score is called the median growth percentile (MGP), and it is useful because, unlike a simple average, it doesn't change much if one or two students do unusually well or unusually poor relative to their peers. The MGP does not account for variations among students or classes, nor does it indicate what caused improvement.

Median growth percentile calculations do not try to adjust for differences in student characteristics. Neither SGPs nor value-added modeling indicates what might have caused improvements, nor do they reveal whether other students would make similar improvements if taught by that teacher.

FIGURE 1.2 STUDENT GROWTH PERCENTILE MODEL



What are the limitations of growth models?

Growth models hold great promise of evaluating schools on the amount of growth their students achieve. But growth models, especially value-added models, are relatively new in the education realm and their limitations are still being debated within the research community. Please note, however, that the research community is almost united in the opinion that growth models provide more accurate measures of schools than the current status models alone. Moreover current status models also suffer from many of the same limitations. While none of these issues should preclude states or districts from considering implementing a growth model, they do need to be acknowledged so the model developed will be the most effective tool for its purpose.

MEASURES OF ACHIEVEMENT CAN BE GOOD, BUT NONE IS PERFECT

This guide doesn't debate the pros and cons of standardized testing; there are plenty of publications that do. But it is necessary to discuss limitations and how they can affect the reliability of a growth measurement.

As discussed earlier, it's important to use tests that are appropriate for growth models. Growth can be measured without tests, but any tests used should have the following features:

- They cover the lower and upper ranges of student performance, rather than cluster test content around the knowledge and skills that constitute "proficient."
- They are vertically aligned and scaled to more accurately measure student achievement from year to year.
- They are aligned with state standards.

Unfortunately, while some tests are clearly better than others, there is no perfect measure of achievement, a statement to which even the most ardent supporter of standardized testing would agree.

One of the problems with tests used for growth models is that gain scores over time tend to be what statisticians call "noisier" than measures of achievement at a single point in time. By this, statisticians mean that gain scores tend to fluctuate even though a student's true ability typically does not change much from year to year. This happens because on any given test a student's performance is a result of his or her true ability and random influences, like distractions, during the test and the selection of items—effects that statisticians call measurement error. When scores from the two tests are subtracted from each other, as in Simple Growth models, the measurement error increases so the "true" performance becomes less clear.

There are statistical adjustments to minimize variance in the data, such as including scores from other subjects and previous years. Another way to minimize the effect of unreliable data is to create rolling averages by averaging growth over multiple years to provide a more stable picture of performance. However, such adjustments will add to the complexity of the growth model and may make it difficult to explain to educators why two schools (or teachers) with similar achievement gains received different ratings of effectiveness.

A SINGLE GROWTH MODEL DOES NOT SERVE ALL PURPOSES

Growth data can be helpful in many ways, and it is tempting to create one growth model and use it for multiple purposes. Policymakers and educators should resist this temptation. Although a single model could save a lot of time and money, many researchers strongly discourage using just one model, because trying to pull distinct pieces of information from one model would likely lead to false conclusions.

COMPARISON OF GROWTH MODELS

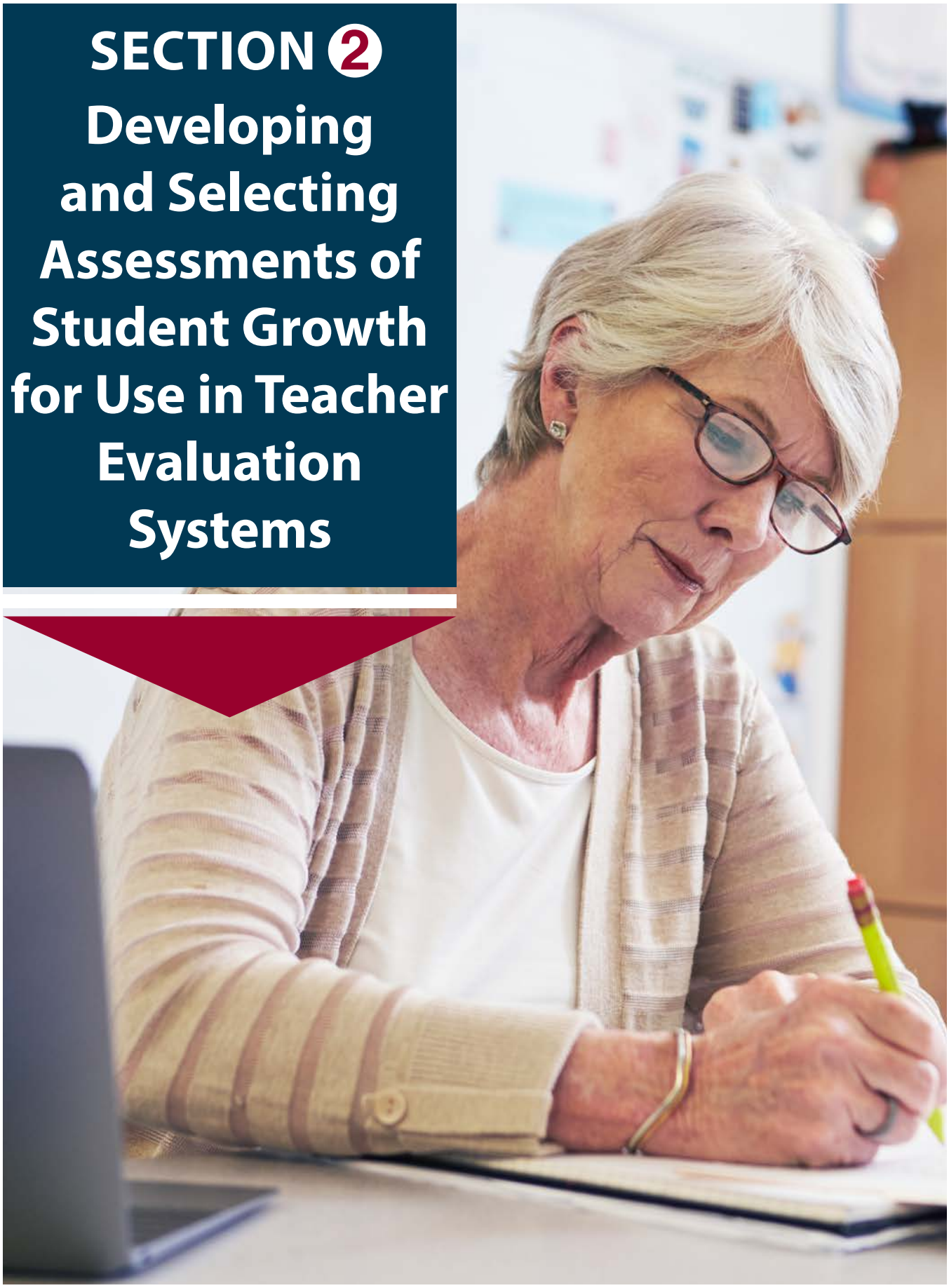
The following table provides a summary of the growth models described above that may assist the reader in making comparisons.

TABLE 1.8 GROWTH MODEL SUMMARY

Model	Description	Primary Question Asked	Benefits
Improvement Model	Compares the percent proficient from one year to the next with the same grade or teacher in that grade.	Are students in the teacher's class showing improvement from year to year?	• Easy to implement/ understand. • Indicates whether more students in a grade level are reaching proficiency from year to year.
Performance Index Model	Compares performance ranges by multiplying the percent proficient by an index score that converts the data to a "Performance Index" score. The performance index score is then compared to a standard or criteria for effectiveness.	How much gain are students showing, overall?	• The performance index can be used to credit gains made by proficient and non-proficient students. • The index creates comparability across teachers and content area.
Simple Growth Model	Describes growth with simple differences or average gains over time.	How much has the student learned on this scale?	• Simple to calculate. • Can be used with individual students or with groups of students.
Growth to Proficiency Model	Compares student gains to overall proficiency targets, including targets for non-proficient students based on projected trajectory to reach proficiency.	Are students proficient and making progress toward proficiency?	• Credits gains for students below proficiency.
Student Learning Objective Model	Establishes growth goals for students.	Are students meeting learning targets?	• Educators define learning targets. • Promotes learning, reflection, and retention of teachers. • Aligns with school improvement high impact initiatives.
Student Growth Percentile Model	Provides a ranking of a student's change in score when grouped with others who started at the same baseline score.	How does this student's growth rank in comparison to others of comparable prior ability?	• The model recognizes gains of all students, regardless of proficiency.

Primary Interpretation	Required Data Features	Setting Standards	Limitations
Growth description. Average group gains typically based on the percent proficient for different cohorts of students.	The use of the same assessment from year to year. Assumes the assessment has a vertical scale to be comparative from year to year.	Requires judgment about adequate gain or average gain. Requires understanding of the scale or can be norm-referenced.	• Does not measure growth for individual students. • Does not measure same cohort of students, change may be due to different student characteristics from year to year, not educator effectiveness.
Growth description. Provides a summary score that can be compared to a standard of effectiveness.	Can be used with raw scores, scaled scores, and percent proficient. Simple to calculate.	Requires setting of standards to determine performance bands, index weighting, and criteria for determining effectiveness of performance index scores.	• Does not measure growth for students from year to year. • Can mask change within score ranges.
Growth description. Simple difference between two points in time, typically compared to a norm or criterion to establish adequacy of the gains.	Vertically scaled data or scaled scores. Must control for instructional intervals between pre- and post-measurement.	Requires data review to establish adequacy of gains	• Requires business rules for missing data—students must take two measures. • Data can be manipulated to enhance gains. Requires control of instructional intervals when making comparisons between teachers.
A combination of growth description and growth prediction. Provides a description of overall proficiency including the accomplishment of projected gains for non-proficient students.	Vertically scaled data to be used to project proficiency in 3 to 5 years. Requires articulated cut scores across grades.	Standards are based on future scale or future standard. Requires setting targets for students based on projections of future performance. Requires setting of standard for effectiveness rating at teacher level.	• Does not attend to growth of consistently proficient students. • Problematic with student groups who will not achieve proficiency due to disability or disadvantage.
Growth prediction. Provides student level information that can be aggregated to reflect teacher performance.	Can be used with criterion referenced and/or standardized assessments at interim, benchmark, summative intervals.	Requires setting standards for future individual student learning objectives and for classroom targets based on a future standard of score.	• Time consuming. • Concerns with comparability of objectives and data sources. • Does not translate to year-to-year comparisons.
Growth description and growth prediction. Is used to interpret "on-track" students.	Requires large data sets to calculate the student growth percentile. Best with scaled score data.	Requires judgment about an adequate Student Growth Percentile or median/ average Student Growth Percentile. Predictions require a future standard and a time horizon to meet the standard.	• Sometimes misinterpreted as the percentile rank of gain scores. • Sometimes over interpreted as supporting value-added inferences. • Can be inflated by dropping initial scores.

SECTION 2
Developing and Selecting Assessments of Student Growth for Use in Teacher Evaluation Systems



New legislation requires districts to establish clear methods whereby the measurement of student growth will be utilized as an important aspect of determining an educator's effectiveness through a rigorous evaluation process. The assessments utilized for the purpose of measuring student growth must be nationally normed or locally adopted assessments aligned to state standards, or based on achievement of individualized education program goals. In this section of the guidance document, considerations are provided to help districts develop and/or select assessments for purposes of supporting the evaluation process.

CRITERIA FOR SELECTING ASSESSMENTS

When selecting assessments for the purpose of measuring student growth, educators should consider the components of the assessment as well as the validity of its intended use. Analyzing the components of the assessment should include two major considerations:

1. **The assessment must align to content standards and learning objectives that will be taught during the interval of instruction that will be used** for assigning an effectiveness label to the educator. When examining assessments for alignment, content teams should consider the following:
    - Items on the test should reflect the essential content standards for the course/grade.
    - No items on the assessment should reflect standards that were not part of the instructional plan for the course/grade.
    - The number of items dedicated to specific standards should mirror the instructional emphasis intended for the course/grade. For example, if the curriculum in a reading course focuses 50% of the instructional time on reading comprehension, then 90% of the assessment items should not be measuring knowledge of vocabulary.
    - The items on the assessment should represent the full range of cognitive thinking required in the course/grade. For example, if the curriculum in a third grade math class requires students to solve word problems and explain reasoning, then the assessment should require students to demonstrate this same level of reasoning.
2. **The assessment must provide enough item flexibility so that both high- and low-achieving students may adequately demonstrate their understanding of the content.** In order to accurately determine student growth, the assessment must provide enough range to demonstrate learning on both ends of the ability continuum. When selecting assessments that include adequate range, educators should consider the following:
    - All students should be able to demonstrate growth on the assessment.
    - The items should include basic knowledge and skills, as well as items that will challenge the highest performing students.

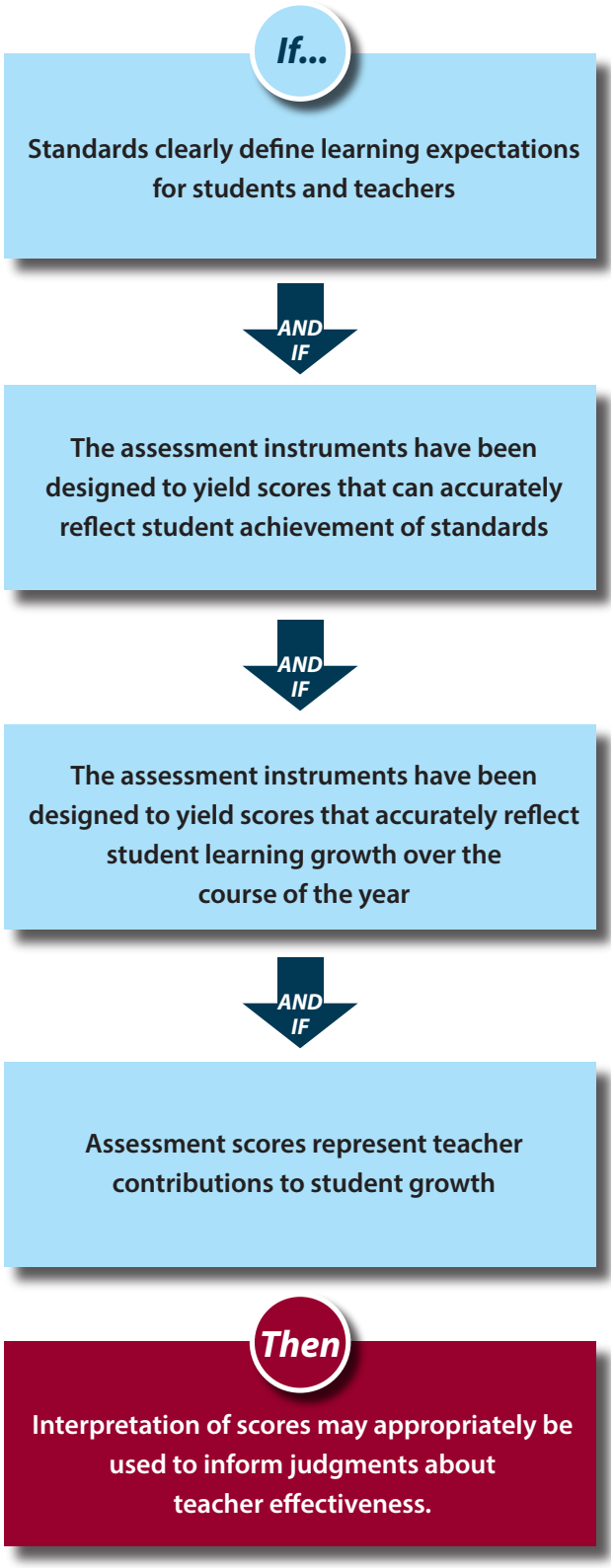
JUSTIFYING THE USE OF ASSESSMENTS IN TEACHER EVALUATION

The primary concept to be considered when selecting assessments for teacher evaluation is validity. **Validity is the extent to which an assessment measures what it is intended to measure and provides sound evidence for use as a part of the evaluation process.** As districts evaluate the purpose of assessments, they must consider validity.

Validation involves justifying the interpretation of scores as a determination of student growth over time and ensuring assessments are designed to report teacher effectiveness. Specific conclusions and decisions will be made on the basis of assessment performance; therefore, the following assurances regarding assessments should be verified:

- The purpose of the assessment is to measure student learning and growth over time.
- The standards of learning are clearly defined for both students and teachers.
- The assessment is of high quality, as measured by specific criteria and assurances. (Table 2.2)
- Student assessment scores accurately and fairly measure what students have learned in a given time period.
- Students' growth scores (based on assessments) accurately and fairly appraises the contributions of individual teachers.

FIGURE 2.1 JUSTIFYING THE USE OF ASSESSMENTS IN TEACHER EVALUATION



Adapted from Herman J.L., Heritage, M., and Goldschmidt, P. (2011). *Developing and Selecting Assessments of Students' Growth for Use in Teacher Evaluation Systems.*

ASSESSMENT INVENTORY

The first step in selecting assessments for the purpose of measuring student growth is to **determine which assessments are currently utilized throughout the district.** The assessment inventory is a tool designed to assist districts with the inventory process.

It is critical that all stakeholders understand the following:

- why specific assessments have been selected,
- the purpose the assessment serves,
- and how the district will use the data to impact student growth.

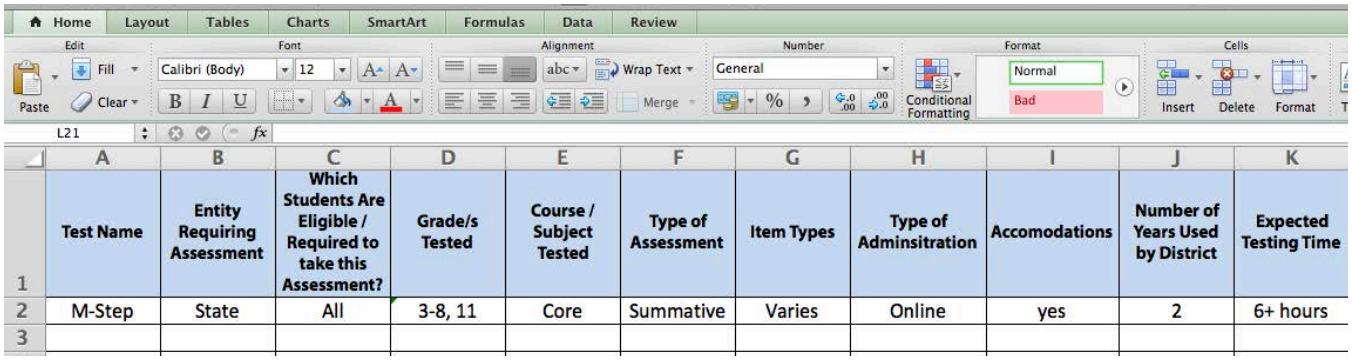
District and school leaders are encouraged to use and adapt this inventory to facilitate conversations about assessments with teachers, parents, and their local school board. This tool is also intended to identify gaps that exist across the district, in which an appropriate measure of student growth will need to be determined. Upon completion of this district-wide inventory, decisions can be made regarding which assessments are appropriate and should be used as a component of the evaluation system.

The assessment inventory is available in two electronic formats – Excel and Google Sheets. There are a variety of elements on the inventory that may be refined or adapted to provide the necessary data for collaborative district decisions. Elements from the assessment inventory include:

- Test Name
- Frequency of testing (i.e. monthly, quarterly, annually)
- Content standard alignment
- Intended purpose of the assessment
- Data collection methods and process for storing the data
- Identification of staff that will use the data
- Perceptions related to the appropriate use of the assessment
- Description of the processes in place to use the data
- Recommended reports to display student results
- Time between test administration and results provided to users of the data
- Intended decisions to be made from the data
- Information related to the assessment vendor/ developer
- Annual cost and funding sources

As district teams review these elements in collaborative teams, recommendations may be shared for the development and/or elimination of district assessments for the purpose of measuring student growth. If there are courses/subjects in which an assessment is not available or appropriate for use in a student growth model, educators should consider using criteria for the development of new assessments.

TABLE 2.1 ASSESSMENT INVENTORY



	A	B	C	D	E	F	G	H	I	J	K
1	Test Name	Entity Requiring Assessment	Which Students Are Eligible / Required to take this Assessment?	Grade/s Tested	Course / Subject Tested	Type of Assessment	Item Types	Type of Adminsitration	Accomodations	Number of Years Used by District	Expected Testing Time
2	M-Step	State	All	3-8, 11	Core	Summative	Varies	Online	yes	2	6+ hours
3											

CRITERIA FOR QUALITY ASSESSMENTS

Quality is a keystone of any assessment that will be used for the purpose of student growth. Districts must commit to **ensuring that all assessments associated with teacher evaluations are of high quality and have been verified as such through a defined process.** This process should be applied to vendor assessments, as well as those that have been locally developed. The following checklist may be used as a means to evaluate existing assessments for the purpose of educator evaluation.

TABLE 2.2 CRITERIA FOR QUALITY ASSESSMENTS

Criteria	Possible Evidence	Verified
Learning expectations are clear and reflect a progression (at minimum, the span of a grade level).	Sensitivity reviews by district team (e.g., curriculum director, department chair) Data analysis indicate consistent trends in achievement without significant fluctuations	
Specifications/blueprint for the assessment reflects the breadth and depth of learning expectations.		
Assessment tasks and items are designed to be accessible and fair for all students.		
Assessments are designed to accurately measure the growth of individual students.		
Assessments are designed to be instructionally sensitive to the standards being taught.		
Assessment items and tasks are consistent with the specifications/blueprint and are designed using item quality guidelines.	Item Quality Checklists *(included)*	
Assessment design, administration, and scoring procedures are likely to produce reliable results	Review of administration and scoring procedures (e.g., building principal, district team)	
Scores derived from the assessment are sufficiently precise and reliable.	Psychometric Analysis *(see Section 3: Measuring Student Growth)*	
Cut scores for student progress are justified.	Standard Setting Process *(see Section 4: Standard Setting for Student Growth)*	

STEPS FOR CREATING QUALITY LOCAL ASSESSMENTS

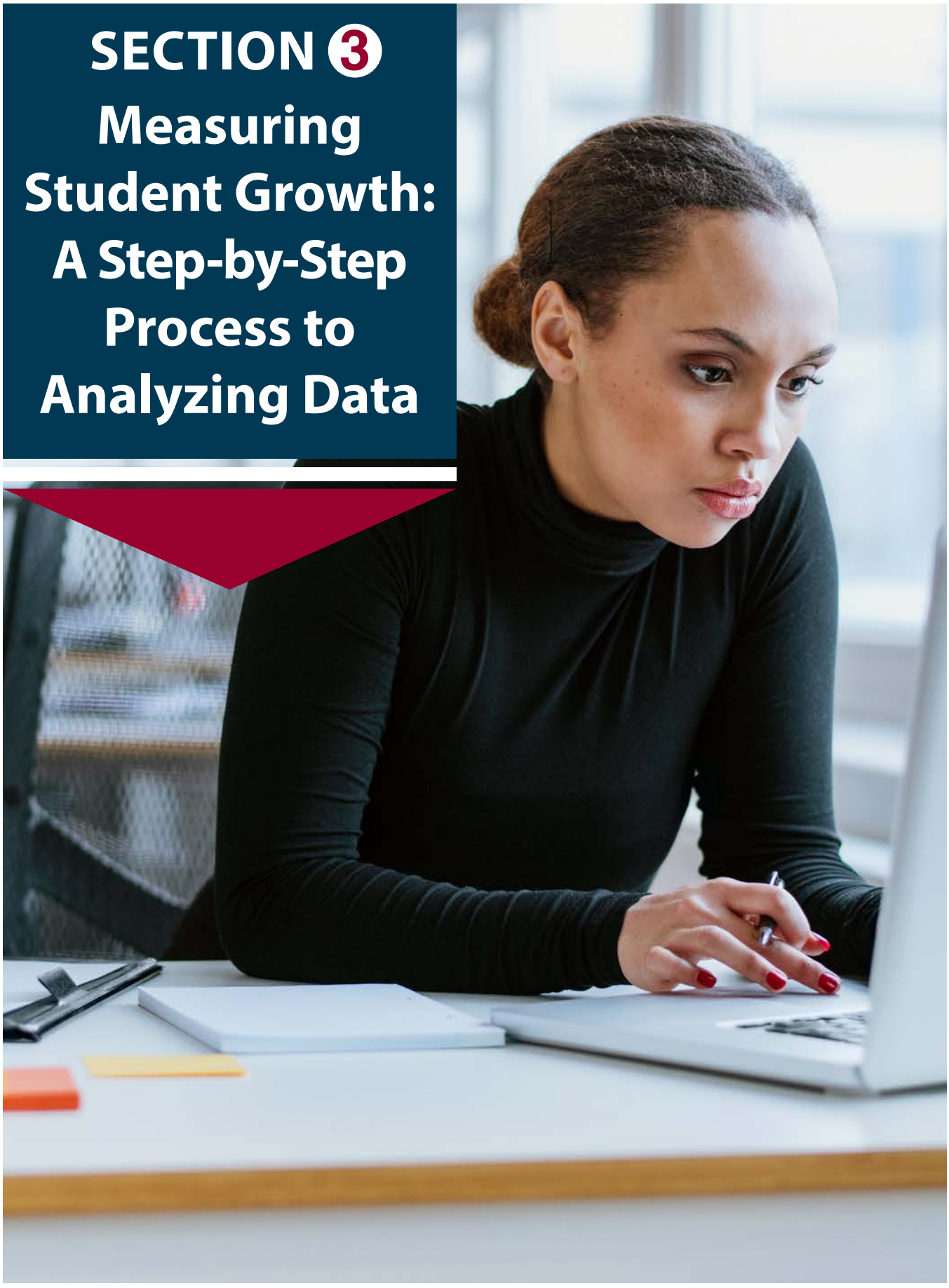
The criteria for quality assessments are recommended for justifying the decisions made in which student growth measures are attributed to teacher effectiveness. This process is essential for both locally developed and commercial assessments. Districts that elect to create new local assessments should employ processes and protocols to ensure consistent and quality results.

#	Step	Description	Review
1	**Assemble a Team**	Choosing a team of teachers that have knowledge in the content area will help verify that the assessment has content validity.	
2	**Establish the Purpose of the Assessment**	Clearly define the purpose of the assessment as a student growth measure.	*Michigan Assessment Consortium (MAC) Module 3:* **Determining the Outcome of the Common Assessment**
3	**Unpack/Deconstruct the Standards to Create Clear Learning Targets**	• **Important** – the standard or target is important to the discipline and/or important to the students' future. • **Leverage** – the standard or target has leverage, it represents enabling knowledge and/or skills. • **Appropriate** – the standard or target has been selected because it reasonably represents what the student can do at a given point in time (relative to the purpose of the assessment (e.g. pre-assessment, final end-of-course assessment). • **Clear** – the standard has been deconstructed so that it can be turned into a learning target and the standard is in language the student understands. • **Measureable** – the standard has been turned into a learning target that contains a verb that accurately represents the type of thinking, performance or disposition desired. • **Balance** – the collection of targets selected for one common assessment, or for a collection of common assessments for one grade and one academic year represent a variety of target types. For example, we are not unwittingly selecting a majority of targets that represent knowledge only.	*Michigan Assessment Consortium (MAC) Modules 4 and 5:* ***Identifying Learning Targets for the Common Assessment*** ***Matching the Assessment Methods to the Learning Target***

#	Step	Description	Review
4	**Ensure Access to the Content for ALL Students**	Provide appropriate accommodations that guarantee access to the assessment for students with special needs.	*Michigan Assessment Consortium (MAC) Module 6:* **Assessing Students with Special Needs**
5	**Create a Test Blueprint**	Design the assessment using the content and rigor of the learning targets. Choose the most appropriate assessment methods and quantity of items that will best reflect the standards taught for the instructional period. Consider the standard alignment, depth of knowledge (DOK), assessment method, item source, item sampling, and testing time.	*Michigan Assessment Consortium (MAC) Module 7:* **Writing the Test Blueprint**
6	**Select Items or Develop New Items**	Review content from item banks and make appropriate selections, based on the blueprint criteria. If there are no existing items to choose from, use quality guidelines for item creation.	*Michigan Assessment Consortium (MAC) Modules for Item Development 8-15:* **Module 8: Writing Selected Response Items – Part 1** **Module 8: Writing Selected Response Items – Part 2** **Module 9: Writing Constructed Response Items** **Module 10: Writing Performance Assessments** **Module 11: Using Portfolios to Assess Students** **Module 12: Developing and Using Scoring Guides and Rubrics** **Module 13: Editing the Draft Assessment Items** **Module 14: Detecting and Eliminating Bias and Distortion** **Module 15: Assembling the Assessment Instrument**

#	Step	Description	Review
7	**Field Test Items**	Evaluate the effectiveness of assessment items by analyzing real student data for item difficulty, item discrimination and bias. Analyze the reliability and validity of the assessment through psychometric principles and content expert reviews.	*Michigan Assessment Consortium (MAC) Modules for Field Testing 16-19:* **Module 16: Field Testing** **Module 17: Looking at Field Test Data** **Module 18: Reliability** **Module 19: Validity**
8	**Determine Standards for Adequate Growth**	Adhere to sound principles for standard setting and apply rigorous procedures to determine intervals of growth.	See section on Standard Setting for Student Growth
9	**Assessment Review**	Review the assessment periodically and revise to maintain alignment with standards and/or curriculum.	School/District Teams
10	**Administer and Score the Assessment**	Develop standardized procedures for the administration and scoring of the assessment. Establish procedures to ensure the validity of score interpretation as a measure of teacher effectiveness.	School/District Teams

SECTION 3
Measuring Student Growth: A Step-by-Step Process to Analyzing Data



The objectives of this section are to provide technical guidance to teachers and administrators with information regarding how to measure student growth, suggest appropriate calculations that purport to measure student learning, and highlight effective ways by which to communicate one's results to stakeholders. We will walk through basic steps to use when analyzing data.

The basic steps are listed:

1. **Locate your data**
2. **Find the central tendency**
3. **Calculate the variability of the scores**
4. **Run the statistics to determine if the growth scores are significant**
5. **Determine the magnitude of student growth**

The types of statistics that can be used to analyze data vary by the method utilized to collect data. For example, you will use a different procedure to understand pre- post- data when compared with progress monitoring data.

This guidance will provide the reader with an explanation of data analysis terms, demonstrate how to use appropriate data analysis methods, suggest resources that can be used to easily make calculations, and ultimately help with the interpretation of results.

GOOD DATA BASICS

Before we begin, it is important to consider good assessment practices and quality data requirements. **The assessment must be aligned to standards. It is also important for the assessment to be used for its intended purpose.** Here are some essential questions when choosing the assessments designed to measure student growth:

Assessment Alignment to Standards

- Is the assessment aligned to our standards?
- Does the assessment measure what is most important for students to learn and be able to do?
- Does the assessment measure what the educators intend to teach?

Assessment Purpose

- What is the purpose of the assessment?
- Does the assessment inform curriculum, instruction and practice?
- Does the assessment provide valuable information about student learning and help identify student progress toward desired targets?
- Does the assessment provide information about the magnitude of the instruction?

Multiple Measures

As a teacher, it is important to use multiple measures to gauge the progress of students and to demonstrate their learning. For the purposes of this document, **student growth is defined as the student's change in academic performance on a given assessment over two or more points in time.** Consider the sources of data available to you. The table below provides an example of the multiple assessments that a classroom teacher may be using to monitor student growth. The array of assessments present challenges in managing, summarizing and analyzing the data designed to represent student growth.

TABLE 3.1 AVAILABLE ASSESSMENTS

Available Assessments	Instructional Interval	Types of Scores
State Assessments	1 Year	Student Growth Percentile
District Tests	3 Times a Year	Gain Scores
Common Assessments	Quarterly	Proficiency Scores
Classroom Assessments	Monthly	Number and Percent Correct
Progress Monitoring	Bi-Weekly	Error Rates

Assessment Quality

Perhaps you have heard the saying, "Garbage in- Garbage out" referring to situations whereby bad quality of data has led to bad quality outcomes. Assessments must have integrity and be of good quality to be fair and appropriate for use with students and by teachers who are concerned with the growth of student learning. Some basic requirements of quality assessment practices and uses of data are listed below:

- **The teacher must have training on assessment literacy**, which refers to the understanding of assessments being used in the classroom.
- **The items on the assessment must be valid** in that they measure the content, as well as the construct the assessment is supposed to be measuring. This expectation applies to both purchased and locally developed assessments.
- **The assessment must provide reliable results.** This means the scores will be consistent on different administrations or when scored by different teachers. This expectation applies to purchased and locally developed assessments.
- **The test administration instructions must be standardized.** Standardized administration instructions, such as only "one correct answer" or "best choice", assure that every student has the same opportunity to understand the test. When test administration directions are altered, the difficulty of the test may be altered.
- **The grading of the test must also be standardized** so that different teachers will score the same. When two teachers score the performance of students, this is called "Inter-Rater Reliability". Reliability is determined with a statistical calculation called a reliability coefficient. If reliability coefficients are available, an acceptable level would be .70 or higher (George and Mallery, 2003).
- Research on effective practices in the measurement of student growth have determined that **there are optimal instructional intervals for determining growth**. A minimum of 2-weeks should lapse before post-testing and should not go beyond 6-months (Nunnally, 1978; Nunnally and Bernstein, 1994; Gay, 1992).

LET'S GET STARTED!

You are a classroom teacher and have been asked by your principal to provide evidence of student growth. **Student Growth will simply be defined as the student's change in academic performance on a given assessment over two or more points in time.** But the following questions come to mind:

- ***Where would I begin?***
- ***How is a typical student performing in my classroom?***
- ***How do individual student scores typically vary from one another?***
- ***How confident am I in the comparisons I am making with my data?***
- ***How do I show that my students made growth when I have pre and post test scores?***
- ***Is there a way I can make a calculation that would mathematically demonstrate the magnitude of growth my students made?***

Where do I begin?

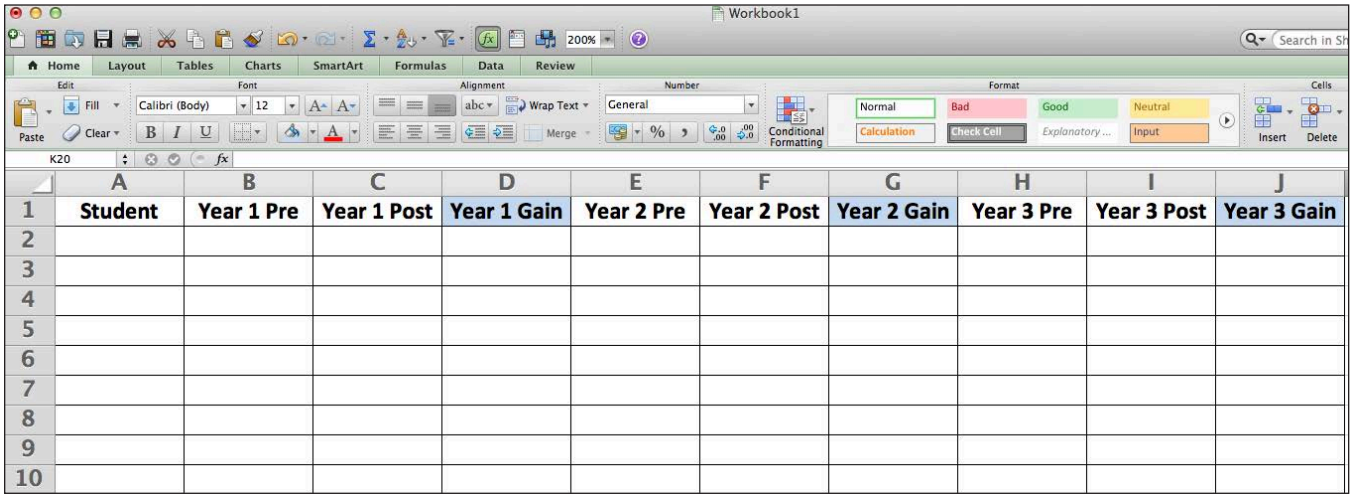
STEP 1 Locate the Data

Begin by locating and reviewing your classroom data. If you have it, you can use 3-years or more of data to demonstrate the growth of your students. In fact, in order to make any trend statement, you will need at least 3 years of data. If your school has developed common assessments, you may also refer to that data on your present and former students. You can summarize your data on a simple data collection form.

TABLE 3.2 EXAMPLE OF GROWTH DATA COLLECTION FORM

Student	Year 1 Pre-Test	Year 1 Post-Test	Year 1 Gain	Year 2 Pre-Test	Year 2 Post-Test	Year 2 Gain	Year 3 Pre-Test	Year 3 Post-Test	Year 3 Gain
Student 1									
Student 2									
Student 3									
Student 4									
Student 5									

FIGURE 3.1 EXAMPLE OF GROWTH DATA COLLECTION FORM—EXCEL



How is a typical student performing in your classroom?

STEP 2 *Find the Central Tendency*

Central tendency is a single student score that is representative or typical of the data set that you are using, in this case, your classroom data. There are three ways that central tendency can be measured: **Mean**, **Median** and **Mode**.

When setting growth goals and learning objectives for your class, you will want to refer to the central tendency of your data. This information will help you to determine typical growth while considering achievement projections that will be fair and rigorous.

Mean is the arithmetic average.

To calculate, add up all of the student scores and then divide by the number of students.

For example:

TABLE 3.3 CENTRAL TENDENCY MEAN EXAMPLE

Student	Score
A	2
B	3
C	4
D	7
Total	16

To Calculate Mean:

Total Scores: 2 + 3 + 4 + 7 = 16

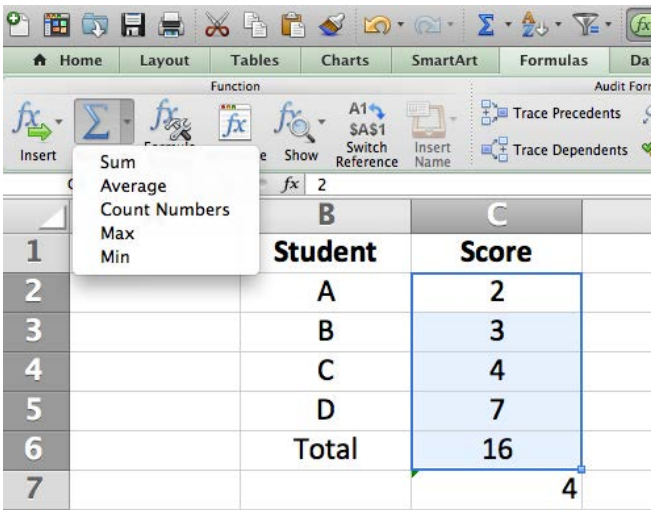
Total Number of Students: 4

16/4 = 4

The Mean = 4

In Excel:

FIGURE 3.2 CENTRAL TENDENCY MEAN EXAMPLE—EXCEL

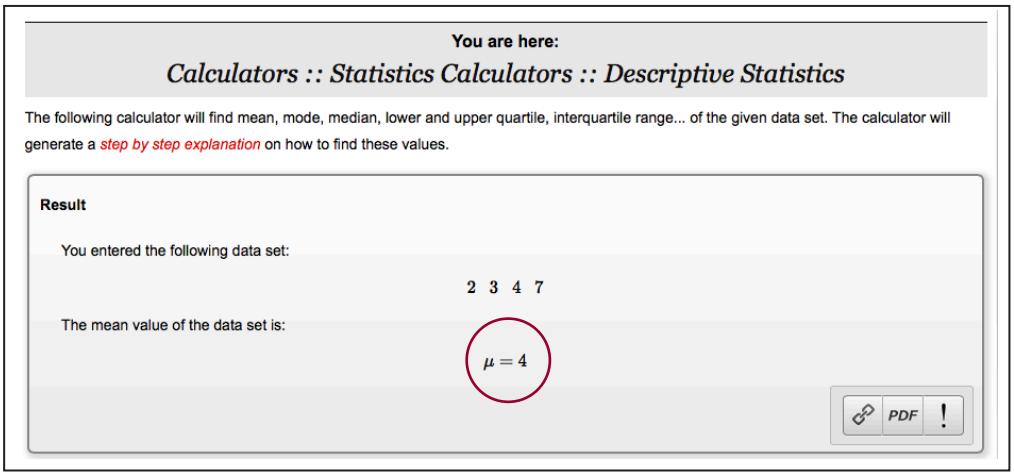


To Calculate Mean with an Excel Spreadsheet:

1. Enter the scores for each student.
2. Select/highlight the scores
3. Go to the tab "Formulas"
4. Click on "Auto Sum"
5. Click on "Average"
6. The spreadsheet will automatically calculate the mean.

With an Online Calculator

FIGURE 3.3 CENTRAL TENDENCY MEAN EXAMPLE—ONLINE CALCULATOR



To Calculate Mean with an Online Calculator

1. Go to the following web page
   http://www.mathportal.org/calculators/statistics-calculator/descriptive-statistics-calculator.php
2. Enter your data, separated by a comma, colon, or semicolon, or blank space.
3. Select "Find Arithmetic Mean".
4. Click on "Calculate".

Talking about the Mean

If you were asked to interpret the **mean score**, you might state:

"On this assessment, the mean score was 4."

Median is the midpoint or score that is in the middle of a list of scores. If you have percentile data, the median is the 50th percentile.

To calculate, order the numbers from lowest to highest and then identify the midpoint.

For example:

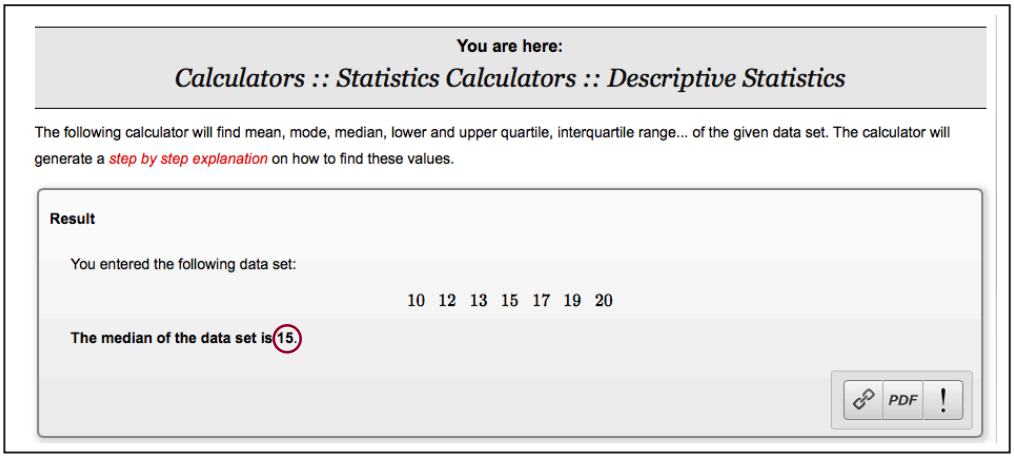
TABLE 3.4 CENTRAL TENDENCY MEDIAN EXAMPLE

Student	Score
A	10
B	12
C	13
D	15
E	17
F	19
G	20

In this example, the median is 15. One-half of the student scores fall below this number and one-half fall above it.

With an Online Calculator

FIGURE 3.4 CENTRAL TENDENCY MEDIAN EXAMPLE—ONLINE CALCULATOR



To Calculate Median with an Online Calculator

1. Go to the following web page:
   http://www.mathportal.org/calculators/statistics-calculator/descriptive-statistics-calculator.php
2. Enter your data, separated by a comma, colon, or semicolon, or blank space.
3. Select "Find Median".
4. Click on "Calculate".

Talking about the Median

If you were asked to interpret the **median** of your data, you might state:

"The median score was 15. This score represents the midpoint in the scores."

Mode is the most frequently occurring student score.

To calculate, order the numbers from lowest to highest and then identify the most frequently occuring score.

For example:

TABLE 3.5 CENTRAL TENDENCY MODE EXAMPLE

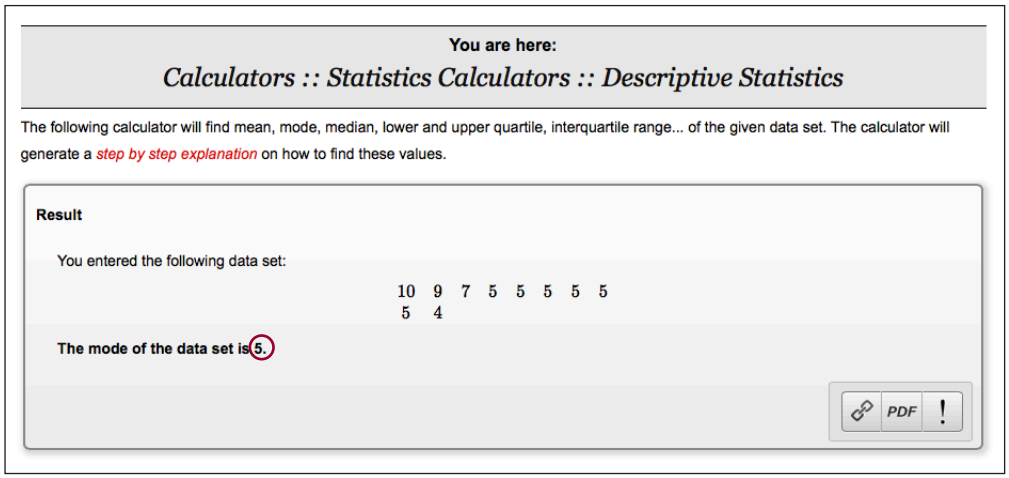
Student	Score
A	10
B	9
C	7
D	5
E	5
F	5
G	5
H	5
I	5
J	4

In the example above, a score of 5 occurs six times. It is more frequent in the data set than any other score.

The Mode = 5

With an Online Calculator

FIGURE 3.5 CENTRAL TENDENCY MODE EXAMPLE—ONLINE CALCULATOR



To Calculate Mode with an Online Calculator

1. Go to the following web page:
   http://www.mathportal.org/calculators/statistics-calculator/descriptive-statistics-calculator.php
2. Enter your data, separated by a comma, colon, or semicolon, or blank space.
3. Select "Find Mode".
4. Click on "Calculate".

Talking about the Mode

If you were asked to interpret the **mode** of your data, you might state:

"On this assessment, the mode score was 5."

Student Performance Summary

Now that you have the Mean, Median, and Mode, you can describe how students typically perform on the assessment. You may summarize your data in a table format. An example follows:

For example:

TABLE 3.6 STUDENT PERFORMANCE SUMMARY

	Grade 4	Grade 5	Grade 6
Central Tendency	**2014 (N=30)**	**2015 (N=30)**	**2016 (N=30)**
Mean	18	18.5	19
Median	13	15	16
Mode	14	15	16
N= the total number of students			

Giving a Student Performance Summary

If you were asked to describe the typical performance of students for the past three years, you might state:

"Three years of cohort data are available, including 2014, 2015, and 2016. There were 30 students tested each year. The mean scores were 18, 18.5 and 19, respectively. Improvements were noted in median performance. The median score in 2014 was 13. The median score was 15 in 2015. The median score in 2016 was 16. The mode scores were 14 in 2014, 15 in 2015, and 16 in 2016."

How do individual student scores typically vary from one another?

STEP 3 *Calculate Your Variability of Scores*

Variability means that things are not equal to one another. Look at your data and you will see that there are high and low scores. There are methods to analyze and understand this variability in student performance.

The three ways to describe and analyze the variability of test scores include the **range, the variance, and the standard deviation**.

This information might help you to set targets that represent the breadth of range and normal expectations for learning.

Range is the highest student score minus the lowest student score.

Be aware of scores that are outliers, either very high or very low and not typical for your data set. An extreme score may distort how you report the range of your student's performance.

To calculate: Subtract the lowest score from the highest score and the difference is the range.

For example:

TABLE 3.7 CALCULATING RANGE

Student	Score
A	95
B	92
C	84
D	82
E	81
F	79
G	75
H	71
I	64
J	60

95–60=35
The Range = 35

Talking about the Range

If you were asked to describe the **range** of your data, you might state:

"The range of scores was 35 points, with 60 as the lowest score and 95 as the highest."

Variance is a standard measure of how individual scores move away from (or disperse) from the mean.

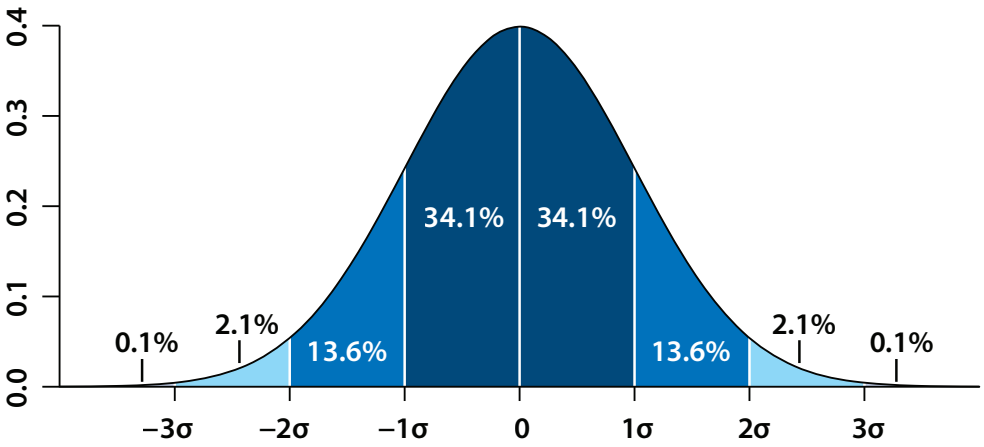
Think of a playground teeter-totter in which the middle (or fulcrum) represents the mean of a test and multiple student scores move away from the mean on both sides of the teeter-totter. **The variance is not reported when discussing student test performance.**

Standard deviation is the square root of the variance.

When reviewing your student scores, 68% of your students will fall within plus or minus one standard deviation. And 95% of your students will fall within plus or minus two standard deviations. You should report the standard deviation when discussing student test performance.

For example:

FIGURE 3.6 STANDARD DEVIATION



In the above graph of the normal distribution, the symbol μ = **Mean**.

The symbol σ = **Standard Deviation**.

-1 σ is 1 standard deviation below the mean while 1 σ is 1 standard deviation above the mean.

Percentile scores between 25 - 75 are the **Average Range**, between -1σ and +1σ.

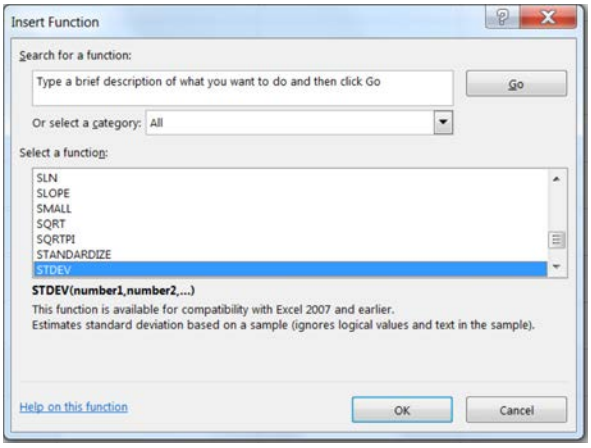
Calculate the Standard Deviation in Excel:

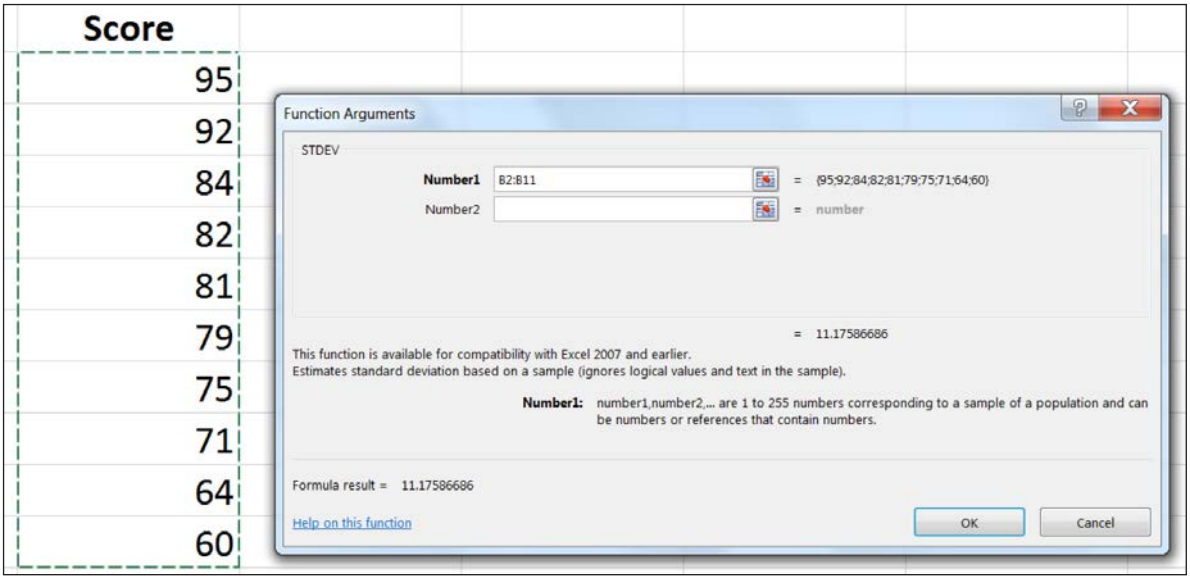
FIGURE 3.7 CALCULATE THE STANDARD DEVIATION IN EXCEL

To Calculate the Standard Deviation with an Excel Spreadsheet:

1. Enter your scores.
2. Go to the "Formula" tab.
3. Select the "Auto Sum".
4. In drop-down menu, Select "Other".
5. You will see a blue box.
6. Select or search for "StDev", which stands for standard deviation.
7. Highlight the column of scores.
8. Click "OK".
9. The "Formula Result" is the Standard Deviation.

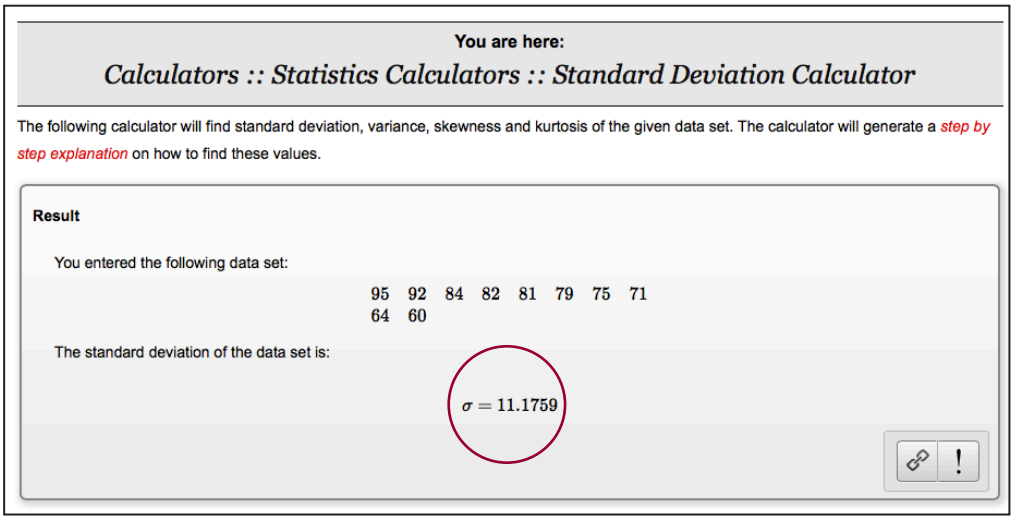
	A	B
1	Student	Score
2	A	95
3	B	92
4	C	84
5	D	82
6	E	81
7	F	79
8	G	75
9	H	71
10	I	64
11	J	60





Calculate the Standard Deviation with an Online Calculator

FIGURE 3.8 CALCULATE STANDARD DEVIATION WITH AN ONLINE CALCULATOR



To Calculate Mode with an Online Calculator

1. Go to the following web page:
   http://www.mathportal.org/calculators/statistics-calculator/standard-deviation-calculator.php
2. Enter your data, separated by a comma, colon, or semicolon, or blank space.
3. Select "Find Standard Deviation (default)".
4. Click on "Calculate".

The symbol σ is used to represent the **Standard Deviation**.

Talking about the Standard Deviation

If you were asked to describe and interpret the **Standard Deviation** of your data, you might state:

"The mean of student scores was 78.3 with a standard deviation of 11.18. The average score varied by as much as 11.18 points. In other words, 68% of the student scores typically fell between 67.3 and 89.5."

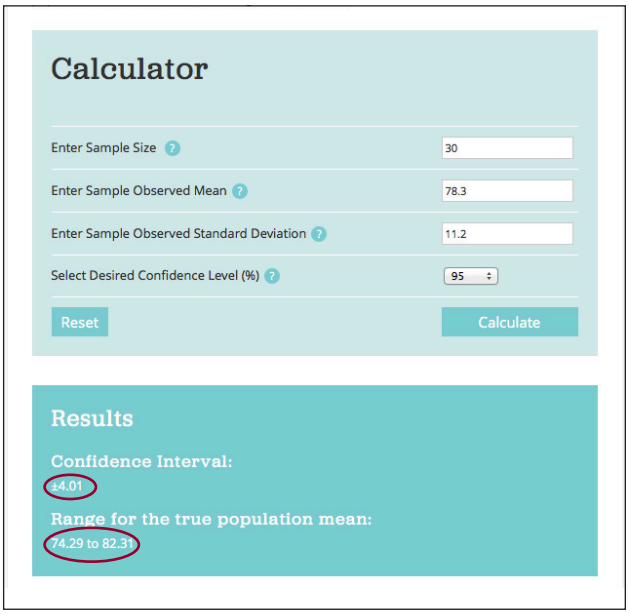
How confident am I in the comparisons I am making with my data?

Confidence Interval represents the probability that the true mean will fall within a range.

The confidence interval is one of the most important statistics to calculate. The 95% confidence interval is simply a range in which a teacher is 95% confident that the true mean lies within the upper and lower limits of the interval.

With an Online Calculator

FIGURE 3.8 CALCULATE CONFIDENCE INTERVAL WITH AN ONLINE CALCULATOR



Go to the following web page:

https://www.mccallum-layton.co.uk/tools/statistic-calculators/confidence-interval-for-mean-calculator/#confidence-interval-for-mean-calculator

1. Locate the **Mean** and **Standard Deviation** for your data.
2. Enter these into the calculator fields.
3. Select 95% Confidence Level (Other options may be used e.g. 90% C.I.)
4. Click on "Calculate".

In this example, the **confidence interval** is + 4.01. The teacher is confident that the true mean is between 74.29 and 82.31.

This is important because if future growth projections are outside of the confidence interval, the target would be highly unlikely to achieve.

Talking about Confidence Interval

If you were asked to describe and interpret the **confidence interval** of your data, you might state:

"I am 95% confident that my true mean score will be between 74.29 and 82.31."

Using Data Carefully

Let's consider how the confidence interval will help an administrator and teacher set realistic and reasonable targets using the data in this example.

Without the confidence interval the administrator and teacher have the following information:

Number of Students	30
Mean	78.3
Standard Deviation	11.2

At first glance, one might think that 67 - 89 is typical performance. Based on this information, the administrator might even set a target above 89, thinking this would be reasonable because it is more than one standard deviation above the mean. This target is problematic. With the confidence interval, the administrator and teacher would be able to establish, with high probability and confidence, that the true mean of the students falls in the range of + 4.01. The true mean is between 74.29 and 82.31 with 95% confidence. This is a more conservative range for target setting. If the administrator were to continue to set the target outside this range, the teacher would have less than a 5% chance of hitting the target. This would be unfair, unrealistic, and a misuse of data.

How do I show that my students made growth when I have pre- and post-test scores?

STEP 4 *Tests to Determine If the Growth Scores Are Significant*

PRE-TEST AND POST-TEST MEASURES

It is a common practice for teachers to assess students in the fall and later in the school year to measure the growth of their students.

REPEATED MEASURES

In situations where the teacher is progress monitoring, s/he is taking what is called "Repeated Measures". Repeated measures permit a teacher to assess the same student on multiple occasions over time.

There are three common statistics for analyzing data that are collected with pre-test/post-test measures and/or repeated measures:

1. **Percent change**
2. **Wilcoxon signed-rank test**
3. **Paired-samples *t*-test**

They are well-suited to measure a student's pre-test/post-test score by determining a difference score over time (Gravetter and Wallnau, 2004). These statistics are especially useful for measuring growth because the student's scores are paired for the analysis. For example, the individual student's pre-test score is paired with his/her post-test score. When the individual student's scores are matched in this way, the data has less error that is introduced by other types of comparisons. Other benefits associated with the repeated measures design include:

- The ability to measure a small number of students
- The ability to identify prior baseline knowledge
- Determining statistical significance (which helps to show the effectiveness of instruction)
- Measuring the difference score or growth
- Calculating percent change
- Measuring the magnitude of instruction by the effect size (ES), which is reported in standard deviation units
- Controlling for extreme scores

How could these tests be used when teachers are working with a common assessment?

In this example, teachers in a building developed a common assessment. The teachers pre-tested all of the students at the beginning of the course. With on-going instruction, the teachers assessed students at a midpoint in the semester and again at the end of the semester.

Let us review the three methods when analyzing data.

Percent Change is the extent to which something gains or loses value.

To calculate percent change, here is a simple formula:

(Post-test Score Minus Pre-test Score) Divided by
(Pre-test Score) x 100 = Percent Change

A **Positive Score means there is a gain** in student achievement.

A **Negative score indicates a decline** in student achievement.

For example:

1. Calculate the mean pre-test score.
2. Calculate the mean post-test score.
3. Subtract the pre-test mean from the post-test score.
4. Divide the results by the pre-test mean.

TABLE 3.8 CALCULATING PERCENT CHANGE

Pre-test	Post-test
2	6
3	7
1	5
4	10
1	6
2	8
2	9
1	3
3	6
2	3

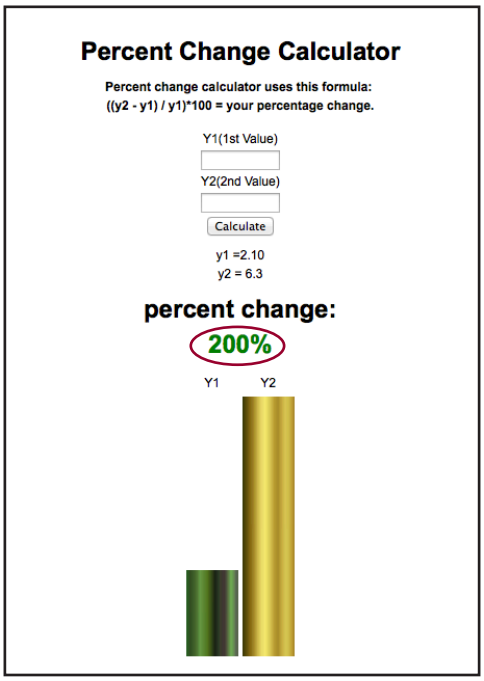
Pre-test Mean: 2.10

Post-test Mean: 6.30

(6.30 - 2.10) = 4.2/2.10 = 2 x 100 = 200 % Change

With an Online Calculator

FIGURE 3.10 CALCULATE PERCENT CHANGE WITH AN ONLINE CALCULATOR



Go to the following web page:

http://www.percent-change.com

1. Enter the pre-test mean data in Y1.
2. Enter the post-test mean data in Y2.
3. Click on "Calculate".

Talking about Percent Change

If you were asked to describe and interpret the **percent change** of your data, you might state:

"The pre-test mean was 2.10. The post-test mean was 6.30. This represents a 200% gain in student scores."

Wilcoxon Signed–Rank Test compares two means using the same student in order to determine if the difference is statistically significant (if the *p* value is < 0.05).

The Wilcoxon signed-rank test should be used any time the student group is not typical. In other words, this test should be used when randomization is absent.

With an Online Calculator

As an example, a veteran teacher has been assigned to teach a large number of at-risk high school students and the teacher has collected pre-test and post-test data for one semester. How will this Wilcoxon Signed-Rank test be calculated?

The best and easiest method to do this is using an online calculator.

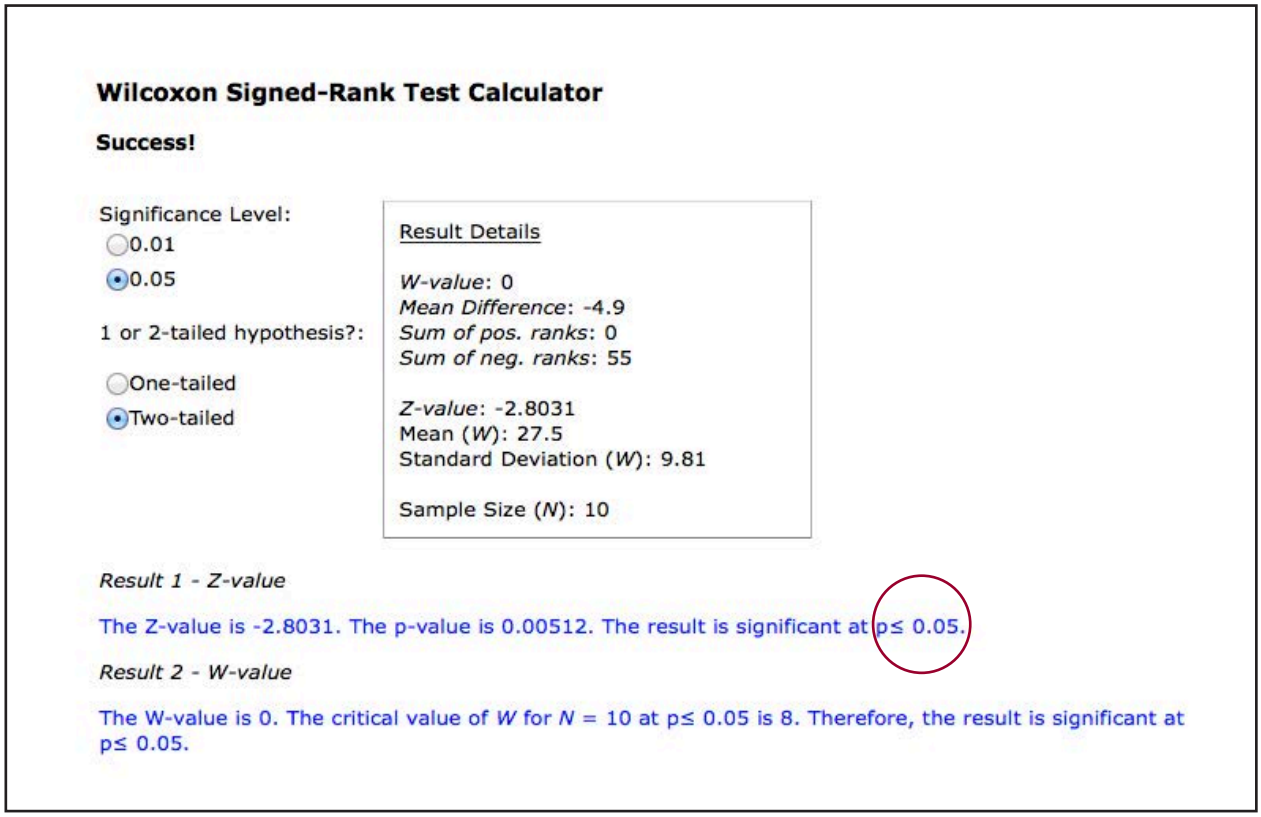
http://www.socscistatistics.com/tests/signedranks/Default.aspx

1. Click on "Take me to the calculator".
2. Enter the pre-test scores in the area "Treatment 1".
3. Enter the post-test scores in the area "Treatment 2".
4. Settings should read "Significance Level .05" and "Two-Tailed".
5. Click on "Calculate".

The results will appear at the bottom of the screen and show whether or not results are significant. The results will be reported as a Z-value and as a W-value, which are simply two different ways to calculate if your results are significant. The *p*-value associated with the W-value should be reported when dealing with a small calss size (<10). Larger class sizes should report the *p*-value associated with the Z-value.

With an Online Calculator—*continued*

FIGURE 3.11 WILCOXON SIGNED-RANK TEST ONLINE CALCULATOR



Talking about the Wilcoxon Signed-Rank Test

If you were asked to describe and interpret the statistical significance of the differences between your pre-test and post-test scores, you might state,

"The Wilcoxon signed-rank test was used to analyze pre-test and post-test scores. The results were statistically significant (p<.05). These results would occur by chance less than 5% of the time. The data suggest that instruction has had a significant impact on student performance."

Paired Samples *t*–test is a statistical technique that is used to compare two population means in the case of two samples that are related.

Paired samples *t*-test is used with pre-test/post-test measures, or when the samples are matched pairs.

The paired samples *t*-test should be used when students' have been 1) randomized to classrooms and, 2) the district's student scores are normally distributed or bell-shaped. This statistic compares two means using the same student in order to determine if the difference is statistically significant (if the *p*-value is < 0.05). A statistically significant result infers that the change was due to the teacher's instruction and not chance factors.

For example:

For this example, the teacher has matched each student's pre-test with his/her post-test scores.

TABLE 3.9 EXAMPLE STUDENT'S PRE-TEST/POST-TEST DATA

Student	Pre-test	Post-test
A	9	16
B	5	10
C	4	8
D	7	19
E	10	17
F	8	12
G	9	7
H	7	11
I	8	20
J	11	18

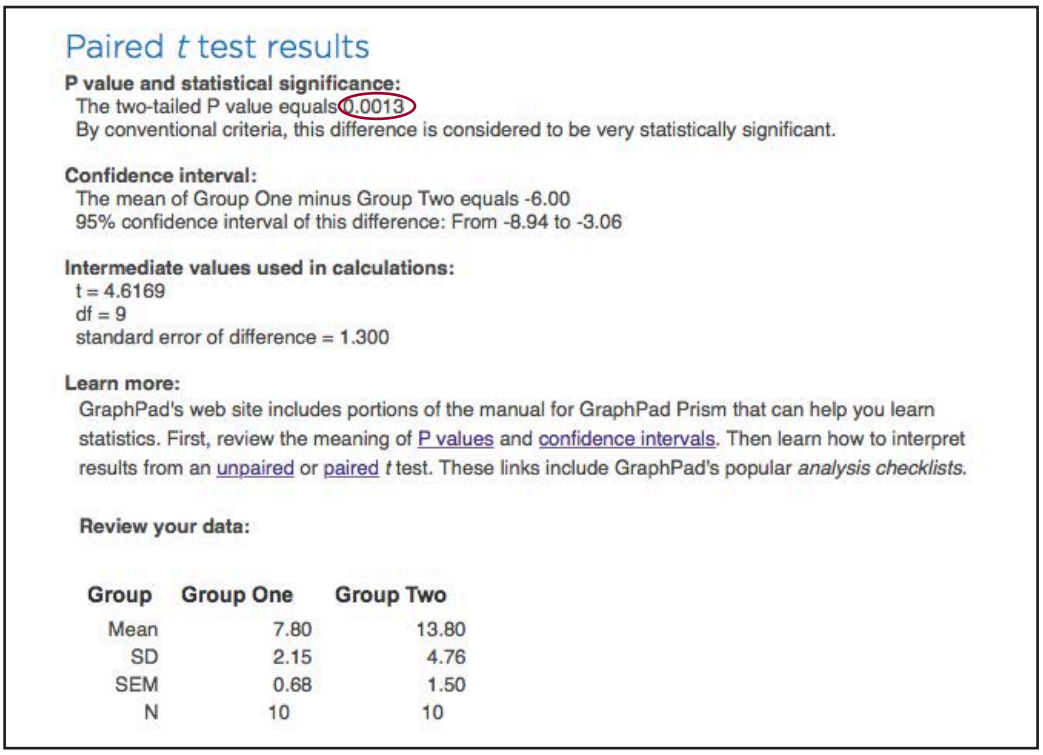
With an Online Calculator

Go to the following web page:

http://graphpad.com/quickcalcs/ttest1.cfm

1. Locate your list of pre- and post-test scores.
2. Enter the data. Pre-test data goes into the column for Group 1 and post-test data goes into the column for Group 2.
3. Select the paired *t*-test option.
4. Click on "Calculate now".

FIGURE 3.12 CALCULATE PAIRED *t*-TEST WITH AN ONLINE CALCULATOR USING DATA FROM TABLE 3.9



Paired *t* test results

P value and statistical significance:
The two-tailed P value equals 0.0013
By conventional criteria, this difference is considered to be very statistically significant.

Confidence interval:
The mean of Group One minus Group Two equals -6.00
95% confidence interval of this difference: From -8.94 to -3.06

Intermediate values used in calculations:
t = 4.6169
df = 9
standard error of difference = 1.300

Learn more:
GraphPad's web site includes portions of the manual for GraphPad Prism that can help you learn statistics. First, review the meaning of P values and confidence intervals. Then learn how to interpret results from an unpaired or paired *t* test. These links include GraphPad's popular *analysis checklists*.

Review your data:

Group	Group One	Group Two
Mean	7.80	13.80
SD	2.15	4.76
SEM	0.68	1.50
N	10	10

Talking about Paired *t*-test results

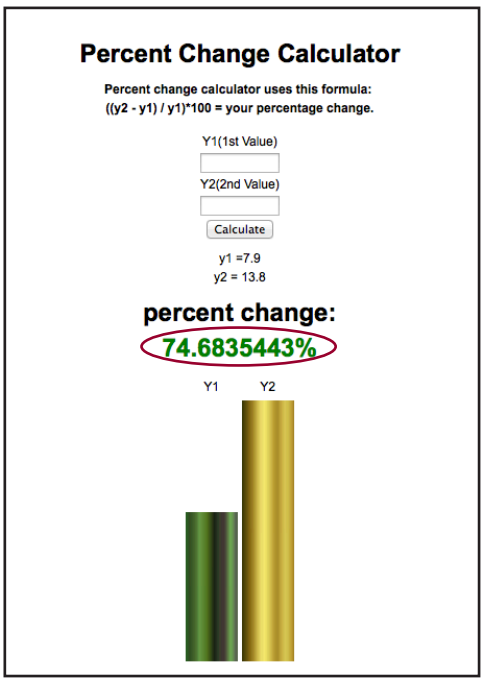
If you were asked to describe and interpret the statistical significance of the differences between your pre-test and post-test scores, you might state the following:

"The mean pre-test score was 7.8, SD = 2.15, and the mean post-test score was 13.8, SD = 4.76. This increase in performance was statistically significant, p= 0.0013".

For Additional Support

For additional support, you may also want to calculate the percent change using the mean values. Recall the above information on how to calculate percent change. In this example, the teacher used the Mean values from Group One and Group Two to obtain the percent change using the online calculator. The results are below:

FIGURE 3.13 CALCULATING THE PERCENT CHANGE



Talking more about Percent Change

With this additional information, the teacher demonstrated 74.7% percent change. To interpret this information, you may state:

"The students demonstrated an increase in scores by 74.7%. These gains were statistically significant and suggest that the instruction was effective."

Is there a way I can make a calculation that would mathematically demonstrate the amount of growth my students made?

STEP 5 *Determine the Magnitude of Student Growth*

Effect Size (Cohen's *d*): An effect size (ES) measures the magnitude of the teacher's instruction.

For example, an effect size of +1.0 would mean that a teacher's students moved one full standard deviation in achievement. It will indicate standardized differences which are reported in standard deviations between two means. Typically, an effect size is only calculated when there is a statistically significant result.

For example:

The formula for the calculation of the effect size is:

$$\frac{\text{Post-test Mean} - \text{Pre-test Mean}}{\text{Pooled Standard Deviation}}$$

An example for calculating ES with this Pre-test/Post-test data:

TABLE 3.10 EXAMPLE STUDENT'S PRE-TEST/POST-TEST DATA

Student	Pre-test	Post-test
A	9	16
B	5	10
C	4	8
D	7	19
E	10	17
F	8	12
G	9	7
H	7	11
I	8	20
J	11	18

Pre-test Mean = 7.80, Standard Deviation = 2.15

Post-test Mean = 13.8, Standard Deviation = 4.76

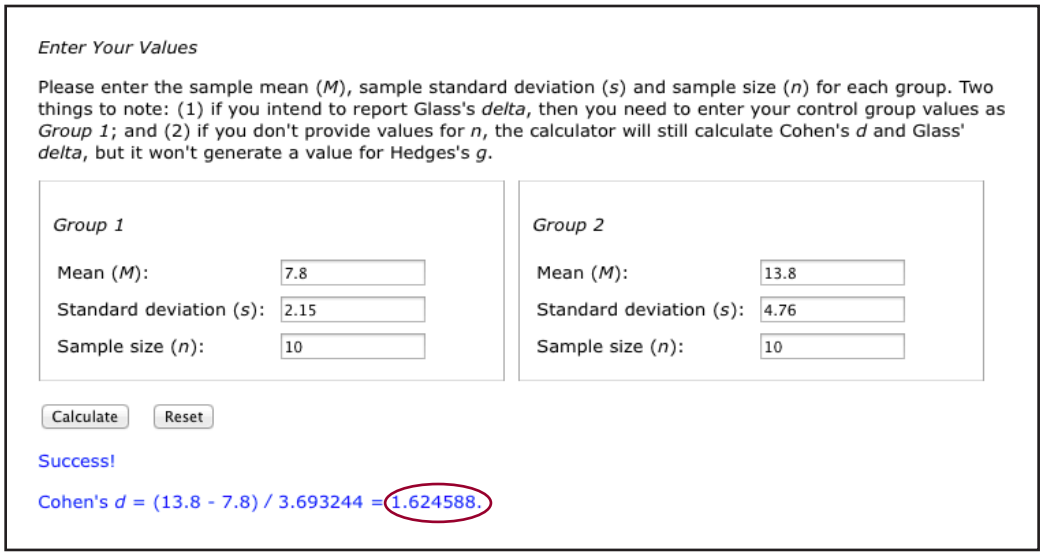
With an Online Calculator

Go to the following web page:

http://www.socscistatistics.com/effectsize/Default3.aspx

1. Enter the pre-test mean, standard deviation, and number of students into Group 1.
2. Enter the post-test mean, standard deviation, and number of students into Group 2.
3. Click on "Calculate".
4. Use the table below to interpret.

FIGURE 3.14 CALCULATE EFFECT SIZE WITH AN ONLINE CALCULATOR



Cohen (1988) interprets effect size as follows:

Effect Size	
Small	0.2
Medium	0.5
Large	0.8

Issac and Michael (1997) note that an effect size greater than 0.33 "to have a practical significance that is large enough to make a practical difference" (p. 209).

Talking about Effect Size

Here is an example of how an effect size could be used to create a rubric:

TABLE 3.11 EFFECT SIZE RUBRIC

	Highly Effective	Effective	Minimally Effective	Ineffective
Effect Size	> 0.80	0.33 - 0.79	0.20 - 0.32	< 0.20

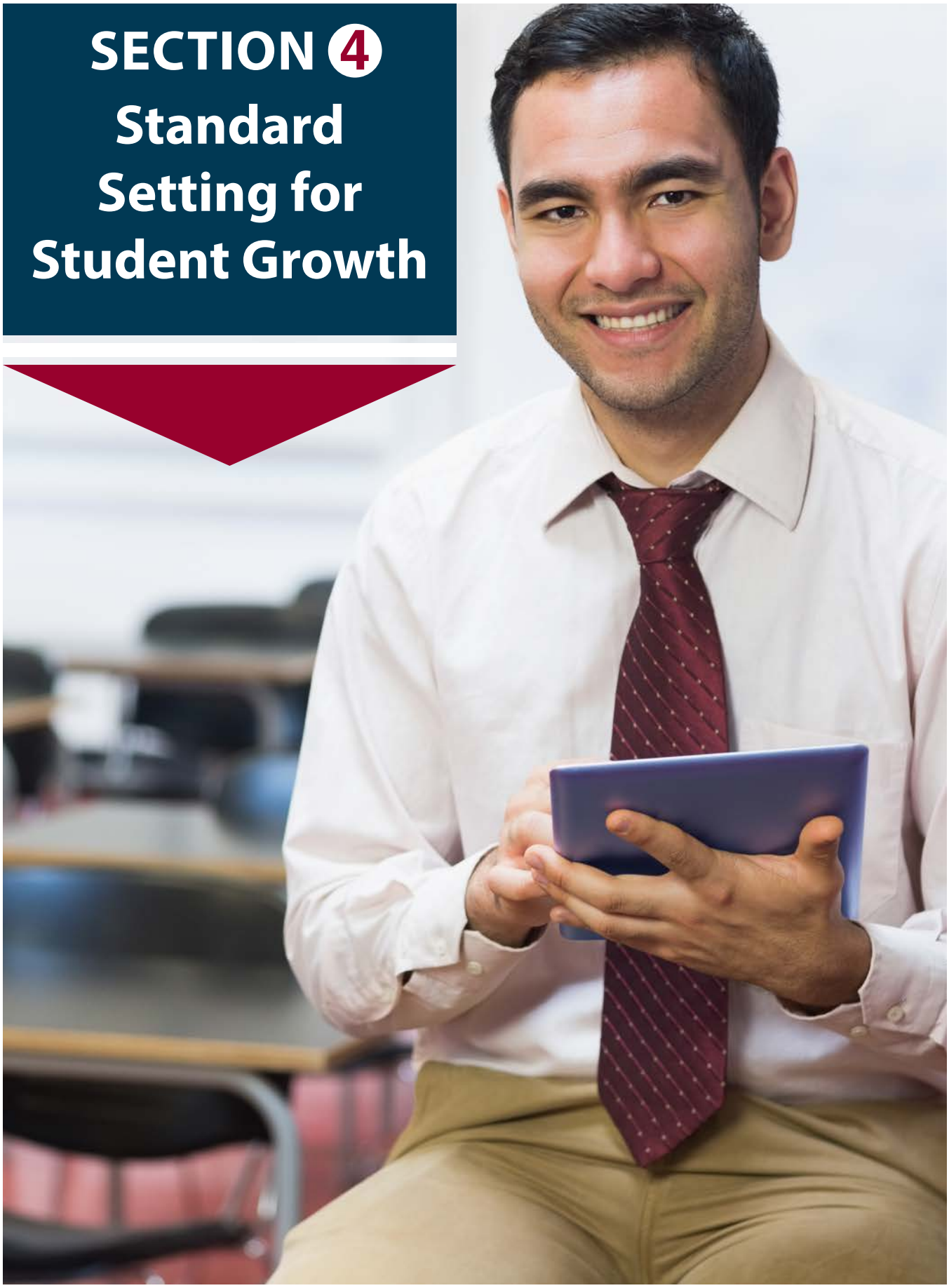
If you were asked to describe and interpret the statistical significance of the differences between your pre-test and post-test scores, you might state the following:

"The mean pre-test score was 7.8, SD = 2.15, and the mean post-test score was 13.8, SD = 4.76. This 74% increase in performance was statistically significant, p= 0.0015. The effect size was 1.62 which can be considered to be a large effect size."

SUMMARY

This section of the guidance document has provided the reader with user- friendly methods to demonstrate student growth. These methods should help with the understanding of data by identifying the central tendency, which describes how typical students perform. Also, this guidance will allow the reader to calculate and analyze the amount of variability in the performance of students, and with confidence, define the range of true scores. The on-line tools will assist the reader with determining if the change in students' scores are statistically significant. In addition, educators will be able to demonstrate the magnitude instruction has had upon student learning. Finally, with each calculation, there are examples of how to summarize and describe findings in order to communicate results with colleagues and administrators.

SECTION 4
Standard Setting for Student Growth



The standard setting process is research-based. It is a valid and reliable method for bringing stakeholders together with the intention of assigning meaning to achievement data.

Standard setting is a process for defining cut scores that is based on local school district achievement data. The standard setting process is used to find the cut scores for teacher evaluation. The purpose of the cut score is to indicate the threshold for each teacher evaluation category. The cut score establishes the difference between effective and minimally effective teacher performance based on student growth data.

The standard-setting steps include:

1. **Assemble the Team**
2. **Document the Process**
3. **Establish Data Sources for Student Growth**
4. **Clarify the selected Model of Growth Assumptions**
5. **Write Performance-Level Descriptions**
6. **Train Committee Members**
7. **Implement the Selected Standard-Setting Procedure**
8. **Evaluate the Process**
9. **Share the Recommendations**

STEP 1 *Assemble the Team*

The standard setting process involves the creation of **a committee of stakeholders whose task is to review the data and establish the cut scores for teacher effectiveness categories.** Committee members may include administrators, teachers, union representatives, and appropriate community members. Representatives from special populations should be included as well. The committee is formed at the district level to define cut scores that will be used across the school district. There may be reasons to apply this process to building data with building-based teams of educators and stakeholders.

- District leadership identifies the facilitator and the roles of committee representatives.
- Identify the facilitator who is trained in the data, the collaborative inquiry process, and the standard setting procedure. (See Appendix A)
- Assign committee representative to appropriate groupings, based on grade level, content area, and relevant considerations.

STEP 2 *Document the Process*

Establish the agenda, meeting norms, and schedule for the committee (See Appendix B).

Define the rules for the data to be included or excluded from the data sets. For example,

- student attendance rates,
- classroom enrollment,
- demographics,
- missing data,
- extenuating circumstances

Also consider how to handle incongruent data, such as how to round cut scores.

Clearly describe the selected standard setting procedure and the forms that will be used to collect data from the committee members.

Test the standard setting method using local data as a trial run for how the committee will proceed with this work.

STEP 3 *Establish Data Sources for Student Growth*

To implement the standard-setting process, the district will need to intentionally identify the data sources that will be used for the student growth component of the evaluation. The number and type of data sources utilized should consider current requirements of the Michigan legislation, but ultimately remains a district decision. Recommended practice would be to **include multiple measures of student growth with procedures designed to assure that the measurement of growth is comparable and consistent across grade levels and content areas.**

Table 1 provides a framework for summarizing the data sources, including consideration of grade level, content or essential/power standards, type of gain or growth score, and the instructional interval.

The instructional interval is an important consideration for such instructional time factors as semester classes or the timeframe of the evaluation benchmark periods. It is important to consider time to teach and opportunity to learn when comparing a teacher's growth data to criteria or norms that are developed based on an annual time frame. If, for example, a teacher has only had one semester to teach an advanced placement course that is used for student growth and the advanced placement test was developed based on norms for students who had one year of instruction, this would place the teacher at a disadvantage in showing comparable achievement.

TABLE 4.1 SUMMARIZING DATA SOURCES FOR GROWTH MEASUREMENT

Data Sources	Grade Level/s	Content	Type of Gain or Growth Score	Instructional Level
State Assessment Data				
District Purchased Assessment				
District Developed Assessment				
Classroom Assessment				
Collection of Evidence ***(e.g., portfolios, IEP goals)***				

STEP 4 *Clarify the Selected Model of Growth Assumptions*

According to Michigan legislation, growth is defined as the difference in student achievement measured at two points in time. There is no discussion of basing student growth on trajectories toward proficiency. **When implementing a standard setting process, the district should clarify assumptions about student growth and the explicit assessments to be used in this evaluation process.** By defining assumptions about growth in the standard setting process, there is the opportunity to differentiate the measurement of growth to address individual student needs. Clarification of assumptions about growth are integral to student learning objectives in teacher evaluation. Table 2 provides examples of assumptions about student growth that may need to be clarified with committee members.

TABLE 4.2 EXAMPLES OF ASSUMPTIONS ABOUT STUDENT GROWTH

Growth Models	Description
Catch Up Growth	Student is not at benchmark and needs to make catch-up growth to get to benchmark.
Keep Up Growth	Used to be known as the "bubble students". The achievement is near to benchmark.
Move Up Growth	Students are at or above benchmark and can be challenged to improve or move up to higher levels of achievement.

Adapted from Fielding, Lynn, Kerr, Nancy, and Rosier, Paul. *Annual Growth for All Students Catch-Up Growth for Those Who Are Behind.* The New Foundation Press, 2007.

STEP 5 *Write Performance Level Descriptions for Student Growth*

The performance level descriptions **define the parameters of student growth measurement that are then referenced by the team to determine data-based cuts for the categories of teacher effectiveness.** Table 4.3 (see page 69) provides examples of growth performance level descriptions that may vary based on the data type. The district team will need to consider their own data to develop meaningful performance level descriptors for your context. Remember, it is the task of the standard setting committee to assign data points, based on scores, percentages, percentiles, etc., to these descriptors.

STEP 6 *Train Committee Members*

Train the committee members on the goal of the committee. **They are there to set the cut scores for teacher effectiveness ratings based on district identified data using the performance level descriptors as the criteria.** Work with committee members to set aside assumptions and to focus on the purpose, tasks and outcomes. Train the committee members on the data reports they will be reviewing, building assessment literacy. Train the committee members on the standard setting protocol. Support the committee members with good facilitation using collaborative norms (See Appendix B).

TABLE 4.3 EXEMPLAR PERFORMANCE LEVEL DESCRIPTIONS BASED ON DIFFERENT DATA SOURCES

Data Source	Ineffective	Minimally Effective	Effective	Highly Effective
State Assessment Data *(Example of performance level description for Student Growth Percentile data)*	Student growth in the ineffective category is defined by MDE as...	Student growth in the minimally effective category is defined by MDE as...	Student growth in the effective category is defined by the MDE as….	Student growth in the highly effective category is defined by MDE as…...
District Purchased Assessment *(Example of Norm-Referenced Growth Data or Gain Scores)*	Student growth in the ineffective category is defined by parameters of an instructional interval (e.g., one year) and **decline, no, or minimal test growth** that is below district identified norms or standards. *(*Considering confidence intervals or standard error)*	Student growth in the minimally effective category is defined by parameters of an instructional interval (e.g., one year) and **marginal** test growth that is below district identified norms or standards. *(*Considering confidence intervals or standard error)*	Student growth in the effective category is defined by parameters of an instructional interval (e.g., one year) and **meets** test growth that is consistent with district identified norms or standards. *(*Considering confidence intervals or standard error)*	Student growth in the highly effective category is defined by parameters of an instructional interval (e.g., one year) and test growth **exceeding** district identified norms or standards. *(*Considering confidence intervals or standard error)*
District Developed Assessment *(Example of Common Assessment or Mastery of Standards)*	Student growth in the minimally effective category is defined by **mastery of few** grade level standards at grade expectancy as defined by a local benchmark assessment.	Student growth in the minimally effective category is defined by **mastery of some** grade level standards at grade expectancy as defined by a local benchmark assessment.	Student growth in the effective category is defined by **mastery of several** grade level standards at grade expectancy as defined by a local benchmark assessment.	Student growth in the highly effective category is defined by **mastery of most** grade level standards at grade expectancy as defined by a local benchmark assessment.
Classroom Assessments *(Example of Student Learning Targets/Objectives/ I Can Statements)*	Student growth in the ineffective category is defined by **mastery of few** learning objectives.	Student growth in the minimally effective category is defined by **mastery of some** learning objectives.	Student growth in the effective category is defined by **mastery of several** learning objectives.	Student growth in the highly effective category is defined by **mastery of most** learning objectives.
Collections of Evidence *(Examples may include Portfolios, Capstone Projects, IEP Goals)*	The **evidence is not present to specific IEP content area goals** or IEP goals do not include content area objectives. Students are unresponsive. Appropriate supports are not provided to the students.	**Evidence of progress is minimal in relation to IEP content area goals**, or IEP goals are not related to content area objectives. Students show little or no evidence of performance of IEP related goals; students perform tasks in a limited range of contexts, tasks are not meaningful or are not age-appropriate, failure to use appropriate supports.	**Evidence of progress demonstrates mastery of several IEP objectives**; IEP goals are relevant to content area objectives. The students show some evidence of performance of goals, in a limited variety of settings with opportunity for some interactions with peers, uses age-appropriate materials to perform some meaningful tasks with real-world applications, uses some appropriate supports.	**There is ample evidence of mastery of IEP goals**, which are clearly related to content area objectives. The student shows considerable evidence of performance related to goals, performs tasks in a variety of settings, engages in social interaction with a diverse range of age-appropriate peers, uses age-appropriate materials to perform meaningful tasks in a real-world context, and consistently uses appropriate supports.

STEP 7 *Implement the Standard Setting Procedure: The Vertical Scaling Method*

One recommended method of standard setting is based on the concept of vertical scaling. **Vertical scaling provides the opportunity to establish a system of cut scores across grade levels within the content area.** In this method the standard setting panel is comprised of representatives from adjacent grades (e.g., K - 1 - 2) who are knowledgeable in the content area and method of assessment for measuring growth. Ideally the vertical scaling teams would have grade level representatives in three year clusters, such as Grades K - **1** - 2; 1 - **2** - 3; 2 - **3** - 4, etc. The bolded grade is the primary focus of the standard setting.

Benefits to this method include the **opportunity for teachers to vertically align and identify non-negotiable standards and expectations for each grade level.** Consider beginning with a limited and specific amount of standards (e.g., 5 - 10 non-negotiable standards within the grade level for that content area) that can be refined in future iterations. For contexts of special education or personal curriculum, adjustments can be made to the amount of standards, which may be 1 - 5 standards).

Even with a rigorous standard setting process, conflicts and inconsistencies may occur for a variety of reasons that make this challenging work. These challenges have been recognized in the research on standard setting with growth measures.

SEVEN STEPS TO IMPLEMENT THE VERTICAL SCALING METHOD

1. **Prepare the data by grade level and content area.** Collect historical and trend data for up to 3 years, if available. Organize the data by year, grade and content area.

2. **Assign participants to grade level and content groupings, including adjacent grades.**
   (Grades K - **1** - 2; 1 - **2** - 3; 2 - **3** - 4, etc.).

3. **Develop the Performance Level Descriptors.**
   Establish the performance level descriptors, considering the historical data. Where evidence are not available, rely on generalization from other measures and discussions with content experts and stakeholders to shape the performance model.

4. **Train the participants on the data they will be reviewing, norms, and the standard setting process.**
   Review the performance level descriptors in detail to assure all participants are clear on the criteria for setting cut scores.

5. **Conduct the Standard-Setting Sessions with Cross Grade Participants.**
   Present the historical data in multiple rounds of the standard-setting procedure, cross grade or cross subject panels or in meta-panels.
   If cut scores are to be articulated across grades, it seems reasonable that the cut scores for a given grade be considered by individuals with strong interests in the performances of students in the adjacent grades. Where possible, all grade review should be included for at least one round, at or near the final round of review.

6. **Individual Setting of Cut Scores.**
   The task is for the participant to answer the question, *"Is it likely that the minimally qualified teachers at this category (e.g., Effective) would have a median score in this range?"* Participants are given data sets with scores vertically scaled from lowest grade to highest grade and from lowest score to highest score. Each participant independently reviews the data and marks the cut points, applying the performance level descriptors to the data.

TABLE 4.4 EXAMPLE OF PARTICIPANT RATING CHART FOR SETTING CUT SCORE USING GROWTH PERCENTILES

	Growth Percentile	Ineffective	Minimally Effective	Effective	Highly Effective
Grade 4	90 - 100				
	80 - 89				●
	70 - 79				
	60 - 69				
	50 - 59			●	
	40 - 49				
	30 - 39		●		
	20 - 29				
	10 - 19	●			
	0 - 9				

The chart is an example of how to organize the data ranges so that the participant would mark the score point or cut score for each category. For example, with fourth grade data, this participant set a cut score of 50 Growth Percentile as Effective.

7. **Summarize each participant's ratings and provide them with the feedback.**
   The purpose of this feedback is for the participants to see their agreement, to consider the impact, to clarify understanding of the data or performance level descriptors and to reach consensus. If discrepancies continue to exist, consider repeating the process. Another option would be to submit the recommendations for independent review as described below.

TABLE 4.5 EXAMPLE OF SUMMARY CUT SCORES FROM INDIVIDUAL PARTICIPANTS

Participant	Ineffective	Minimally Effective	Effective	Highly Effective
1	20 - 29	40 - 49	60 - 69	90 - 100
2	10 - 19	30 - 39	50 - 59	80 - 90
3	10 - 19	40 - 49	50 - 59	90 -100
4	60 - 69	70 - 79	80 - 89	90 - 100
5	50 - 59	60 - 69	70 - 79	90 - 100

As part of providing the participants with feedback, consider providing data that demonstrates the impact, based on recommended cut points. For example, provide them with the number or percent of teachers who would fall within each category based upon their cut scores. This additional review of the impact data will inform them as whether their recommendations will lead to desired outcomes when put into practice.

STEP 8 *Evaluate the Process*

Ask the committee members to provide reflection and feedback on the process and outcome of the standard setting. The information is essential to check for validity and to demonstrate the degree of support for the cut scores.

Ask them to answer the questions:

1. **Was it reasonable?**
2. **Can it be replicated?**

STEP 9 *Share the Recommendations*

Prepare a summary report to be shared with district constituents that includes the following components:

- Brief overview of the standard setting process
- Committee members involved in the standard setting
- Data and impact
- Cut scores for effectiveness categories
- Revisions per state law
- Include recommendations for future review committees

When the process is completed, the standard setting process may yield a rubric that includes the data points to be used for categorizing student growth data. An example of what this rubric might look like is demonstrated in Table 4.6:

TABLE 4.6 EXAMPLE OF STUDENT GROWTH RUBRIC FOR TEACHER EFFECTIVENESS CATEGORIES

Ineffective	Minimally Effective	Effective	Highly Effective
Less than X% of students demonstrated adequate growth on the ABC assessment. The standard setting team, based upon trend data at the district/grade or building level, will determine growth on an annual basis using three years of the most recent data (if available).	**Between X% and X% of students demonstrated adequate growth** on the ABC assessment. The standard setting team, based upon trend data at the district/grade or building level, will determine growth on an annual basis using three years of the most recent data (if available).	**Between X% and X% of students demonstrated adequate growth** on the ABC assessment. The standard setting team, based upon trend data at the district/grade or building level, will determine growth on an annual basis using three years of the most recent data (if available).	**At least X% of students demonstrated adequate growth** on the ABC assessment. The standard setting team, based upon trend data at the district/grade or building level, will determine growth on an annual basis using three years of the most recent data (if available).

FOLLOW UP WITH REVIEW AND ADJUSTMENTS

The cut scores may be reviewed and revised by officials or in situations in which the standard setting was disparate, a third party must make the final determinations. Furthermore, as there are changes with legislation, assessments, or observed data trends, it would be appropriate to revisit the standard setting process to make the needed adjustments to cut scores.

STANDARD SETTING PROTOCOL FOR EDUCATIONAL SETTINGS THAT ARE BASED ON COLLECTION OF EVIDENCE

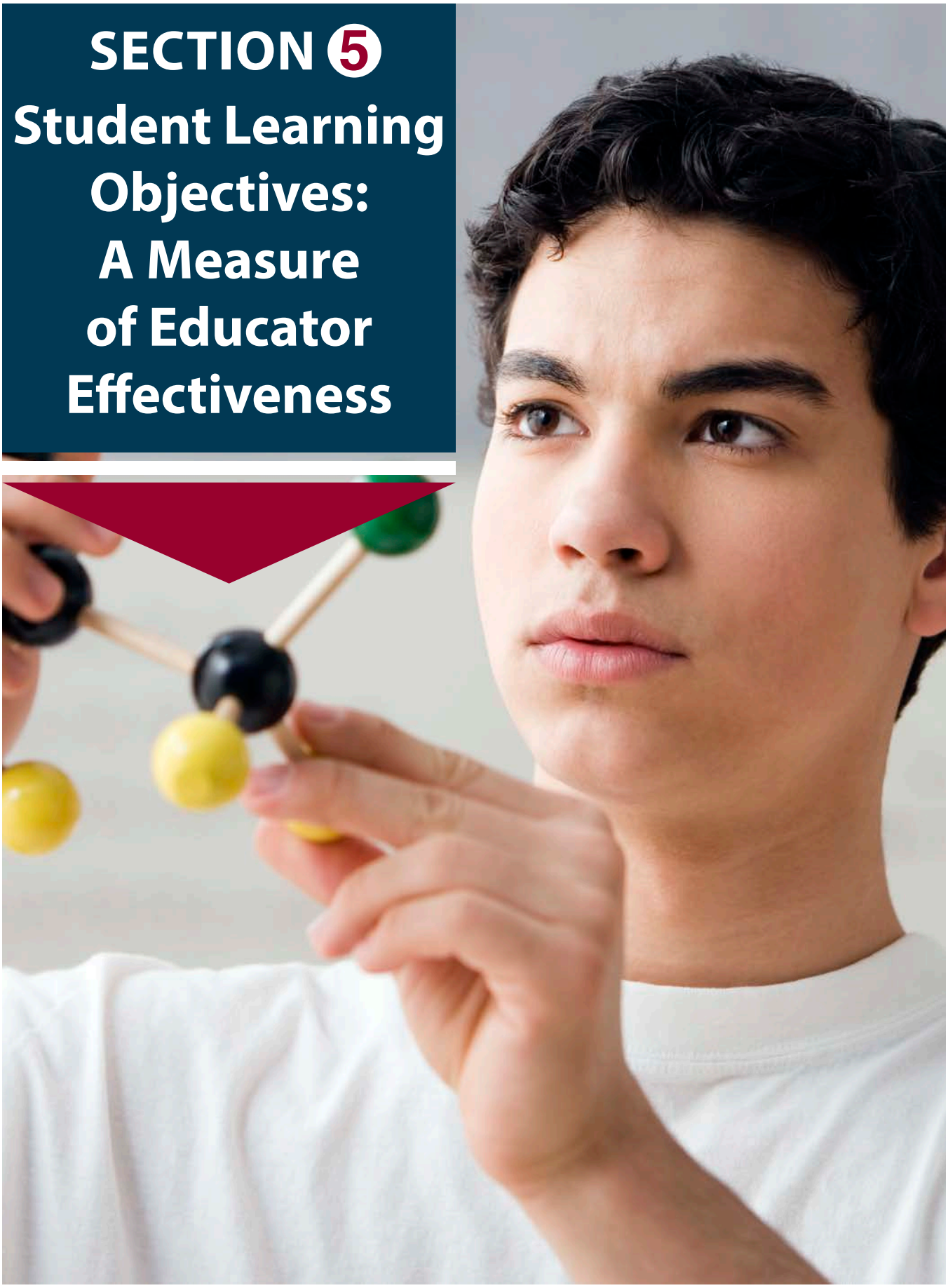
Examples of settings in which achievement measures are based on collection of evidence may include: Alternative Education Schools, Career Technical Schools, Special Education Center Programs, Co-Op Programs, Special Disciplines, e.g., the Arts.

Examples of collection of evidence may include Portfolio Assessment Systems, IEP Goals, or other Performance data. The preferred method of standard setting with collections of evidence is the Generalized Holistic Method. In this method, the collections of evidence are scored by established criteria along essential dimensions that align to designated standards. Each collection of evidence is then independently rated as meeting performance level descriptions for the teacher effectiveness categories. Cut scores are based on mid-points between adjacent categories.

General steps to take with collection of evidence are described:

1. **Identify the artifacts or evidence** to be included.
2. **Review alignment of the artifacts/evidence to the state content standards.**
3. **Establish a scoring criteria for the collection of evidence**, appropriate to the skills and instructional objectives/benchmarks for the population or discipline.
4. **Create a rubric** for the scoring of the collection of evidence.
5. **Prepare the Performance Level Description** based on your data sources and scores.
6. **Follow a Generalized Holistic Method of Standard Setting.**
    a. **Organize the collection of evidence** by content area/grade range.
    b. **Establish rules for data to include/exclude and how to handle score incongruities**, e.g., rounding of scores.
    c. **Each collection of evidence is reviewed by 3 raters.**
    d. **Each participant reviews up to 8 collections of evidence,** making independent ratings. Using the rubric scores of the collections of evidence to establish cut points for effective, minimally effective, highly effective, and ineffective categories. (See Appendix C)
    e. **Provide data analysis and feedback to the participants,** including impact data.
    f. **Calculate the cut scores** based on mid-points between adjacent categories.
    g. **Facilitate discussion on the results,** emphasize variability in ratings.
    h. **Arrive at consensus or defer for independent review.**
7. **Assess participant ratings.**
8. **Summarize and report out.**

SECTION 5
Student Learning Objectives: A Measure of Educator Effectiveness



INTRODUCTION

As states and districts develop or adopt approaches to measuring student growth, they face many challenges, including subject matter or grade levels not being tested on the state summative assessment. To compound these challenges, when student growth measures are utilized to assign effectiveness labels, fairness and comparability become critical to employment decisions. The use of Student Learning Objectives (SLO) has proven to be a promising approach to measuring student achievement for purposes of conducting educator evaluation. As of 2016, over half of the states in the US have either required, recommended, or identified student learning objectives as a measure of student growth for use in educator evaluation. The SLO practice began in Denver Public Schools in 1999 and continues to be expanded, studied, and used successfully by many districts. The purpose of this section of the guidance document is to provide relevant information designed to assist the reader with the decision-making process, as well as understand concepts related to basic implementation strategies.

WHAT TO CONSIDER

When considering the use of Student Learning Objectives, it is important for district decision makers to explore the:

- **purpose** of SLOs,
- **steps** in the SLO development process,
- **function** within the evaluation cycle,
- **current research findings, as well as benefits and challenges** associated with the SLO process

PURPOSE

The Purpose and Use of Student Learning Objectives: What is a SLO?

A Student Learning Objective (SLO) is a measure of a teacher's impact on student learning within a given interval of instruction. **An SLO is a measurable, long-term academic goal informed by available data that a teacher or teacher team sets at the beginning of an instructional interval for all students or subgroups of students.** The teacher and students work toward the SLO growth targets throughout the instructional interval and use interim, benchmark, and formative assessments to measure progress toward the goal. At the end of the interval of instruction, the teacher meets with a principal or building team to discuss attainment of the SLO and determine the teacher's impact upon student learning.

WHAT ARE THE NECESSARY COMPONENTS OF A SLO?

The development and use of Student Learning Objectives is part of an overall process that exists within the context of a larger teacher evaluation system. Though the actual SLO can take a variety of forms, they typically contain the following components:

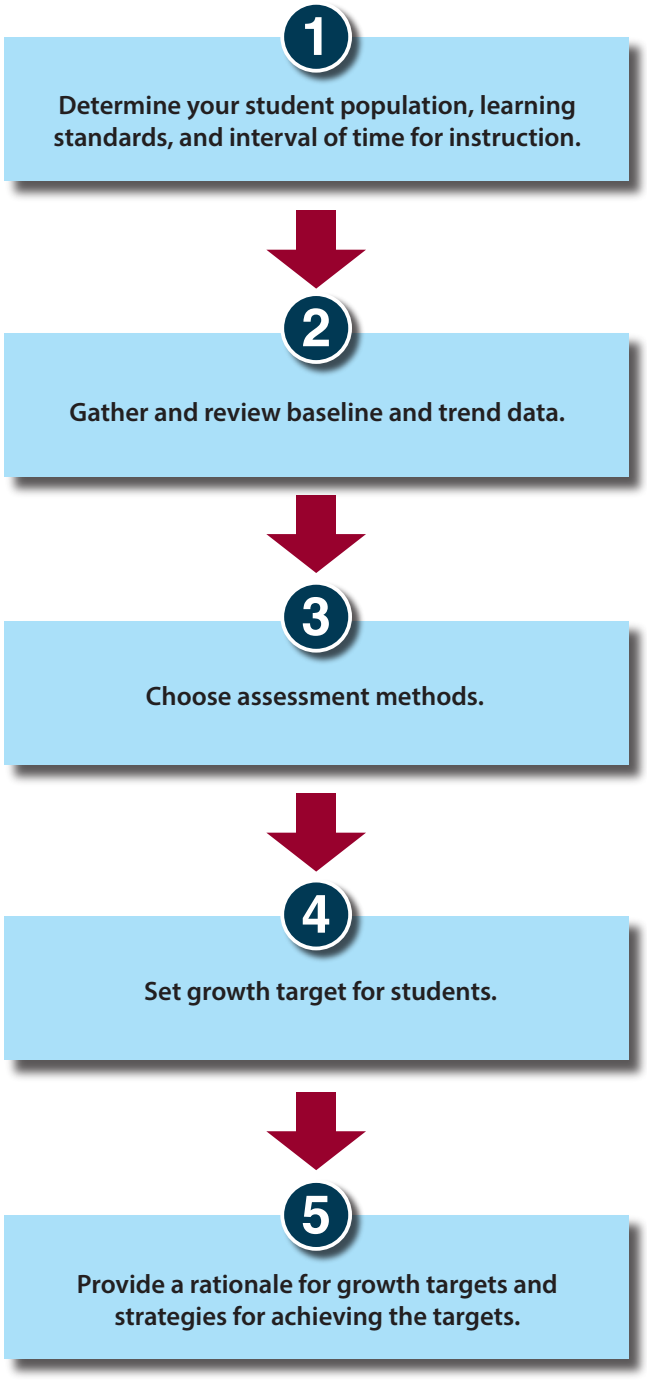
Components of a SLO	
Student Population	Characteristics that describe the student population served by the SLO, including the number of students that have special needs relevant to the SLO
Learning Standards	A list of key standards that are connected to the learning content
Interval of Time	The interval of instruction based on the course structure that includes the start and end dates
Baseline Data	A description of the data reviewed in the creation of the SLO and an explanation of how the data supports the SLO
Assessment Choice	The assessment instrument that will be used to measure the outcome of the SLO
SLO Growth Targets	The quantitative targets that will demonstrate achievement of the SLO—targets should be rigorous yet attainable and can be tiered for specific students
Rationale	Precise statements that describe student needs and explain in detail how the baseline and trend data informed the development of growth target(s)

PROCESS

The SLO Development Process

The writing of effective Student Learning Objectives involves a process that requires and supports assessment literacy and instructional practices. In order for the SLO process to be meaningful, educators should be well-trained regarding development, execution, implementation, and measurement. They should be given models, samples, clear instructions, and supported practice with feedback. The SLO development process can provide a structure to promote collaboration and reflective practice among teachers. Allowing and structuring Professional Learning Community (PLC) time during the school day for these conversations to take place shows the value and commitment to intentional and purposeful planning and data-based goal setting (See Appendix D for template.)

FIGURE 5.1 THE SLO DEVELOPMENT PROCESS



Assessment Choice in the SLO Process

At the core of the SLO process we must consider the following questions:

- **Is the assessment appropriate to be used when measuring student growth?**
- **Have I selected assessments that accurately reflect the standards I will teach throughout an appropriate interval of time?**

Identifying rigorous assessments that are aligned to the content is a critical and often unanticipated challenge encountered when developing an SLO. It is important to note that SLOs are only as good as the baseline, trend, and assessment data upon which they are built. Most states strongly recommend that districts not use assessments created by individual educators who may lack expertise in assessment writing and alignment. If districts choose to use locally developed assessments, those assessments should be carefully developed by teacher teams and then reviewed at the district level by content experts or staff with training in assessment literacy.

Assessment options may include:

- Performance-based assessments, such as presentations, projects, and tasks graded with an approved rubric
- Portfolios of student work, with samples throughout the year that illustrate knowledge and skills before and after a learning experience. A rubric is also needed for this type of assessment.
- Nationally normed tests created by vendors
- Educator, school-created, or district-created tests

Educators should identify assessments that are:

- **Aligned to state standards and the SLO growth target** (meaning that they measure the skills or content addressed by the SLO)
- **Reliable**, meaning that they produce accurate and consistent results
- **Valid**, meaning that they measure what they are designed to measure
- **Realistic** in terms of the time required for administration

(See the ***Developing and Selecting Assessments of Student Growth for Use in Teacher Evaluation Systems*** section of this guidance document for more detailed information.)

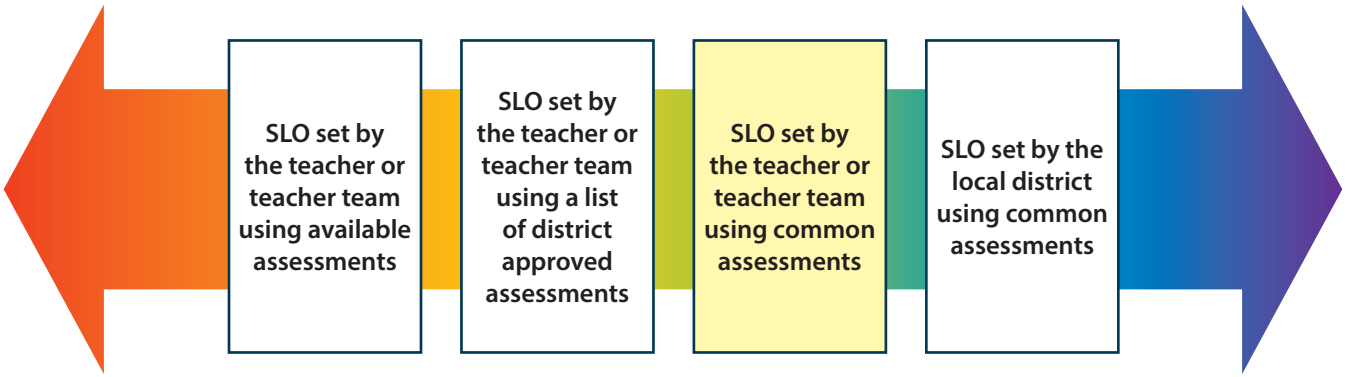
TABLE 5.1 CHECKLIST FOR SELECTING ASSESSMENTS TO MEASURE STUDENT GROWTH THROUGH THE SLO PROCESS

Course/Subject:			
Assessment Name:			
Alignment to Standards: ***Is the Learning Objective clearly reflected in the assessment measure?***			
YES	**SOMEWHAT**	**NO**	
			Test items align to the standard(s) addressed in the SLO.
			The assessment measure addresses the full range of topics and skills included in the SLO.
			The focus of the assessment reflects the focus of the curriculum and standards.
			The items or task represent the range of cognitive thinking (Depth of Knowledge) required during the course.
Range: ***Will students be able to demonstrate growth on this assessment?***			
YES	**SOMEWHAT**	**NO**	
			The test includes items that cover requisite knowledge and skills from prior years and appropriate content-relevant items that will challenge all students.
			Appropriate accommodations are available to make the assessment accessible to all students.
			Test items cover knowledge and skills that will be of value beyond the school year.
Validity and Reliability: ***Is the assessment a valid and reliable tool for the intended purpose?***			
YES	**SOMEWHAT**	**NO**	
			The assessment uses grade-level appropriate vocabulary.
			Items or tasks are written clearly and concisely.
			Clear scoring rubrics, guidance, and/or student exemplars exist for open-ended questions or performance-based assessments.
			Assessment administration is standardized across all classes.
			Assessment provides information to help improve teaching and learning.
Comments:			

Adapted from the *Ohio Department of Education Guidance on Selecting Assessments*

There is a spectrum of approaches that districts can take when creating their guidelines regarding assessment selection within the SLO process. District leaders must consider the balance between teacher agency and comparability when deciding the approach that best fits their philosophy and need. **Teacher agency refers to the amount of autonomy teachers have in creating SLOs, specifically in terms of assessment choice and setting growth targets. Comparability refers to how similar SLOs are among teachers who teach the same grade or subject across classrooms, buildings, or the district.**

FIGURE 5.2 SLO ASSESSMENT APPROACHES



INCREASING TEACHER AGENCY

INCREASING SLO COMPARABILITY

Adapted from: Lachlan-Hache,L. Matlach, L, Reese, K., Cushing, E., and Mean, M. *Student Learning Objectives: Early Lessons*. Teacher Incentive Fund Technical Assistance Network, 2011.

On the left side of the above spectrum, teachers decide what baseline data to use. They select their own growth targets and decide or create the assessment to measure that growth. This approach allows for a great deal of autonomy. At the right side of the spectrum, the district decides what baseline data will be utilized, what growth targets should be set, and how they will be measured. This approach decreases teacher agency, but allows for more comparability. When creating guidance, **most states—including the state of Michigan—opt for an approach that encompasses both teacher agency and comparability**, such as the approach highlighted above, whereby teacher teams are encouraged to work together to set growth targets using district created or vendor purchased common assessments.

Considerations for Setting Growth Targets

Educators must understand assessment data and identify student achievement trends to set rigorous yet realistic student growth targets that align with state standards, district priorities, and course objectives. These growth targets should include specific indicators of growth that demonstrate learning between two points in time. (See the *Standard Setting for Student Growth* section of this guidance document for a more detailed and robust explanation of this process)

TABLE 5.2 GROWTH TARGET TYPES

GROWTH TARGET TYPES	
Course-level SLOs	Growth targets that encompass all students in a class or course.
Tiered targets within a course-level SLO	Growth targets that reflect a range in skill and ability within a class or course. A teacher should consider tiered targets if a handful of students have already demonstrated mastery of the standards being assessed. This group may have a separate growth target than the rest of the class.
Targeted SLOs	Growth targets that are adapted to subgroups of students based on need. Writing separate SLOs can focus on specific content or specific groups of students who need additional attention or more focused attention in particular skill areas.

The Use of SLOs within the Evaluation Cycle

The developing of student learning objectives is part of the SLO evaluation cycle, which is a deliberate process that fits into the larger context of the educator evaluation system. The cycle outlines a sequence of five events that start with the SLO development and end with a final review, scoring and conversation with the teacher and the evaluator. Ultimately, all of the steps in the SLO evaluation cycle can lead to reflection, collaboration, and the improvement of educator practice.

FIGURE 5.3 THE SLO EVALUATION CYCLE

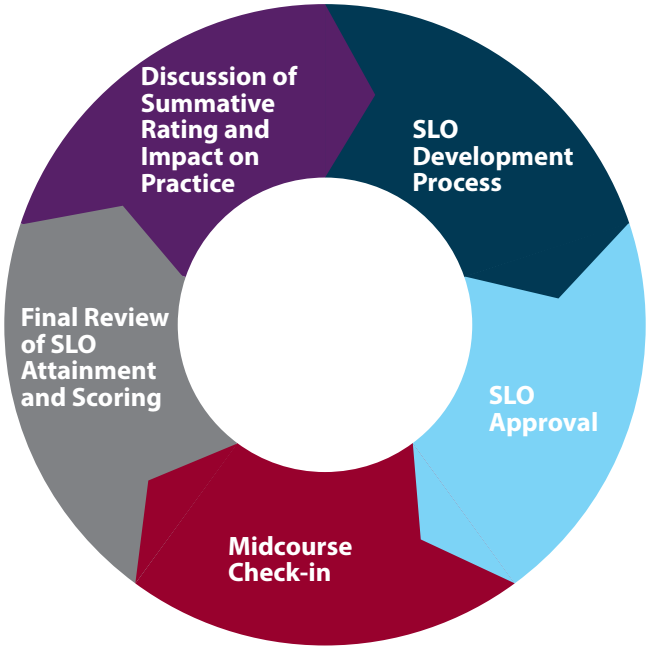


FIGURE 5.4 THE SEQUENCE OF EVENTS IN THE SLO EVALUATION CYCLE

The Benefits of Using a SLO Process

WHY USE SLOS?

The use of Student Learning Objectives has many benefits. The SLO process promotes learning; reflective teaching practices, the retention of teachers, and it aligns with many quality administrative and school improvement high impact initiatives.

The SLO process promotes learning by:
• providing a system for educators to measure and monitor student learning and the academic growth of their students
• adapting to changes in curriculum, changes in student population, and student needs
• allowing for the use of multiple assessment measures
• maximizing the use of assessment data to target learning needs
The SLO process promotes quality instruction by:
• encouraging educators to set goals for students, using data to determine student progress and making important decisions regarding instructional practice
• guiding educators through the steps of using the data to impact instruction
• allowing all educators, including non-tested grades and subjects, to demonstrate their impact on student learning
• allowing educators to focus on objectives that are most relevant for their specific student population and content areas
• encouraging educators to use more evidence-based practices and providing clear and measurable connections to instruction.
• providing the opportunity for the creation of realistic goals based on student population and educators' specific course loads
• promoting collaboration and opportunities for educators to reflect on instructional practices

The SLO process promotes administrative initiatives and aligns to School Improvement Goals by:
• being versatile, which make them appropriate for measuring student growth for all educators, including those in non-tested subjects and grades
• being highly adaptable, making it possible to quickly respond to changes in curriculum and/or standards
• encouraging evaluators to offer feedback and suggestions regarding the individual SLOs; thus establishing administrators as educational leaders
• providing a concrete process for data analysis, goal setting, goal refinement, and evaluation—similar to the School Improvement Process which focuses on an entire school's goal attainment, the SLO process narrows the focus by teacher and/or course (Michigan Department of Education, 2015)
• expanding the scope and application of the School Improvement Process
• providing flexibility in terms of deciding which educators will utilize SLOs for the student growth measure
• reinforcing the credibility of the evaluation process by fostering ownership of student results
• providing another measure to be used for teacher evaluation, meeting the state's expectation for the employment of multiple measures
• creating opportunities for meaningful dialogue between administration and teachers for the purpose of cultivating student growth and measuring success
The SLO process promotes the retention of teachers by:
• providing teachers the opportunity to have a voice in setting their achievement standards; thus the teachers play a critical role in determining how they will be evaluated.
• allowing administrators to have data that accurately measures teachers' impact on student learning; thus allowing recognition for positive results.
• allowing teachers to have a voice in how learning will be measured and how teachers will be evaluated.
• creating opportunities for purposeful dialogue between administration and teachers, which will allow for improving student growth and measuring success.
• allowing teachers to be more engaged in and feel more connected to their own evaluation process.

Limitations and Challenges of the SLO Process

Though the use of Student Learning Objectives is a promising practice, it is not without its challenges. SLOs can be a powerful solution if implemented with care and purpose, but they are not an easy solution. There can be a misconception that SLOs are the quick and easy fix to the challenge of assessing student growth, but in reality, much time and effort is required to execute the SLO process in a credible manner. Time and effort are necessary for planning, communicating, training, and monitoring SLO implementation to make the hoped for improvement in teacher effectiveness and student learning.

FIGURE 5.5 OTHER CHALLENGES OF THE SLO PROCESS

Other challenges include:
• identifying or developing high-quality assessments for all grades and students
• creating appropriate growth targets for classrooms that include students who are starting at different achievement levels
• setting ambitious yet attainable targets including identifying the proper gain size of an objective
• addressing the school and district culture change that will result from implementing SLOs
• advancing educator practice and the continuous improvement of the SLO process

Research on the Use of Using Student Learning Objectives

Student Learning Objective research is limited. The following chart synthesizes key findings from 20 studies of SLO measures. Though there are few definitive conclusions about the efficacy and implementation of SLOs as measures of student learning and there is the need for more research, there are some lessons learned to date that can assist states and districts in making decisions about their use and implementation.

TABLE 5.3 CURRENT RESEARCH RESULTS ON IMPLEMENTATION

Lessons Learned from Early Research on the Implementation of Student Learning Objectives		
Research Questions	**Initial Findings**	**Implications for Practice and Implementation**
Do teachers perceive SLOs as affecting their practice?	Though findings are mixed, many teachers report that they believe the use of SLOs has improved their teaching, and has caused them to increase their focus on long-term student achievement and data analysis	• Collect data on teacher perceptions about the usefulness of SLOs and use this information to inform continuous improvement of your system
According to teachers, what are the benefits of the SLO process?	In multiple studies, teachers reported how the SLO process lead to an increased use of student data, promoted goal setting, teamwork, and collaboration with colleagues and a focus on quality assessments	• Highlight the value of data-driven instruction within the SLO process. If you already have effective data practices in place, leverage them • Align the SLO process with existing structures that focus on data-driven instruction
What challenges do teachers encounter in analyzing data and setting goals as part of the SLO process?	Teachers frequently cite accessing and analyzing data as the most challenging aspects of writing SLOs, and cite the need for more support	• Make data readily available to teachers and administrators in a timely manner. Make effective use of a data management system that can support data collection • Provide assessment literacy training to develop teacher and administer skill and confidence in data use
What challenges do educators face in working with assessments?	Finding, creating, or updating assessments is a time-consuming and challenging process for teachers. In some cases, teachers lack confidence in their ability to find or develop high-quality assessments and lack the expertise	• Do a district-level inventory to determine gaps in subjects and grades where high quality assessments are not readily available • Select high-quality, standards-aligned assessments upon which to base SLOs. Build capacity with district-level experts to support decisions regarding sound assessments • Provide assessment literacy training and support for teachers and administrators
What challenges in communication and support do educators report in working with SLOs?	Teachers cite ineffective communication regarding the SLO process as a major challenge to successful implementation	• Develop a communication plan and implement feedback loops • Create clear talking points and documents that identify the key messages of SLO implementation. Share how SLOs integrate into the larger system, provide context, and be sure that all stakeholders are receiving the same information about expectations and content
Does SLO quality improve over time?	Findings suggest that SLO quality, as measured by rubrics or lists of quality criteria, generally improve over time. Teachers become more comfortable with SLOs after repeated exposure and use	• Pilot using SLOs without any high stakes decisions involved initially. Allow for experience in a low-stakes environment the first year as you move to effective implementation
Is there a relationship between SLO attainment and achievement?	Limited research finds inconsistent correlations between SLO attainment and student performance on standardized tests. Though one theory of action is that teachers who produce higher quality SLOs will have engaged in thoughtful analysis and reflection and will be able to draw on this to help their students reach greater levels of achievement	• Create an intentional framework to examine overall impact of SLO implementation on multiple outcomes. Collect and compare data to determine the impact on school climate, educator engagement and retention, and the closing of achievement gaps to determine for yourself if the SLO process supports your overall district goals

Adapted from Lachlan, Lisa. *The Art and Science of Student Learning Objectives: A Research Synthesis.* American Institute of Research, 2015.

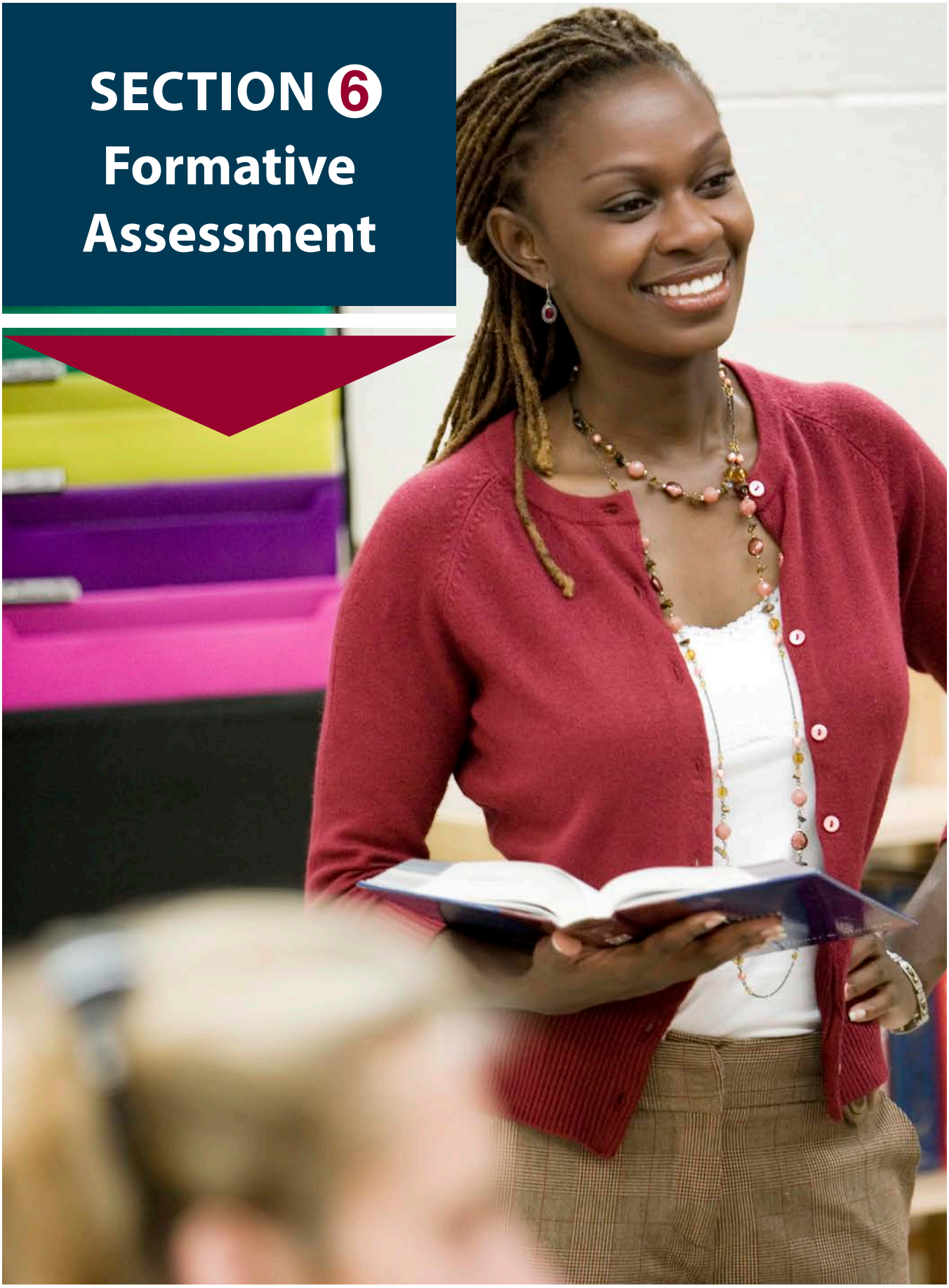
Important Considerations for Implementation

It is important to note that SLOs will only be useful if they inform educator practice. They "will not change the quality of instruction if they remain inactive documents disconnected from action. Therefore, conversation and thought around how the SLO is enacted are essential" (*The Basics*, Lachlan-Hache p.5.).

When planning for implementation, district leaders should consider the following:

- **Create a plan and start small** with a few subject areas and grades.
- **Provide professional development and use models and guidance documents** to support understanding and consistency.
- **Have early adopters document and share their experiences** and lead the conversation and momentum in the district.
- **Plan for multiple years of implementation.** This practice requires a shift in culture.
- **Provide detailed training** to ensure rigor, fairness, and comparability. These investments take time.
- **Seek results from additional studies and different approaches** which could inform implementation practices.
- **Conduct your own internal studies and data collection on the effects of pedagogy, teacher morale, and commitment.**

SECTION 6
Formative Assessment



Recent History of Formative Assessment

The Race to the Top (RTTT) Assessment Program has funded two state consortia to develop new assessment systems that measure student skills against a common set of college and career-ready standards in mathematics and English Language Arts (U.S. Department of Education, 2010). The initial RTTT invitation to submit proposals prompted extensive discussion about a vision for next-generation assessment systems intended to play a critical role in supporting students to be college and career ready. To contribute to the vision, Council of Chief State School Officers (CCSSO) published a white paper on comprehensive assessment systems to support high-quality learning. The paper called for assessment systems that supported multiple purposes at different levels of the educational enterprise and that included multiple forms of assessment, incorporating "formative as well as summative measures" (Darling-Hammond, 2010, p. 1).

The thesis of this paper is that, despite the pioneering efforts of CCSSO and other organizations in the U.S., we already risk losing the promise that formative assessment holds for teaching and learning. **The core problem lies in the false, but nonetheless widespread, assumption that formative assessment is a particular kind of measurement instrument, rather than a process that is fundamental and indigenous to the practice of teaching and learning.** Margaret Heritage (2010) says, "This distinction is critical, not only for understanding how formative assessment functions, but also for realizing its promise for our students and our society."

Effective Formative Assessment

A major landmark in the emergence of formative assessment as an explicit domain of practice was a synthesis of research findings conducted by Paul Black and Dylan Wiliam in 1998. This synthesis built on prior reviews (Crooks, 1988; Natriello, 1987) and encompassed "diverse bodies of research including studies addressing: teachers' assessment practices, students' self-perception and achievement motivation, classroom discourse practices, quality of assessment tasks and teacher questioning, and the quality of feedback" (Shepard, 2009, p. 32). From their review, Black and Wiliam (1998b) proposed that **effective formative assessment involves teachers making adjustments to teaching and learning in response to assessment evidence; students receiving feedback about their learning with advice on what they can do to improve, and students' participation in the process through self-assessment.**

They concluded that the student learning gains triggered by formative assessment were amongst the largest ever reported for educational interventions with the largest gains being realized by low achievers (1998b). This was, and remains, a powerful argument for formative assessment.

Formative Assessment and Student Achievement

Formative Assessment is a process used by educators and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes. (CCSSO, 2008) By using evidence from formative assessment processes, **educators can quickly make decisions to adjust instruction that will meet student needs while the learning is still in progress.** Students benefit from this because they can use the results to make decisions regarding the adjustment and improvement of their own learning. During the formative assessment process, feedback from the educator should be

descriptive, telling the student what was done well and what steps to take in order to improve. **Research shows descriptive feedback to be the most significant instructional strategy to move students forward in their learning.**

When incorporated into classroom practice, the formative assessment process provides information needed to adjust teaching and learning while they are still happening. The process serves as practice for the student and a check for understanding during the learning process. The formative assessment process guides educators in making decisions about future instruction. Here are a few examples that may be used in the classroom during the formative assessment process to collect evidence of student learning.

Formative Assessment and Teacher Evaluation

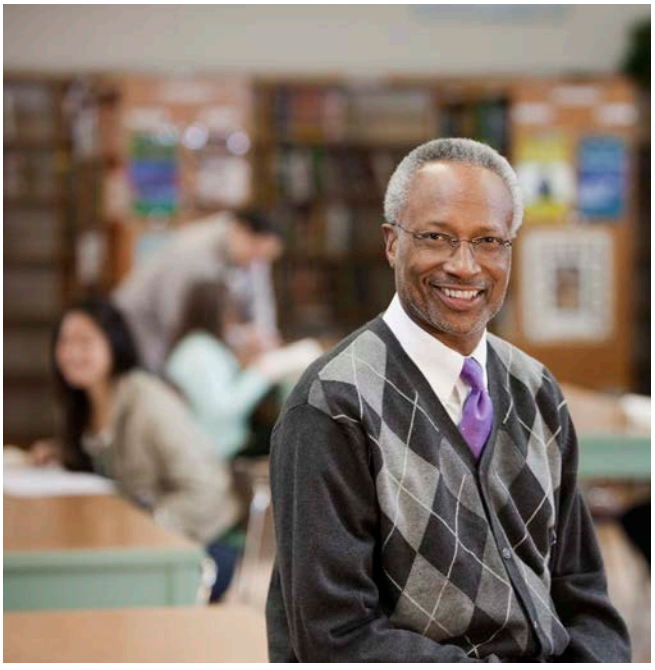
The Marzano Center believes that the best foundation for educator evaluation is Formative Assessment. **Formative Assessment is simply the measurement of student progress over time using multiple measures.** Instead of 'thick slice' assessment (data taken from a single point in time, and/or data with very large blocks of time between measurements), **formative assessments are conducted at meaningful 'thin slice' points measured throughout the academic year.** Formative assessment is, therefore, much closer to the classroom and reflects changes in instructional practices. **By making student progress part of your evaluations** (as opposed to being strictly based on a single point-in-time test score), **your own progress in improvement is factored in.**

The Marzano Center approach to evaluation:

- Identifies the direct cause-and-effect relationship between teaching practices and student achievement
- Helps educators and leaders make informed decisions to yield the greatest benefits for their students
- Is based on 40 years of collected research and five years of real-classroom experimental/control studies
- Is tested for inter-rater reliability and aligned with intensive training for accuracy and fairness
- Makes steady, measurable increases in student achievement on an achievable goal

According to the formative assessment expert, Rick Stiggins, Defensible Teacher Evaluation, weaving measurement of student learning growth into educator evaluation, requires certain criteria:

- The specific academic achievement standards for which each educator is to be held accountable must be identified and agreed to by both educator and supervisor
- Those standards must align with and sample the range of the educator's normal instructional responsibilities
- Each achievement standard must be accompanied by a detailed assessment plan and high-quality assessments
- Those assessments must be conducted in a pre-test/post-test manner to demonstrate changes in student achievement
- Educators should be given to the opportunity to describe factors, positive or negative, which influenced results



Committee Findings

After a thorough review by the Student Growth Committee, a decision was made in terms of the use of formative assessment related to measuring student growth. Essentially the committee decided, after careful analysis of the literature, that **formative assessment is not a tool or methodology designed to measure student growth in terms assigning an effectiveness label to a teacher or administrator, instead formative assessment is a multilevel process that is embedded into effective classroom instruction and educator evaluation practices.** This thinking is reflected by the language used within the four Michigan approved educator evaluation tools (See Appendix E: Danielson, 5-Dimension, Marzano, Thoughtful Classroom). Table 6.1 below represents the frequency with which indicators demonstrate alignment between the approved Michigan teach tools and formative assessment practices.

TABLE 6.1 FORMATIVE ASSESSMENT ALIGNMENT WITH MICHIGAN APPROVED TEACHER EVALUATION MODELS

Observation Tool	Number of Indicators/ Components/Elements	Number Aligned with Formative Assessment Practices
Thoughtful Classrooms	75	33
5 Dimensions	30	20
Danielson's Framework	22	14
Marzano's Model	60	24

References

SECTION ONE: Growth Models

Hull, Jim. Center for Public Education. Taken from **http://www.centerforpubliceducation.org/Main-Menu/Policies/Measuring-student-growth-At-a-glance/Measuring-student-growth-A-guide-to-informed-decision-making.html#sthash.ObX8ECGd.dpuf**

SECTION TWO: Developing and Selecting Assessments of Student Growth for Use in Teacher Evaluation Systems

Achieve. Student Assessment Inventory for School Districts. (2016) Retrieved March 23, 2016: **http://www.achieve.org/files/AchieveStudentAssessmentInventory.pdf**

Herman, J. L., Heritage, M., & Goldschmidt, P. (2011). *Developing and selecting assessments of student growth for use in teacher evaluation systems* (extended version). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Michigan Assessment Consortium. (2016). *Common Assessment Modules.* Retrieved on May 25, 2016: **http://www.michiganassessmentconsortium.org**

Ohio Department of Education. (2015). *Student Learning Objectives: A Guide to Using SLOs as a Locally-Determined Measure of Student Growth*. Retrieved on March 23, 2016: **https://education.ohio.gov/getattachment/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/Student-Growth-Measures/Student-Learning-Objective-Examples/SLO_Guidebook.pdf.aspx**

Perie, M., Marion, S., Gong, B., & Wurtzel, J. (2007). *The role of interim assessments in a comprehensive assessment system*. Achieve/Aspen/Center of Assessment Policy Brief. Retrieved from: **http://www.achieve.org/files/TheRoleofInterimAssessments.pdf**

SECTION THREE: Measuring Student Growth: A Step-by-Step Process to Analyzing Data

Anastasi, A. & Urbina, S. (1997). *Psychological Testing*, 7th ed. Prentice-Hall. Upper Saddle River, NJ.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Lawrence Erlbaum, NJ.

Confidence Interval Calculator: www.mccallum-layton.co.uk.
Taken from: **https://www.mccallum-layton.co.uk/tools/statistic-calculators/confidence-interval-for-mean-calculator/#confidence-interval-for-mean-calculator**

Gay, L.R. (1992). *Educational Research: Competencies for Analysis and Application,* 4th ed. Macmillan, N.J.

George, D. & Mallery, P. (2003). *SPSS for Windows step-by-step: A simple guide and reference* 11.0 update. 4th ed. Boston, MA: Ally and Bacon.

Gravetter F.J., Wallnau, L.B. (2004). *Statistics for the Behavioral Sciences*, 6th ed. Thomson Wadsworth, Australia.

Issac, S. & Michael, W.B. (1997). *Handbook in Research and Evaluation*, 3rd ed. EdITS/Educational and Industrial Testing Services, San Diego, CA.

Math Portal.org. Calculators: Statistics Calculators: Descriptive Statistics. Taken from: **http://www.mathportal.org/calculators/statistics-calculator/descriptive-statistics-calculator.php**

Nunnally, J.C. (1978). *Psychometric Theory.* McGraw-Hill, N.Y.

Nunnally, J.C. & Bernstein, I.H. (1994). *Psychometric Theory.* McGraw-Hill, N.Y.

Percent Change Calculator. **http://www.percent-change.com/**

QuickCalcs. t-test calculator. GraphPad Software.
Taken from **http://www.graphpad.com/quickcalcs/ttest1.cfm** .

Social Science Statistics. Effect Size Calculator.
Taken from **http://www.socscistatistics.com/effectsize/Default3.aspx** .

Social Science Statistics. Wilcoxon Signed-Rank Test
Taken from **http://www.socscistatistics.com/tests/signedranks/Default.aspx** .

SECTION FOUR: Standard Setting for Student Growth

Cizek, Gregory J. and Bunch, Michael B. *Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests*. Sage Publications, 2007.

Fielding, Lynn, Kerr, Nancy, and Rosier, Paul. *Annual Growth for All Students Catch-Up Growth for Those Who Are Behind*. The New Foundation Press, 2007.

Garmstom, Robert J. and Wellman, Bruce M. *The Adaptive School: A Sourcebook for Developing Collaborative Groups*. Rowland & Littlefield Pub., 2013.

How to Design and Agenda for an Effective Meeting. Harvard Business Review. Taken 4/4/2016 from: **https://hbr.org/2015/03/how-to-design-an-agenda-for-an-effective-meeting** .

Protocols for Question Types in Group Facilitation. School Reform Initiative. Taken 4/4/2016 from: **http://www.nsrfharmony.org/free-resources/protocols/a-z** .

SECTION FIVE: Student Learning Objectives: A Measure of Educator Effectiveness

Lachlan, Lisa. *The Art and Science of Student Learning Objectives: A Research Synthesis*. American Institute of Research, 2015.

Lachlan, Lisa, Cushing, Ellen, and Bivona, Lauren. *Student Learning Objectives: Benefits, Challenges, and Solutions*. American Institute of Research, 2012.

Lachlan, Lisa, Cushing, Ellen, and Bivona, Lauren. *Student Learning Objectives as Measures of Educator Effectiveness: The Basics*. American Institute of Research, 2012.

Lachlan, Lisa, Cushing, Ellen, and Bivona, Lauren. *Student Learning Objectives as Measures of Educator Effectiveness: Early Lessons.* Teacher Incentive Michigan Department of Education. (2015).

Measuring student growth: An introduction to Student Learning Objectives. Retrieved on March 23, 2016: **https://www.michigan.gov/documents/mde/SLO_FAQ_11.25.15_507156_7.pdf**

Ohio Department of Education. (2015). Student Learning Objectives: A Guide to Using SLOs as a Locally-Determined Measure of Student Growth. Retrieved on March 23, 2016: **https://education.ohio.gov/getattachment/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/Student-Growth-Measures/Student-Learning-Objective-Examples/SLO_Guidebook.pdf.aspx**

Slotnik, William J. *Targeting Growth: Using Student Learning Objectives as a Measure of Educator Effectiveness.* Community Training and Assistance Center. Retrieved on May 24, 2016: **http://www2.ed.gov/programs/racetothetop/communities/stargetinggrowth.pdf**

Wormeli, R. *Differentiation: From Planning to Practice Grades 6-12.* Portland, ME: Stenhouse, 2007.

SECTION SIX: Formative Assessment

Black, P. J., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5, 7-73.

Black, P. J., & Wiliam, D. (1998b). Inside the Black Box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80, 139-48.

CCSSO (2008). *Educational Leadership Policy Standards: ISLLC 2008* as Adopted by the National Policy Board for Educational Administration

Crooks, T. J. (1988). The Impact of classroom evaluation on students. *Review of Educational Research*, 58, 438-481.

Darling-Hammond, L. (2010). *Performance counts: Assessment systems that support high-quality learning.* Washington, DC: Council of Chief State School Officers.

Heritage, M (2010). National Center for Research on Evaluation, Standards, and Student Testing (CRESST) Graduate School of Education and Information Studies University of California, Los Angeles. Paper prepared for the Council of Chief State School Officers September 2010

Marzano retrieved March 1, 2016. **http://www.marzanocenter.com/blog/article/why-formative-assessment-is-a-big-win-in-teacher-evaluation/#sthash.W1tcA3AY.dpuf**

Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist,* 22, 155-175.

RTT(2010). "Race to the Top Program Executive Summary" U.S. Department of Education. Retrieved January 26, 2010.

Shepard, L. A. (2009). Commentary: Evaluating the validity of formative and interim assessment. *Educational Measurement: Issues and Practice*, 28(3), 32-37.

Stiggins, R. (2014). *Defensible Teacher Evaluation, Student Growth Through Classroom Assessment*. Corwin Thousand Oaks, CA.

Wayne Count RESA 5 Pillars of the Formative Assessment Process Poster

West Virginia Retrieved February 29, 2016.
http://wvde.state.wv.us/teach21/FormativeAssessment.html

GLOSSARY

American Institutes for Research. Progress Monitoring Taken on April 27, 2016 from
http://www.rti4success.org/essential-components-rti/progress-monitoring

American Institutes for Research. Student Learning Objectives: Benefits, Challenges, and Solutions. November 2012. Taken on April 27, 2016 from
https://westcompcenter.org/wp-content/uploads/2013/04/SLO-Benefits-and-Challenges-American-Institutes-for-Research.pdf .

Assessment Literacy: Making Sense of K-12 Assessment for Learning. Taken on April 27, 2016 from
http://assessmentliteracy.org/about/ .

Castellano, Katherine E. and Ho, Andrew D. *A Practitioner's Guide to Growth Models*. Council of Chief State School Officers, 2013.

Wood, John. The Importance of Comparability in Today's Assessment Environment. Taken April 27, 2016 from
http://assessmentliteracy.org/blog/importance-data-comparability-todays-assessment-environment/ .

Glossary

Terms for Statistics and Measurements	Definition	Answers the Question	Pros in Growth Measurement	Cons in Growth Measurement
Assessment Literacy	Refers to an educator's comprehensive understanding of assessment and its role in learning.	How well do I use assessment to improve the learning of my students?	• Is essential for teachers and administrators to understand the assessment data they have available and are using to define and analyze student growth.	• Requires professional development and opportunities to apply understandings of assessment in a meaningful context. • Requires time. • Requires motivation of educators to participate in, learn and apply assessment literacy to their work.
Confidence Interval	A range represented by a lower limit number and upper limit number.	How confident are you that the true mean falls between the two numbers? We say we are 95% confident.	• Provides a good visual for a measure of central tendency (true mean).	• It is not symmetric around the mean resulting in a possible low normal and a high normal.
Criterion Referenced Data	Tests and assessments are designed to measure student performance against a fixed set of predetermined criteria or learning standards.	What are students expected to know and be able to do at this point in their education?	• Criterion referenced assessments are preferable in comparing student performance to previous learning or rating performance aligned to a learning expectation.	• Criterion assessments can be time-consuming and complex, expensive to implement, and do not readily allow comparisons among students.
Interim Assessments	Assessments that are administered between annual assessments. For example, an interim assessment might occur in the fall, winter, and spring to be compared to annual spring assessments.	Is student learning on track toward annual performance expectations? Is sufficient curriculum being covered for students to meet annual assessment expectations?	• Interim assessments provide the ability to gather and compare data within a single year and over the course of multiple years. • The data provide longitudinal information for making comparisons over time. • Administrators often use the data to track student growth.	• There is concern with the amount of time that students spend taking tests with interim assessments. • Time for teachers to review the data and to understand how to use the data to adjust curriculum and instruction can be a problem. • The method assumes that growth is linear when that may not be the best trajectory for the student's developmental level or the skills being assessed.

Terms for Statistics and Measurements	Definition	Answers the Question	Pros in Growth Measurement	Cons in Growth Measurement
Mean	Represents the arithmetic average of scores. It is a measure of central tendency.	What is the average gain for the data on hand?	• Easy to calculate. • Can be used when identifying growth based on average number of students or averages of norm referenced data.	• Masks trends in the distribution of student gains from high to low. • Does not describe range of data. It is affected by extreme scores (outliers).
Median	Represents the mid-point in a distribution of scores. One-half of the scores fall below it and above it. It is a measure of central tendency.	What is the mid-point within the data set? Or what is the 50th percentile score?	• Requires the ranking of the data (or scores) from lowest to highest. It is a stable measure because it is not impacted by extreme scores (outliers). • It permits one to determine at which point a child is represented in terms of percentiles. • Can be more "fair" in representing data trends within the distribution of scores than a solitary mean score. • Most useful with student growth percentile data.	• Represents aggregate data. One should conduct quality assurance checks to ensure that the data entry was correct prior to calculating. • Should use a software with large data sets (Excel).
Mode	The mode is the value that appears most often in the data set.	What is the most common gain observed within the data set?	• Identifies the gain that is most commonly demonstrated across students.	• Time to organize the data for analysis and interpretation. • Does not represent the range of gains in student growth. It may take on a bi-modal shape or two modes. • Requires a context to be meaningful, e.g., a specific teacher's data set with additional explanation of factors.

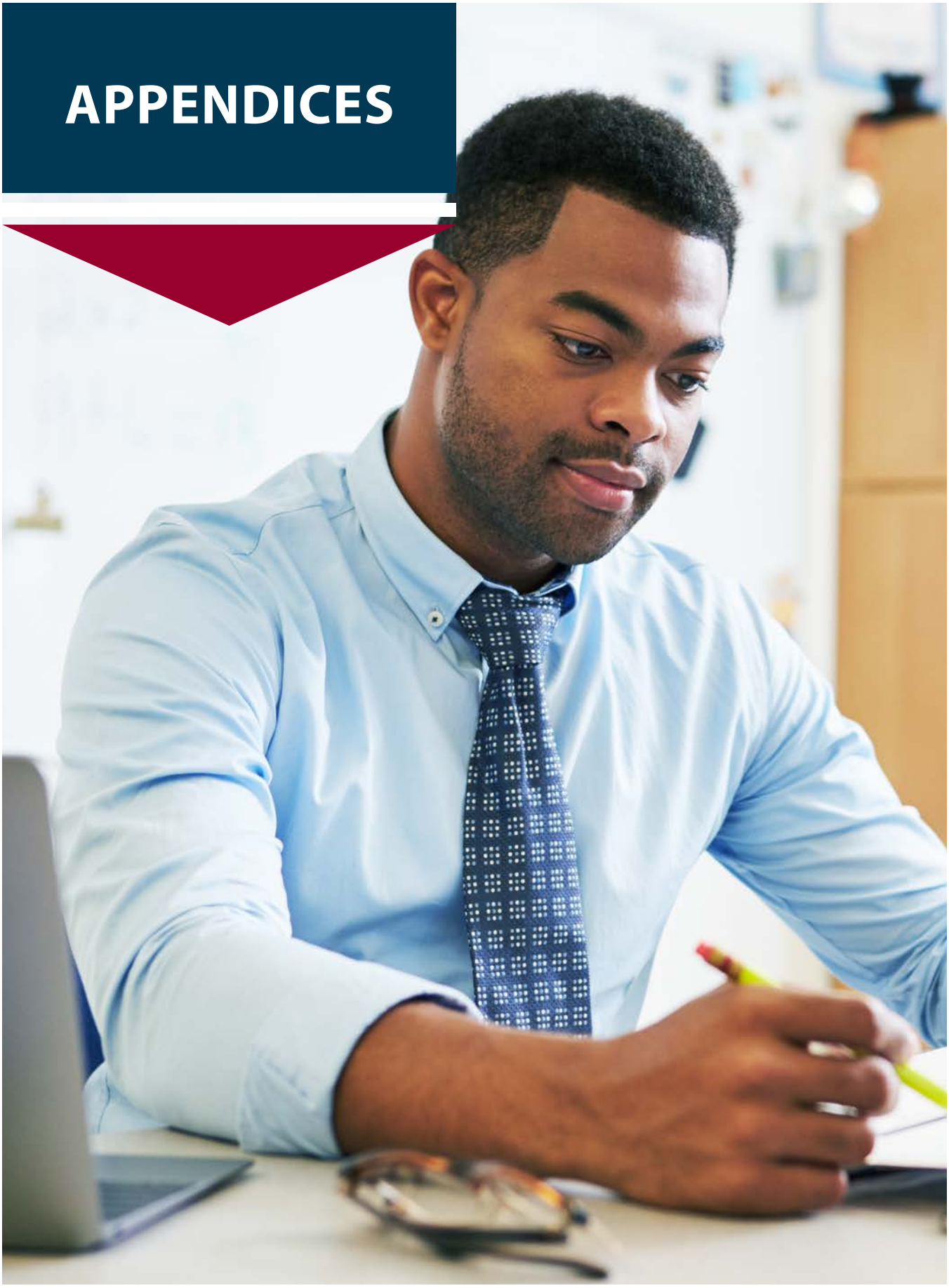
Terms for Statistics and Measurements	Definition	Answers the Question	Pros in Growth Measurement	Cons in Growth Measurement
Norm Referenced Data	Norm-referenced data compares the individual's performance to that of others, usually of the same age or grade level.	How does this individual's performance compare to others?	• Data can be compared across individuals. • Data can be represented in equal interval units, such as standard scores or percentiles. • There is control for central tendency.	• Norm-referenced data may be too far removed from classroom instruction to be appropriate in teacher evaluation. • The representativeness of the sample may not match the local norms in performance or sampling. It makes no mention of content mastery, rather, it asks how a student did compared to her norm.
Percentile	A score that represents the ranking of scores from highest to lowest. For example, a score at the 75th percentile means that the score is greater than or equal to 75% of the persons taking the test.	How does this individual's score rank in comparison to others?	• The percentile provides a ranking or comparison that describes the relative standing of the individual in terms of the percent who performed equal and less well on the task. • Can be simple to calculate. It is misleading if examining scores from a highly gifted student population.	• Is often confused with a percentage score. • The percentile does not communicate the spread of scores from one another but the placement of the individual's score from high to low. • Calculation tools may vary in regard to central tendency in score dispersion.
Percentage	A ratio or number that expresses a fraction of 100.	What is the ratio of success on this task?	• The percent is simple to calculate. • The percent can be used to represent the ratio of students meeting certain criteria or levels of performance. Is often used by teachers when grading students. • Can be helpful to monitor growth and to summarize performance.	• Can be misused as a target for educator evaluation purposes, especially when used without a context of past performance, years of trend data, and analysis of what is reasonable within growth measurement timeframes.
Performance Level Descriptor	The performance level descriptor is the written criterion for the categories of a rubric.	What is the criterion that distinguishes each category?	• Is customized to the context of data, content, and categories. • Provides a standard against which raters classify data into categories.	• Requires clearly written descriptors.

Terms for Statistics and Measurements	Definition	Answers the Question	Pros in Growth Measurement	Cons in Growth Measurement
Predicted Score	A method of growth measurement in which past scores are used as a basis for projecting future scores.	Given the student's past scores or patterns of scores in the past, what is the predicted score for the future?	• Requires the setting of a future standard of performance and a time frame to meet the standard.	• Predicted scores can be confused with "trajectory". • Emphasis on predicted scores can diminish incentive to work with low achieving students.
Progress Monitoring	A method of assessing a student's academic performance, to quantify a student's rate of improvement or responsiveness to instruction, and to evaluate the effectiveness of instruction. Can be implemented with individual students or a class.	Is the student making progress with instruction and/or intervention?	• Repeated brief and targeted assessments are used that are aligned directly to the instruction of skill(s). • Can be easily represented in graphs. • Can be used with targets or goals.	• Identifying a method of progress monitoring that aligns with instruction. • The focus of the progress monitoring may be too narrow for educator evaluation purposes. • Requires training and monitoring of how the data are used to adjust instruction. • There is no gold standard in the number of observations needed to witness growth (e.g., 3 or 10 observations?)
Reliability	Reliability refers to the consistency of scores over time or the ability of a measure to be repeated with the same or similar results. It is inappropriate to say that a test is reliable because reliability is a function of data or scores on hand.	Are the data from this assessment consistent? If I did this again, would I get the same results?	• Relatively easy to calculate. • Strong reliability indicates that the method is stable.	• Requires some statistical calculation skill or access to calculation tools. • Tests or assessments that are highly reliable may not be sensitive to changes that are age/grade appropriate and meaningful to the individual. • Tests or assessments that have low reliability cannot be trusted to yield consistent information. It is a paradox when attempting to measure change. • High stakes testing requires reliability coefficients ≥ .90.

Terms for Statistics and Measurements	Definition	Answers the Question	Pros in Growth Measurement	Cons in Growth Measurement
Standard Deviation	A statistical method of analyzing the amount of variance around a score.	How much might the score vary due to factors other than ability?	• The standard deviation is an important statistic for describing the amount of error surrounding a score. • It is useful in understanding change in test scores between administrations. For example, if two scores are within the same standard deviation that would indicate that there was little change and the difference in scores may be due to normal fluctuations in the test scores/ data.	• The standard deviation is often not used, not available, or not referenced when analyzing test score data. • Requires some understanding of test scores and statistics to analyze and reference in the context of student growth measurement.
Standard Setting	Process for defining gains that requires judgment about adequate gain or adequate average gain. Requires understanding of the measurement scale or can be norm-referenced.	What are the cut points for differentiating teacher effectiveness categories using student growth data?	• A cut score is established based on performance level criteria. • Involves stakeholders. • Can be revised based on new information. • Provides a context for understanding data and making meaning of growth data categories.	• Can be a time-consuming process. • Requires training and understanding of data, measurement, and performance criteria. • Requires attention to business rules and clarity of terms.
Student Learning Objective (SLO)	A specific learning goal and a specific measure of student learning used to track progress toward the goal.	What is the expectation of learning and method of tracking progress toward that goal?	• The SLO in the context of educator evaluation reinforces best teaching practice, encourages collaboration, relies on teacher skill, and is considered to be helpful in connecting teacher practice to student skill.	• It can be difficult to identify and develop high quality assessments across all grades and subjects. • There are challenges to creating appropriate growth targets for classrooms in which students are starting at different achievement levels. • There are challenges to setting attainable yet rigorous targets with the proper "gain" size.

Terms for Statistics and Measurements	Definition	Answers the Question	Pros in Growth Measurement	Cons in Growth Measurement
Trajectory	A trajectory extends gains or average gains in a predictable, usually linear fashion into the future. Trajectories may be used when using growth-to-benchmark models or gain-score models.	If this student continues on this trajectory, where is she likely to be in the future?	• The trajectory is set by defining a future standard and a time horizon to meet the standard.	• The prediction is descriptive and aspirational. • Requires defensible vertical scaling over many years. • Can be inflated by dropping initial scores.
Validity	Validity is the extent to which a concept, conclusion or measurement is well-founded and corresponds accurately to the real world.	Does the assessment measure the skill, construct, or content it purports to measure?	• Validity is important to ensure the test is measuring the intended content.	• Sometimes persons mistake face validity as sufficient to determine the quality of the content.
Vertical Scaling	Vertical Scaling is the method based on Item Response Theory for assuring the items of a test are aligned to show growth.	Does the vertical scaling represent the developmental appropriateness of performance standards progression over grade levels?	• Vertical scaling provides consistent scores across grade levels and is advantageous for measuring growth.	• The procedure requires sophisticated statistical methods.

APPENDICES



APPENDIX A *Skills for Standard Setting Facilitators*

The collaborative inquiry and problem solving protocol is a preferred method for facilitating groups to thoughtfully arrive at consensus when working with data.

PURPOSES

- ***To provide the participant an opportunity to discuss and understand data presented***
- ***To ensure the standard setting methodology is adhered to and implemented with fidelity***

PROTOCOLS FOR QUESTION TYPES IN GROUP FACILITATION

Question Type	Description	Facilitator Tips
Clarifying Questions	Participants ask clarifying questions of the facilitator. If there is an expert on the technical aspects, the facilitator may refer to them for answers. Questions are asked and answered.	The facilitator should ensure the process is not one of challenging the data or questioning where the data come from. It is imperative that the participants are reflective and stay on task to follow the procedure of the standard setting process.
Probing Questions	Probing questions serve to give participants the opportunity to focus on details or comments made in their group work.	The purpose of the probing questions is to push the group to think more deeply or extensively about the data and their questions. The facilitator may focus the attention of the group with the use of probing questions.
So What? **What Else?** **Now What?**	The facilitator gives the group the opportunity to answer ***"So What?", "What Else?", and "Now What?"*** questions based upon their own perceptions and thinking. **So What:** • What strikes me as significant? • What concerned me? **What Else:** • What is the impact of my ratings? • Do I have enough information? **Now What:** • Define actions to be taken within the standard setting process.	The purpose of these questions are to expand the group's understanding of the data and process, not to give advice. The facilitator should intervene if suggestions about action steps are generated.

FEEDBACK ON PROCESS

The facilitator asks the group to reflect on the process and outcome of the standard setting activities.

- ***Was this reasonable?***
- ***Can this be replicated?***

Adapted and taken 4/4/2016 from http://www.nsrfharmony.org/free-resources/protocols/a-z

APPENDIX B *Meeting Agendas and Norms*

TIPS FOR MEETING AGENDAS

When setting meeting agendas for the team, here are some tips from the *Harvard Business Review* (March 19, 2015)

- Provide a schedule in advance.
- Specify how to prepare for the meeting.
- Establish confidentiality of meeting discussions.
- Keep topics and activities to one hour intervals. If working for more than one hour, plan for breaks and allow for appropriate time.
- List agenda topics as questions the team needs to answer.
- Clearly communicate the purpose to participants.
- End the meeting with reflection, summation, and positive feedback.

Adapted and taken 4/4/2016 from https://hbr.org/2015/03/how-to-design-an-agenda-for-an-effective-meeting .

ESTABLISHING MEETING NORMS

The **7 Norms of Collaboration** are helpful in committee meetings. These skills prepare participants to be effective communicators. The norms serve to keep the communications focused on the topic and task of the committee.

1. **Pausing:** Listen attentively; allow thinking time; reword in our own mind what others are saying; wait for others to finish speaking.
2. **Paraphrasing:** Paraphrase to acknowledge and clarify content, summarize and organize, shift conversation.
3. **Probing:** Seek agreement on what words mean; asks questions to clarify and solicit points-of-view.
4. **Presuming positive intention:** Act as if others mean well; restrain from emotional responses; use positive presuppositions when responding to others.
5. **Put ideas on the table and pull them off:** State intention of communication; reveal relevant information; provide fact, inferences, ideas, opinions, point of view; explain reasons behind comments and actions; communicate modifications.
6. **Pay attention to self and others:** Maintain awareness of own and others thoughts and feelings; maintains awareness of groups task, mood, and contributions.
7. **Promoting inquiry:** inquire about perceptions, assumptions, interpretations.

In addition, the committee will need to establish norms for committee behavior that may include timeliness, attendance, confidentiality, and preparedness for activities.

For more information, refer to: Garmstom, Robert J. and Wellman, Bruce M. *The Adaptive School: A Sourcebook for Developing Collaborative Groups.* Rowland & Littlefield Pub., 2013.

APPENDIX C

Collection of Evidence: Example of How to Summarize Rating to Establish Cut Points

When working with collections of evidence, such as portfolios or IEP evidence, **the first step is to create a rubric for scoring the evidence.** The evidence are scored independently; then in the standard setting process, each committee member reviews the collection of evidence with the associated score and indicates the category of effectiveness rating, based on the performance level descriptor for each category. The data are then summarized, as demonstrated in the table below. In this example, the collection of evidence may yield scores that range from 15-20. The frequency of rating at each score is summarized. The scores are totaled within each category and a mean is established. The cut score is established as the midpoint among the ratings.

EXAMPLE OF COLLECTION OF EVIDENCE SUMMARY CUT SCORES

	Score Points Associated with Frequency of Participant Ratings								
Rater Scoring	**15**	**16**	**17**	**18**	**19**	**20**	**Total**	**Mean**	**Midpoint**
Highly Effective				1	3	2	6	19	**19**
Effective			2	2		1	5	18	**18**
Minimally Effective	2	1	1				4	16	**16**
Ineffective	1	2					3	15.5	**15**

***Note:** *It is often necessary to determine what to do when dealing with averages that are not even numbers. In this example, the decision was to round down.*

APPENDIX D

Student Learning Objective (SLO) Template

Teacher/Teacher Team: _____

Content Area: _____ **Grade Level:** _____ **Academic Year:** _____

Type of SLO: ☐ **Class-Level** ☐ **Course-Level** ☐ **Targeted** ☐ **Tiered**

Student Population:

Describe the characteristics of the student population including how many students have special needs relevant to the SLO (e.g., *I have 4 students with reading disabilities, 2 English Language Learners*…).

Learning Standards:

List the key standards that are connected to the learning content.

Interval of Time:

Describe the duration of instruction based on the course structure. Include start and end dates.

Assessment:

Name the instrument that will be used to measure the outcome of the SLO.

SLO Growth Targets:

Identify the quantitative targets that will demonstrate achievement of the SLO.

Rationale:

Explain your rationale for setting the above targets for student growth; How do the targets connect with the school improvement plan?

SLO Approval Committee

1.	4.
2.	5.
3.	6.

APPENDIX E *Alignment Document*

The Formative Assessment Process and Four Teacher Observation Tools

- *The Formative Assessment Practices Teacher Survey*
- *Charlotte Danielson's Framework for Teaching (2014)*
- *5 Dimensions of Teaching and Learning*
- *Marzano's Art and Science of Teaching Framework*
- *Silver Strong & Associates: Thoughtful Classroom*

Alignment: The Formative Assessment Process and Danielson 2014		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**Danielson Rubric Component**
Planning	Understand and reflect on the planning requires throughout the Formative Assessment Process	**Domain 1: Planning and Preparation** • **1a:** Demonstrate Knowledge of Content and Pedagogy • **1c:** Setting Instructional Objectives • **1e:** Designing Coherent Instruction • **1f:** Designing Student Assessments **Domain 3: Instruction** • **3e:** Demonstrating Flexibility and Responsiveness (Lesson Adjustment)
Learning Targets	Have an awareness of the importance of clear learning targets for the purpose of teacher planning and student clarity	**Domain 1: Planning and Preparation** • **1c:** Setting Instructional Objectives **Domain 3: Instruction** • **3a:** Communicating with Students (Expectations for Learning)
	Develop and implement ways to embed learning targets within instruction	**Domain 1: Planning and Preparation** • **1c:** Setting Instructional Objectives **Domain 3: Instruction** • **3d:** Using Assessment in Instruction (Monitoring Student learning)
	Develop clear criteria for success	**Domain 1: Planning and Preparation** 1c: Setting Instructional Objectives **Domain 2: The Classroom Environment** • **2b:** Establishing a Culture for Learning (Expectations for Learning & Achievement) **Domain 3: Instruction** • **3a:** Communicating with Students (Expectations for Learning) • **3d:** Using Assessment in Instruction (Monitoring Student learning)
Questioning/ Discussion	Effective Questioning Strategies (HOTS, Wait Time, No Opt Out, Cold Calling) are used throughout the lesson	**Domain 3: Instruction** • **3b:** Using Questioning and Discussion Techniques (Quality of Questions/Prompts & Student Participation)
	Teacher uses a variety of Question Types (Clarifying; Probing)	**Domain 3: Instruction** • **3b:** Using Questioning and Discussion Techniques (Quality of Questions/Prompts & Student Participation)
	Effective Questions lead to Quality Discussion	**Domain 3: Instruction** • **3b:** Using Questioning and Discussion Techniques (Quality of Questions/Prompts & Student Participation)
Formative Feedback	Provide Students with Appropriate Feedback on Formative Tasks	**Domain 3: Instruction** • **3d:** Using Assessment in Instruction (Monitoring Student learning)
	Provide students Opportunities to Reflect on Feedback and make Revisions	**Domain 3: Instruction** • **3d:** Using Assessment in Instruction (Monitoring Student learning)

Alignment: The Formative Assessment Process and Danielson 2014		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**Danielson Rubric Component**
Students as Self/Peer Assessors	Students Reflect on Data and Feedback to Improve Their Learning	**Domain 3: Instruction** • 3d: Using Assessment in Instruction (Monitoring of Student Learning)
	Students engage in activities with peers to Reflect on Data and Feedback to Improve Learning	**Domain 3: Instruction** • **3c:** Engaging Students in Learning (Grouping of Students) • **3d:** Using Assessment in Instruction (Student Self-Assessment and Monitoring of Progress)
Collection and Use of Data	Develop and Implement Products, Tools, Observations, and Conferences that Gather Evidence of Student Learning During the Learning Process.	**Domain 1: Planning and Preparation** • **1f:** Designing Student Assessments (Congruence with Instructional Outcomes, Criteria and Standards, Design of Formative Assessments, Use for Planning)
	Teachers use Formative Data to Adjust Instruction When Necessary	**Domain 3: Instruction** • **3c:** Engaging Students in Learning (Structure and Pacing) • **3e:** Demonstrating Flexibility and Responsiveness (Lesson Adjustment, Response to Students, Persistence) **Domain 4: Professional Responsibilities** • **4a:** Reflecting on Teaching (Accuracy) • **4b:** Maintaining Accurate Records (Student Progress in Learning)
	Students use Formative Data to Adjust Learning Tactics when necessary	**Domain 3: Instruction** • **3e:** Demonstrating Flexibility and Responsiveness (Lesson Adjustment, Response to Students, Persistence)
	Assessments are Developed and Evaluated Collaboratively to Ensure Quality and Appropriate Use	**Domain 4: Professional Responsibilities** • **4a:** Reflecting on Teaching (Accuracy) • **4d:** Participating in a Professional Community • **4e:** Growing and Developing Professionally

Alignment: The Formative Assessment Process and 5D+ Teacher Evaluation Rubric		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**5D+ Rubric Component**
Planning	Understand and reflect on the planning requires throughout the Formative Assessment Process	• **P1** The lesson is based on grade level standards and the learning target(s) align to the standard. The lesson is consistently linked to broader purpose or a transferable skill. • **CP1** Instructional materials and tasks align with the purpose of the unit and lesson. • **A3** Teacher provides a variety of strategies for formative assessment that align with the learning targets.
Learning Targets	Have an awareness of the importance of clear learning targets for the purpose of teacher planning and student clarity	• **P1** The lesson is based on grade level standards and the learning target(s) align to the standard. The lesson is consistently linked to broader purpose or a transferable skill. • **CP1** Instructional materials and tasks align with the purpose of the unit and lesson. Materials and tasks consistently align with student's level of challenge. • **A2** Assessment tasks are aligned with the learning targets and allow students to demonstrate complex understanding and/or skill related to the learning targets. • **A6** Teacher uses formative assessment data to make in-the-moment instructional adjustments, modify future lessons and give targeted feedback aligned with the learning target to individual students.
	Develop and implement ways to embed learning targets within instruction	• **P4** Teacher communicates the learning target(s) through verbal and visual strategies, checks for student understanding of what the target(s) are and references the target throughout instruction. • **A1** Students consistently assess their own learning in relation to the success criteria and can determine where they are in connection to the learning target.
	Develop clear criteria for success	• **P5** The success criteria for the learning target(s) are clear to students. The performance tasks align to the success criteria. Students refer to success criteria and use them for improvement.

Alignment: The Formative Assessment Process and 5D+ Teacher Evaluation Rubric		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**5D+ Rubric Component**
Questioning/ Discussion	Effective Questioning Strategies (HOTS, Wait Time, No Opt Out, Cold Calling) are used throughout the lesson	• **SE1** Teacher frequently asks questions to probe and deepen students' understanding or uncover misconceptions. Students ask questions of one another to probe for deeper thinking.
	Teacher uses a variety of Question Types (Clarifying; Probing)	• **SE1** Teacher frequently asks questions to probe and deepen students' understanding or uncover misconceptions. Students ask questions of one another to probe for deeper thinking. • **A1** Students consistently assess their own learning in relation to the success criteria and can determine where they are in connection to the learning target.
	Effective Questions lead to Quality Discussion	• **CEC6** Classroom discourse and interactions reflect high expectations and beliefs about all students' intellectual capabilities and create a culture of inclusivity, equity and accountability for learning. • **SE4** Engagement strategies encourage equitable and purposeful student participation and ensure that all students have access to, and are expected to participate in, learning. • **SE6** Student talk embodies substantive and intellectual thinking.
Formative Feedback	Provide Students with Appropriate Feedback on Formative Tasks	• **A6** Teacher uses formative assessment data to make in-the-moment instructional adjustments, modify future lessons and give targeted feedback aligned with the learning target to individual students.
	Provide students Opportunities to Reflect on Feedback and make Revisions	• **A6** Teacher uses formative assessment data to make in-the-moment instructional adjustments, modify future lessons and give targeted feedback aligned with the learning target to individual students.

Alignment: The Formative Assessment Process and 5D+ Teacher Evaluation Rubric		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**5D+ Rubric Component**
Students as Self/Peer Assessors	Students Reflect on Data and Feedback to Improve Their Learning	• **P5** Students refer to success criteria and use them for improvement. • **SE2** Teacher consistently provides opportunities and strategies for students to take ownership of their learning. Most locus of control is with students in ways that support students' learning. • **CP6** Teacher provides scaffolds and structures that are clearly related to and support the development of the targets concepts and/or skills. Students use scaffolds across tasks with similar demands. • **A1** Students consistently assess their own learning in relation to the success criteria and can determine where they are in connection to the learning target. • **A5** Students consistently use assessment data to assess their own learning, determine learning goals and monitor progress over time.
	Students engage in activities with peers to Reflect on Data and Feedback to Improve Learning	• **P5** Students refer to success criteria and use them for improvement. • **SE2** Teacher consistently provides opportunities and strategies for students to take ownership of their learning. Most locus of control is with students in ways that support students' learning. • **CP6** Teacher provides scaffolds and structures that are clearly related to and support the development of the targets concepts and/or skills. Students use scaffolds across tasks with similar demands. • **A1** Students consistently assess their own learning in relation to the success criteria and can determine where they are in connection to the learning target. • **A5** Students consistently use assessment data to assess their own learning, determine learning goals and monitor progress over time.

Alignment: The Formative Assessment Process and 5D+ Teacher Evaluation Rubric		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**5D+ Rubric Component**
Collection and Use of Data	Develop and Implement Products, Tools, Observations, and Conferences that Gather Evidence of Student Learning During the Learning Process.	• **A3** Teacher provides a variety of strategies for formative assessment that align with the learning target. • **A4** Teacher has an observable system and routines for (collecting) and recording formative assessment data, uses multiple sources and consistently uses the system for instructional purposes.
	Teachers use Formative Data to Adjust Instruction When Necessary	• **CP5** Teacher consistently uses strategies that differentiate for individual learning strengths and needs. • **CP6** Teacher provides scaffolds and structures that are clearly related to and support the development of the targets concepts and/or skills. • **A6** Teacher uses formative assessment data to make in-the-moment instructional adjustments, modify future lessons and give targeted feedback aligned with the learning target to individual students.
	Students use Formative Data to Adjust Learning Tactics when necessary	• **CP6** Teacher provides scaffolds and structures that are clearly related to and support the development of the targets concepts and/or skills. • **A5** Students consistently use assessment data to assess their own learning, determine learning goals and monitor progress over time.
	Assessments are Developed and Evaluated Collaboratively to Ensure Quality and Appropriate Use	• **CEC1** Teacher collaborates and engages in reflective inquiry with peers and administrators for the purpose of improving instructional practice, and student and teacher learning. • **CEC2** Teacher develops and sustains professional and collegial relationships for the purpose of student, staff, or district growth. • **CEC5** Teacher supports and looks for opportunities to take on leadership roles in developing and implementing school, district, and state initiatives.

Alignment: The Formative Assessment Process and Marzano Art and Science of Teaching Framework		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**Marzano Art & Science of Teaching Framework**
Planning	Understand and reflect on the planning requires throughout the Formative Assessment Process	• **D1-1:** Providing Clear Learning Goals and Scales (Rubrics) • **D1-6:** Identifying Critical Information • **D2-1:** Planning and Preparing for Effective Scaffolding of Information within Lessons • **D2-2:** Planning and Preparing for Lessons within a Unit that Progress Toward a Deep Understanding and Transfer of Content • **D2-3:** Planning and Preparing for Appropriate Attention to Established Content Standards
Learning Targets	Have an awareness of the importance of clear learning targets for the purpose of teacher planning and student clarity	• **D1-1:** Providing Clear Learning Goals and Scales (Rubrics) — Learning Goal is Posted — Learning goal is a clear statement of knowledge or information as opposed to an activity or assignment • **D1-6:** Identifying Critical Information — Teacher begins the lesson by explaining why upcoming content is important
	Develop and implement ways to embed learning targets within instruction	• **D1-1:** Providing Clear Learning Goals and Scales (Rubrics) — Teacher references learning goal throughout the lesson — Teacher makes reference to scale or rubric throughout the lesson • **D1-2:** The teacher facilitates tracking of student progress — Teacher helps students track individual progress on the learning goal
	Develop clear criteria for success	• **D1-1:** Providing Clear Learning Goals and Scales (Rubrics) — Teacher has a scale or rubric that relates to the learning goal posted

Alignment: The Formative Assessment Process and Marzano Art and Science of Teaching Framework		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**Marzano Art & Science of Teaching Framework**
Questioning/ Discussion	Effective Questioning Strategies (HOTS, Wait Time, No Opt Out, Cold Calling) are used throughout the lesson	• **D1-11:** Teachers help students Elaborate on New Information — The teacher asks questions • **D1-26:** Managing Response Rates — Teacher uses response-rate techniques to maintain student engagement in questions
	Teacher uses a variety of Question Types (Clarifying; Probing)	• **D1-11:** Teachers help students Elaborate on New Information — The teacher asks questions; students explain and defend their inferences
	Effective Questions lead to Quality Discussion	• **D1-11:** Teachers help students Elaborate on New Information — The teacher asks questions; students explain and defend their inferences
Formative Feedback	Provide Students with Appropriate Feedback on Formative Tasks	• **D1-23:** Providing Resources and Guidance — Teacher acts as resource provider and guide as students engage in cognitively complex tasks
	Provide students Opportunities to Reflect on Feedback and make Revisions	• **D1-20:** Helping Students Revise Knowledge — Teacher engages students in revision of previous knowledge

Alignment: The Formative Assessment Process and Marzano Art and Science of Teaching Framework		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**Marzano Art & Science of Teaching Framework**
Students as Self/Peer Assessors	Students Reflect on Data and Feedback to Improve Their Learning	• **D1-2:** The teacher facilitates tracking of student progress — Teacher helps students track individual progress on the learning goal • **D1-9 & 10 :** The teacher chunks content into "Digestible Bites' — Teacher stops at strategic points to check for understanding • **D1-13:** Helping Students Reflect on their Learning — Teacher engages students in activities that help them reflect on their learning • **D1-18:** Helping Students Examine Errors in Reasoning — Teacher helps students deepen their knowledge by examining their own reasoning or the logic of the information as presented to them • **D1-19:** Helping Students Practice Skills, Strategies, and Processes — Teacher engages students in practice activities that help them develop fluency
	Students engage in activities with peers to Reflect on Data and Feedback to Improve Learning	• **D1-12:** Helping Students Record and Represent Knowledge — Teacher engages students in activities that help them record their understanding of new content • **D1-15:** Organizing Students to Practice and Deepen Knowledge — Teacher uses grouping in ways that facilitate practicing and deepening knowledge

Alignment: The Formative Assessment Process and Marzano Art and Science of Teaching Framework		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**Marzano Art & Science of Teaching Framework**
Collection and Use of Data	Develop and Implement Products, Tools, Observations, and Conferences that Gather Evidence of Student Learning During the Learning Process.	• **D1-8:** Previewing New Content — Teacher engages students in activities that help them link what they already know to new content • **D1-13:** Helping Students Reflect on their Learning — Teacher engages students in activities that help them reflect on their learning and the learning process
	Teachers use Formative Data to Adjust Instruction When Necessary	• **D1-24:** Noticing when Students are not Engaged — Teacher makes note of when students are not engaged and taking overt action • **D3-1:** Identifying Areas of Pedagogical Strength and Weakness within Domain 1 — Teacher identifies specific strategies and behaviors on which to improve • **D3-2:** Evaluating the Effectiveness of Individual Lessons and Units — Teacher determines how effective a lesson or unit of instruction was in terms of enhancing students achievement and identifies causes of success or difficulty
	Students use Formative Data to Adjust Learning Tactics when necessary	• **D1-20:** Helping Students Revise Knowledge — Teacher engages students in revision of previous knowledge
	Assessments are Developed and Evaluated Collaboratively to Ensure Quality and Appropriate Use	• **D4-1:** Promoting Positive Interactions with Colleagues • **D4-3:** Seeking Mentorship for Areas of Need or Interest • **D4-4:** Mentoring other Teachers and Sharing Ideas and Strategies • **D4-6:** Participating in district and School Initiatives

Alignment: The Formative Assessment Process and Thoughtful Classroom Teacher Effectiveness Framework		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**Thoughtful Classroom Teacher Effectiveness Framework**
Planning	Understand and reflect on the planning requires throughout the Formative Assessment Process	• **2.3** Differentiating instruction and assessment so students of all styles and ability levels can experience the joys of success. • **5.1** Selecting relevant standards that are appropriate to the content and grade level. • **5.2** "Unpacking" standards and turning them into clear and measurable learning goals and targets. • **6.6** Using a variety of questions and response techniques (e.g., signaling, surveying, whiteboard-response systems, Think-Pair-Share, provisional writing) to check for understanding in real time. • **7.1** Identifying critical junctures in the learning sequence, establishing targets that students must achieve at each juncture, and using a variety of formative assessment activities to help students assess their progress toward the target. • **7.8** Assigning purposeful and grade-appropriate homework for students to practice and reinforce learning.
Learning Targets	Have an awareness of the importance of clear learning targets for the purpose of teacher planning and student clarity	• **8.7** Making sure students understand what's expected of them and providing feedback as they work. • **9.5** Helping students review learning goals and targets, assess their level of achievement, and "close the gap" when goals are unmet.
	Develop and implement ways to embed learning targets within instruction	• **5.6** Assessing students' background knowledge, skill levels, and interests relative to learning goals and targets. • **7.3** Building in periodic review and guided practice opportunities to help students master key skills and content. • 8.7 Making sure students understand what's expected of them and provide feedback as they work.
	Develop clear criteria for success	• **6.5** Using modeling and think-aloud to help students understand the thinking skills, processes, and procedures they'll need to master.

Alignment: The Formative Assessment Process and Thoughtful Classroom Teacher Effectiveness Framework		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**Thoughtful Classroom Teacher Effectiveness Framework**
Questioning/ Discussion	Effective Questioning Strategies (HOTS, Wait Time, No Opt Out, Cold Calling) are used throughout the lesson	• **4.2** Engaging students in extended, higher-order thinking challenges (e.g., inquiry, investigation, problem-based learning, action research projects). • **6.6** Using a variety of questions and response techniques (e.g., signaling, surveying, whiteboard-response systems, Think-Pair-Share, provisional writing) to check for understanding in real time.
	Teacher uses a variety of Question Types (Clarifying; Probing)	• **4.4** Probing, extending, and clarifying student responses using effective questioning and recognition techniques.
	Effective Questions lead to Quality Discussion	• **2.5** Designing learning experiences that call for high levels of collaboration, discussion, and interaction among students. • **4.5** Encouraging discussion, dialogue and debate around important ideas. • **5.3** Posing essential questions to guide learning and promote deep thinking. • **5.4** Beginning lessons and units with engaging "hooks"—thought provoking activities or questions that capture student interest and activate their prior knowledge. • **7.7** Providing students opportunities to process new knowledge deeply through questions, discussions, and critical thinking activities.
Formative Feedback	Provide Students with Appropriate Feedback on Formative Tasks	• **7.3** Building in periodic review and guided practice opportunities to help students master key skills and content. • **7.4** Providing clear and descriptive feedback to help students refine their use of key skills and/or deepen their comprehension. • **8.7** Making sure students understand what's expected of them and providing feedback as they work.
	Provide students Opportunities to Reflect on Feedback and make Revisions	• **4.8** Teaching students how to use strategies on their own, as tools and frameworks for thinking and learning. • **5.8** Encouraging students to develop personal learning goals and plans for achieving them.

Alignment: The Formative Assessment Process and Thoughtful Classroom Teacher Effectiveness Framework		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**Thoughtful Classroom Teacher Effectiveness Framework**
Students as Self/Peer Assessors	Students Reflect on Data and Feedback to Improve Their Learning	• **5.8** Encouraging students to develop personal learning goals and plans for achieving them. • **6.5** Using modeling and think-aloud to help students understand the thinking skills, processes, and procedures they'll need to master. • **7.4** Providing clear and descriptive feedback to help students refine their use of key skills and/or deepen their comprehension. • **8.6** Equipping students with the planning, thinking, and self-assessment skills they need to analyze and address task demands. • **9.2** Providing students with opportunities to look back on the content so they can make generalizations, develop new insights, and/or formulate questions. • **9.3** Helping students reflect on their own learning process to identify what they did well and where they'd like to improve. • **9.4** Creating an environment that takes metacognition—or thinking about thinking—seriously. • **9.5** Helping students review learning goals and targets, assess their level of achievement, and "close the gap" when goals are unmet. • **9.6** Working with students to set future performance goals.
	Students engage in activities with peers to Reflect on Data and Feedback to Improve Learning	• **2.4** Building a classroom community that insists on respect and mutual support for each student's learning and provides opportunities for students to become familiar with each other. • **7.1** Identifying critical junctures in the learning sequence, establishing targets that students must achieve at each juncture, and using a variety of formative assessment activities to help students assess their progress toward the target. • **7.4** Providing clear and descriptive feedback to help students refine their use of key skills and/or deepen their comprehension. • **7.5** Using heterogeneous and homogeneous groups to maximize student learning.

Alignment: The Formative Assessment Process and Thoughtful Classroom Teacher Effectiveness Framework		
Formative Assessment Element/Pillar	**FA Standard & Indicators**	**Thoughtful Classroom Teacher Effectiveness Framework**
Collection and Use of Data	Develop and Implement Products, Tools, Observations, and Conferences that Gather Evidence of Student Learning During the Learning Process.	• **5.6** Assessing students' background knowledge, skill levels and interests relative to learning goals and targets. • **6.6** Using a variety of questions and response techniques (e.g., signaling, surveying, whiteboard-response systems, Think-Pair-Share, provisional writing) to check for understanding in real time. • **8.8** Differentiate assessment tasks so that students can show what they know in different ways.
	Teachers use Formative Data to Adjust Instruction When Necessary	• **2.3** Differentiating instruction and assessment so students of all styles and ability levels can experience the joys of success. • **10.1** Self-assessing and working to improve his or her own classroom practice.
	Students use Formative Data to Adjust Learning Tactics when necessary	• **7.1** Identifying critical junctures in the learning sequence, establishing targets that students must achieve at each juncture, and using a variety of formative assessment activities to help students assess their progress toward the target. • **7.4** Providing clear and descriptive feedback to help students refine their use of key skills and/or deepen their comprehension.
	Assessments are Developed and Evaluated Collaboratively to Ensure Quality and Appropriate Use	• **10.1** Self-assessing and working to improve his or her own classroom practice. • **10.3** Seeking out professional development and continuous learning opportunities. • 1**0.4** Working with colleagues to improve practice throughout the building as part of a professional learning community.

Notes

WAYNE RESA
Leading... Learning for All

SERVICE
LEADERSHIP
COLLABORATION
EXCELLENCE

Wayne RESA
33500 Van Born Road • Wayne, MI 48184-2497
734.334.1300 • 734.334.1620 FAX • www.resa.net

Board of Education
James S. Beri • Kenneth E. Berlinn • Mary E. Blackmon
Lynda S. Jackson • James Petrie

Randy A. Liepa, Ph.D., Superintendent

This work is covered under the following Creative Commons license: Attribution, Non-Commercial, ShareAlike 4.0 International
The photographs in this document are copyright protected and ***not*** covered under the Creative Commons license.