

Measuring Student Growth:

A Practical Guide to Educator Evaluation

WAYNE
RESA
Leading... Learning for All

SERVICE
LEADERSHIP
COLLABORATION
EXCELLENCE



SECTION 6: Formative Assessment

Table of Contents

■ Participants	4
■ Opening Letter by Dr. Paul Salah, Associate Superintendent of Educational Services.....	6
■ List of Tables	8
■ SECTION SIX: Formative Assessment.....	9
• Recent History of Formative Assessment	9
• Effective Formative Assessment	9
• Formative Assessment and Student Achievement.....	9
• Formative Assessment and Teacher Evaluation	10
• Committee Findings.....	11
■ References	12
■ Glossary of Terms	14

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

33500 Van Born Road
Wayne, Michigan 48184
www.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

Table #	Table Title	Page
Table 6.1	<i>Formative Assessment Alignment with Michigan Approved Teacher Evaluation Models</i>	11



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 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 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?	<ul style="list-style-type: none"> Is essential for teachers and administrators to understand the assessment data they have available and are using to define and analyze student growth. 	<ul style="list-style-type: none"> 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.	<ul style="list-style-type: none"> Provides a good visual for a measure of central tendency (true mean). 	<ul style="list-style-type: none"> 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?	<ul style="list-style-type: none"> Criterion referenced assessments are preferable in comparing student performance to previous learning or rating performance aligned to a learning expectation. 	<ul style="list-style-type: none"> 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?	<ul style="list-style-type: none"> 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. 	<ul style="list-style-type: none"> 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?	<ul style="list-style-type: none"> • Easy to calculate. • Can be used when identifying growth based on average number of students or averages of norm referenced data. 	<ul style="list-style-type: none"> • 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?	<ul style="list-style-type: none"> • 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. 	<ul style="list-style-type: none"> • 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?	<ul style="list-style-type: none"> • Identifies the gain that is most commonly demonstrated across students. 	<ul style="list-style-type: none"> • 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?	<ul style="list-style-type: none"> • 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. 	<ul style="list-style-type: none"> • 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?	<ul style="list-style-type: none"> • 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. 	<ul style="list-style-type: none"> • 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?	<ul style="list-style-type: none"> • 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. 	<ul style="list-style-type: none"> • 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?	<ul style="list-style-type: none"> • Is customized to the context of data, content, and categories. • Provides a standard against which raters classify data into categories. 	<ul style="list-style-type: none"> • 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?	<ul style="list-style-type: none"> Requires the setting of a future standard of performance and a time frame to meet the standard. 	<ul style="list-style-type: none"> 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?	<ul style="list-style-type: none"> 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. 	<ul style="list-style-type: none"> 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?	<ul style="list-style-type: none"> Relatively easy to calculate. Strong reliability indicates that the method is stable. 	<ul style="list-style-type: none"> 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 $\geq .90$.

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?	<ul style="list-style-type: none"> 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. 	<ul style="list-style-type: none"> 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?	<ul style="list-style-type: none"> 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. 	<ul style="list-style-type: none"> 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?	<ul style="list-style-type: none"> 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. 	<ul style="list-style-type: none"> 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?	<ul style="list-style-type: none"> The trajectory is set by defining a future standard and a time horizon to meet the standard. 	<ul style="list-style-type: none"> 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?	<ul style="list-style-type: none"> Validity is important to ensure the test is measuring the intended content. 	<ul style="list-style-type: none"> 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?	<ul style="list-style-type: none"> Vertical scaling provides consistent scores across grade levels and is advantageous for measuring growth. 	<ul style="list-style-type: none"> The procedure requires sophisticated statistical methods.





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