



New York State Education Department  
Office of Special Education  
**Educational Partnership**



# Assessment Essentials

## Part 1



Produced by the Technical Assistance Partnership for Academics at the University at Albany

# Disclaimer

The resources shown are designed to provide helpful information. Resources are provided for instructional use purposes only and do not constitute NYSED endorsement of any vendor, author, or other sources. To the best of our knowledge, the resources provided are true and complete.

# Introductions

- Name
- Role
- District
- School
- Population Served

# Meeting Norms

- Take care of your needs (water, food, restroom, etc.)
- Speak your truth – Use “I” statements
- Ask what you need to understand and contribute
- Listen with respect
- Push your growing edge
- Participate and struggle together
- Expect a lack of closure
- Respect each other's needs and learning styles

# Agenda

- Welcome
  - Introduction, Inclusion, and Norms
- The Purpose of Assessment
- Assessment Terminology
- Assessment Data
- Technical Adequacy
- Selecting Assessments
- Wrap-up and Survey

# Learning Objectives

Participants will be able to:

- Explain the importance of assessment for providing high quality instruction to all students.
- Describe a variety of assessments using essential terminology.
- Explain the resulting scores from various assessment types.
- Use a process for identifying high quality assessment tools that match student needs.

# Blueprint for Improved Results for Students with Disabilities



## Self-Advocacy

Students engage in self-advocacy and are involved in determining their own educational goals and plan.



## Family Partnership

Parents, and other family members, are engaged as meaningful partners in the special education process and the education of their child.



## Specially-Designed Instruction

Teachers design, provide, and assess the effectiveness of specially-designed instruction to provide students with disabilities with access to participate and progress in the general education curriculum.



## Research-Based Instruction

Teachers provide research-based instructional teaching and learning strategies and supports for students with disabilities.



## Multi-tiered Support

Schools provide multi-tiered systems of behavioral and academic support.



## Inclusive Activities

Schools provide high-quality inclusive programs and activities.



## Transition Support

Schools provide appropriate instruction for students with disabilities in career development and opportunities to participate in work-based learning.



---

# The Purpose of High-Quality Assessment

Why invest in this knowledge?

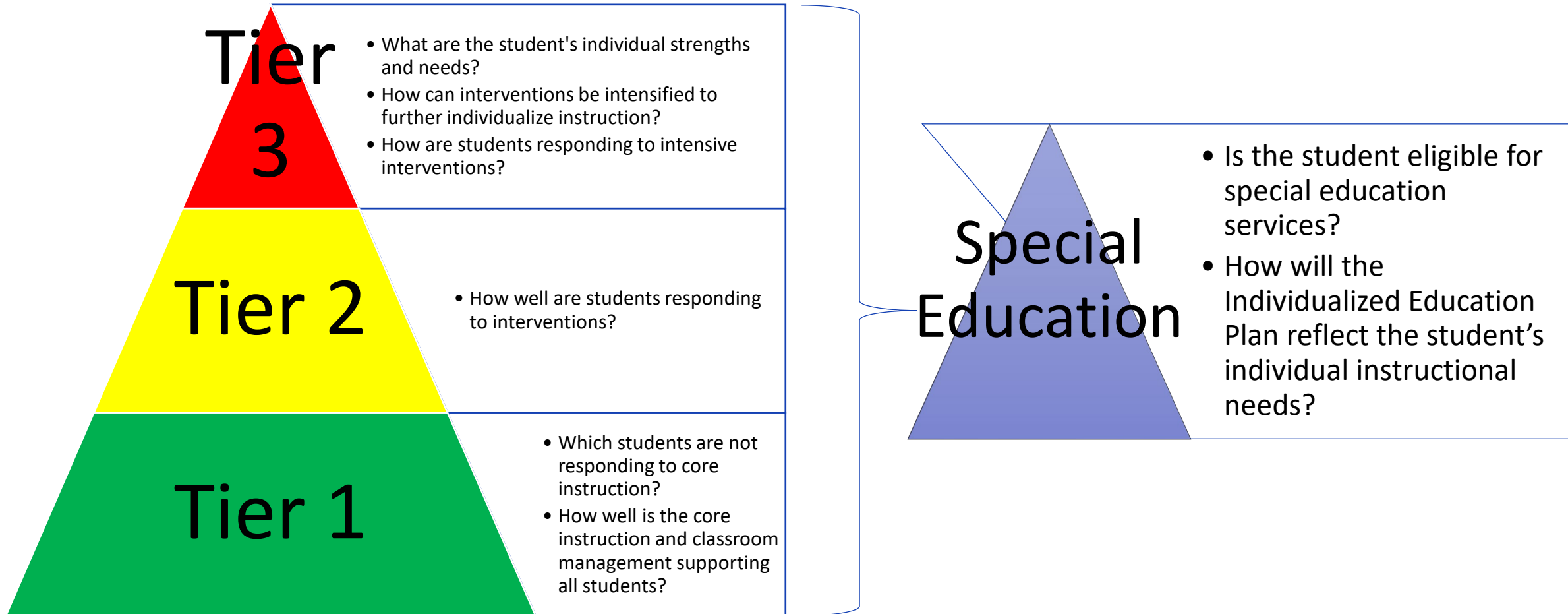


*On a scale of  
sheep, how does  
the word  
“assessment”  
make you feel and  
why?*



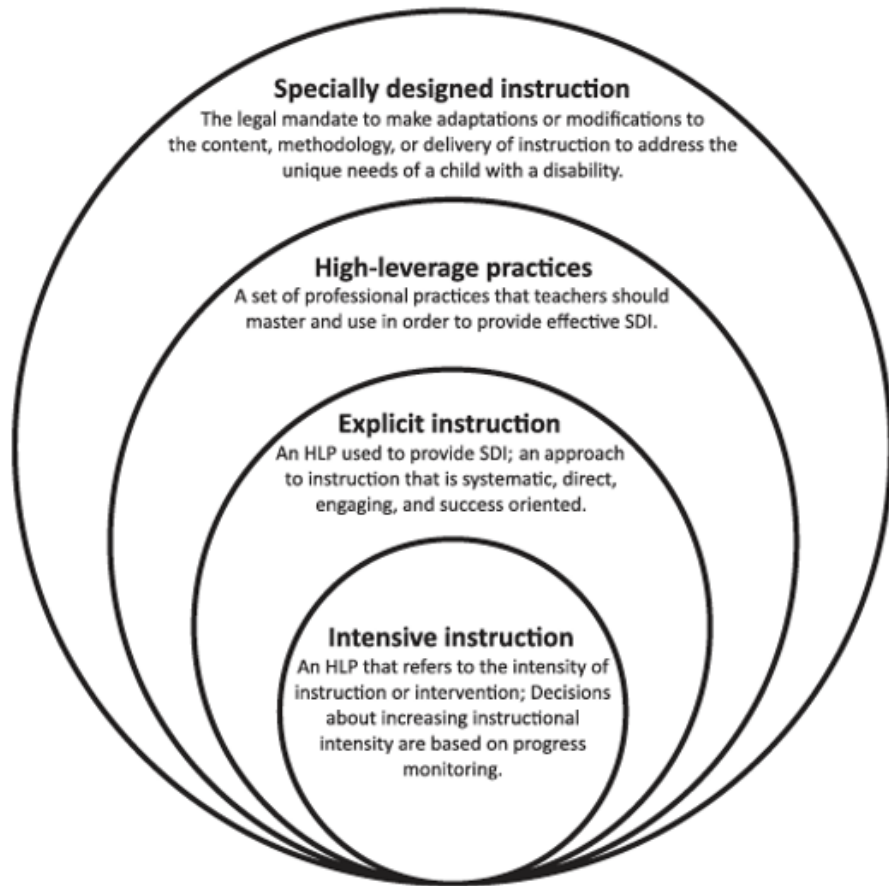
# Assessment Data Provides Answers

## Questions at Each Level of Support



# Specially Designed Instruction (SDI)

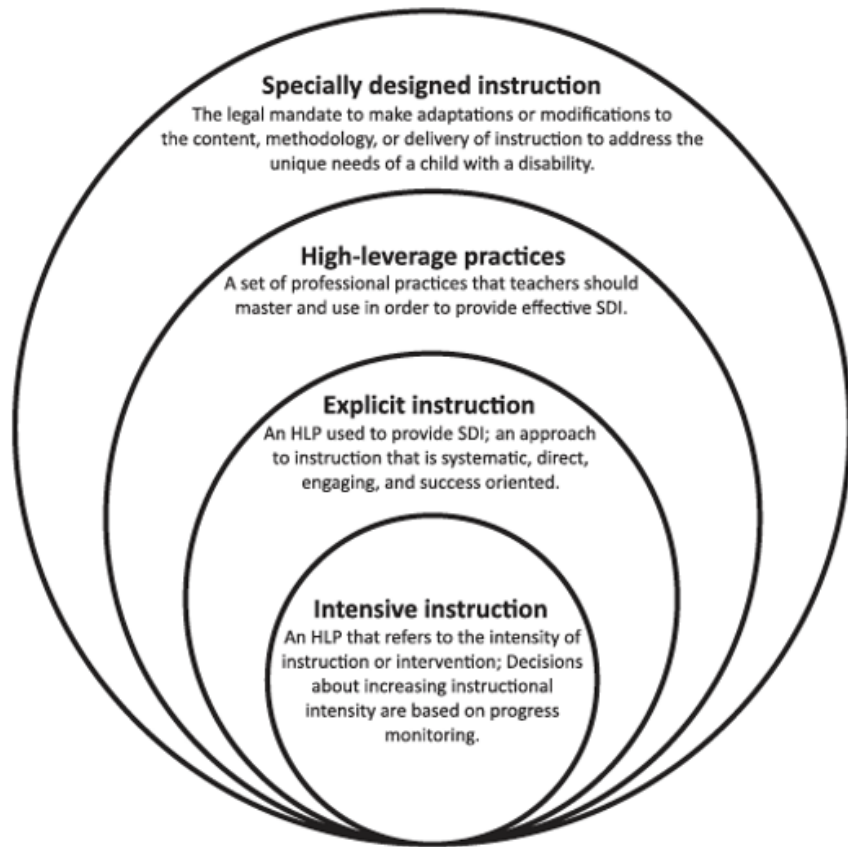
## The Nested Structure of Special Education Terms



*High-leverage practices (HLPs) “can be used as a starting point for selecting, designing, and implementing SDI.”*

Riccomini, Morano & Hughes, 2017, p. 21





## What HLPs apply to assessment?

- **HLP 4:** Use multiple sources of information to develop a comprehensive understanding of the student's strengths and needs.
- **HLP 5:** Interpret and communicate assessment information with stakeholders to collaboratively design and implement educational programs.
- **HLP 6:** Use student assessment data, analyze instructional practices, and make necessary adjustments that improve student outcomes.
- **HLP 20:** Provide intensive instruction.

## Stop & Think



## Read and Consider



1. Review one of the four assessment-related HLPs listed on the previous slide.
2. Respond to the following questions:
  - What skills and knowledge does an educator need to demonstrate this HLP?
  - What challenges to implementing these practices do you encounter with your educational organizations?

# The Effect of Meaningful Assessment

What does the research say?

- Students are more successful when assessed using high quality measurement.
- Teachers are more successful when using high-quality measurement.
- Teachers save time because instruction is more efficient.

# Assessment Terminology

---

Understanding the  
Language of Assessments





# The Language of Assessment

What is this word soup?!

Criterion-Referenced

Summative Assessment

Diagnostic Assessment

Universal Screening

High-Stakes Assessment

Standardized

Norm-Referenced

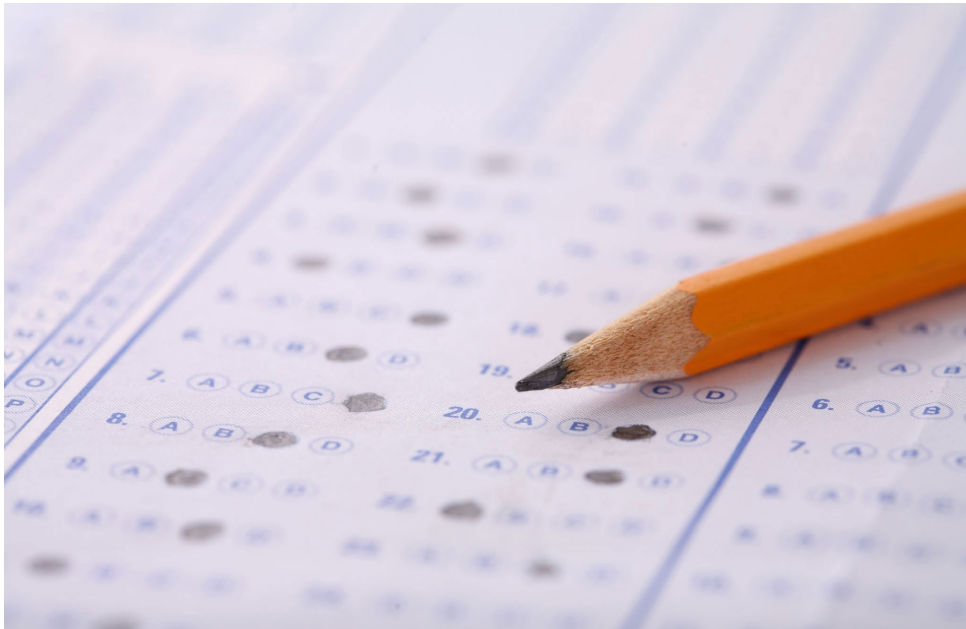
Low-Stakes Assessment

Formative Assessment

Progress Monitoring

# High-Stakes Assessment

## Describing Assessments



- Used to make high-stakes decisions, such as diagnosing a disability, or to classify a student as eligible for special education services.
- Are often individually administered using standardized procedures during longer sessions (an hour or more) and require a high level of training.

# High-Stakes Assessment

## Examples

- A student is administered the Kaufman Test of Educational Achievement (KTEA) as part of a comprehensive psycho-educational evaluation **to determine if they are eligible for special education services.**
- A school psychologist administers a Wechsler Intelligence Scale for Children (WISC) **to determine if a student demonstrates an intellectual disability.**
- A student takes the New York State (NYS) Regents Exams to determine if they will meet requirements **to receive a high school diploma.**



# Low-Stakes Assessment

## Describing Assessments



- Used to make low-stakes decisions, such as informing where to target instruction or how modify it to better meet a student's needs.
- Often require less training to learn and administer.
- Can be standardized or teacher-developed.

# Low-Stakes Assessment

## Examples

- A special education teacher gives a third-grade student with a reading disability a Dynamic Indicators of Basic Early Literacy Skills (DIBELS) oral reading fluency (ORF) test weekly **to determine progress toward a year-end reading goal on the IEP.**
- A middle school math teacher gives their students a chapter/unit test on fractions **to determine if the class is ready to move on to the next chapter/unit.**
- An ELL teacher asks students to correctly use target vocabulary words in sentences **to determine whether to move to word/meaning instruction on a new set of words.**
- A school leadership team evaluates *Acadience* Math universal screening data **to determine if current instruction is meeting the needs of students in the building.**

# Summative, Formative & Diagnostic Assessment

## Describing Assessments

Summative Assessment	Formative Assessment	Diagnostic Assessment
One-time or less frequent evaluations administered <i>after</i> content or skills are taught to determine mastery.	More frequent, often brief, evaluation of learning that occurs <i>during</i> instruction to provide feedback to learners and educators.	Evaluation of specific skills used to diagnose disabilities or more precisely guide instruction.

# Summative Assessment

## Examples

- A student is administered the Kaufman Test of Educational Achievement (KTEA) to evaluate their math and reading performance relative to other same-aged peers.
- A high school student takes the NYS Regents Exam for Algebra II at the end of the course.
- A middle school math teacher gives their students a chapter/unit test on fractions to determine if the class is ready to move on to the next chapter/unit.
- A school leadership team evaluates Acadience Math universal screening data to determine if the current curriculum is meeting the needs of students in the building.

# Formative Assessment

- Examples



- A special education teacher gives a third-grade student with a reading disability a DIBELS ORF test weekly to determine their progress toward a year end reading goal on the IEP.
- A second-grade classroom teacher asks students to each read a word with a taught spelling pattern before lining up for lunch to determine lesson mastery.



# Diagnostic Assessment

## Examples

- A psychologist administers the Autism Diagnostic Observation Schedule (ADOS) to determine if a child meets criteria for a diagnosis of Autism.
- An interventionist or special educator administers a phonics inventory to identify which phonic rules a student has mastered so that intensive instruction/intervention can be targeted to those skills.
- An interventionist or special educator analyzes the errors a student has been making in weekly math computation progress monitoring measures to identify which skills have been mastered and which skills the student needs additional support to develop.



## What exactly does “standardized” mean?

*“A standardized test is a test that is administered, scored, and interpreted in a standard manner. The goal of standardization is to ensure that the testing conditions are as close to identical as possible for all examinees.”*

*Reynolds & Livingston, 2014*



# Standardized Tests

## Examples vs. Nonexamples

Formative Assessment: Standardized	Formative Assessment: Not Standardized
<p>A special education teacher gives a fourth-grade student with a reading disability the DIBELS Maze Curriculum-Based Measure (CBM) to assess their progress toward a year-end reading comprehension goal on the IEP.</p>	<p>A fourth-grade teacher gives a student a chapter/unit test to assess comprehension/mastery of content.</p>

# What Terms Apply?

## Practice Describing Common Assessment Tools



- Let's review a series of common assessment tools to practice applying what we've learned!
  1. Review the description of the tool.
  2. Decide which terms you could use to describe the tool.

# Wechsler Individual Achievement Test (WIAT-4)

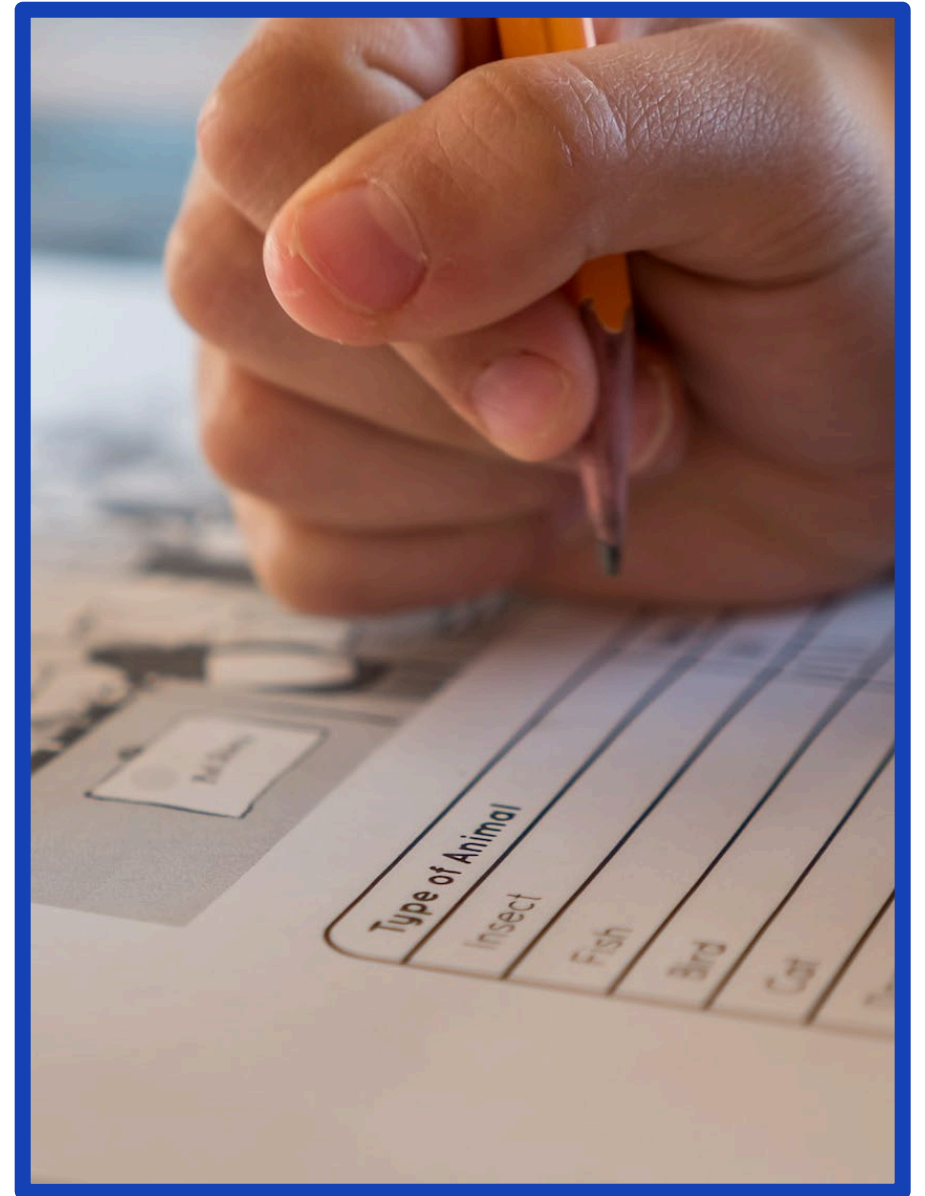
- Lengthy administration (1 hour +).
- Requires high level of training and credentials to give and interpret.
- Standardized on a large national sample of students.
- Can be used for:
  - Determining if a student is eligible for special education services.
  - Evaluating a student's overall academic performance (reading, math, writing).



<b>High-Stakes Assessment</b>	<b>Low-Stakes Assessment</b>
<b>Summative Assessment</b>	<b>Formative Assessment</b>
<b>Diagnostic Assessment</b>	<b>Standardized Assessment</b>

# Pop Quiz

- Teacher-developed, brief, and given to all students in a class/group.
- Can be used for:
  - Informing decisions about how to adjust instruction based on student performance.
  - Determining if students are acquiring the skills or knowledge currently being taught.

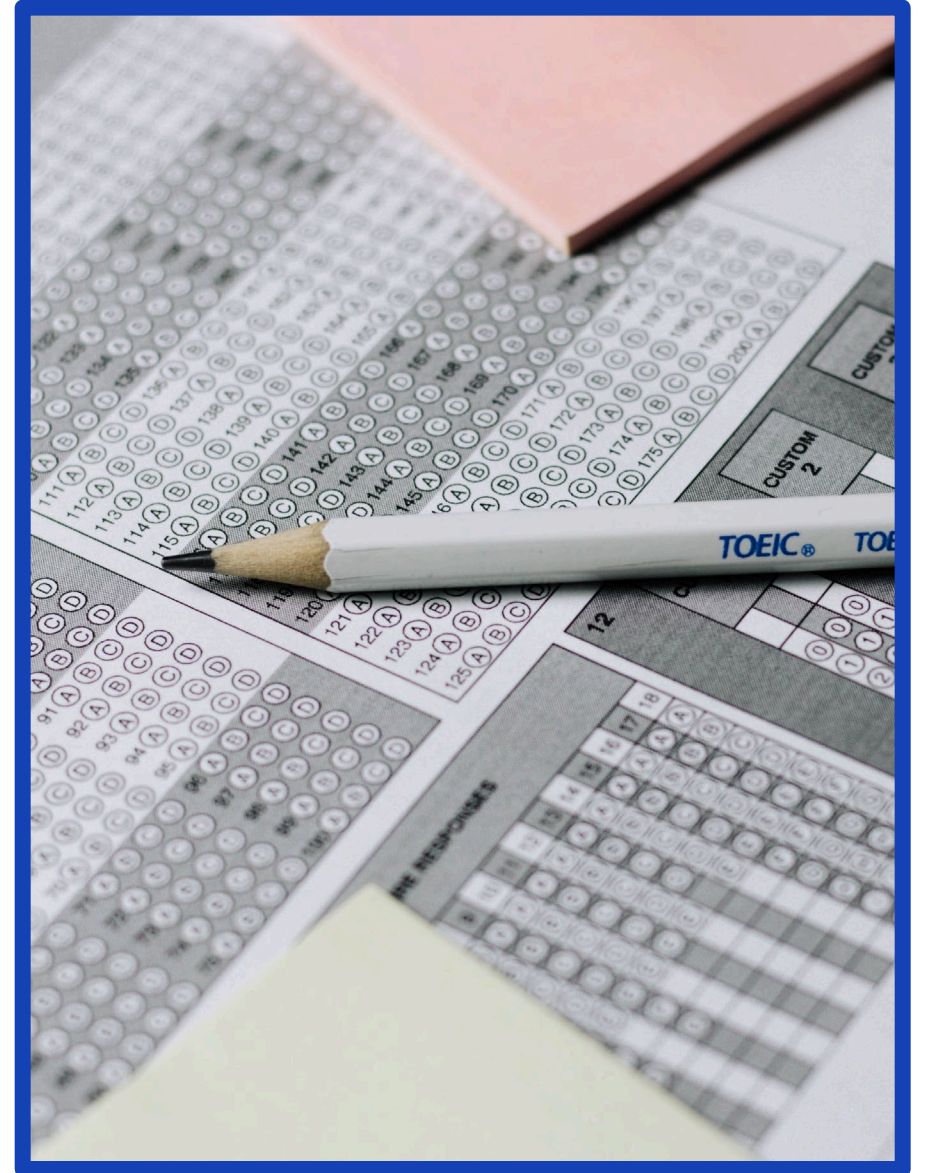


<b>High-Stakes Assessment</b>	<b>Low-Stakes Assessment</b>
<b>Summative Assessment</b>	<b>Formative Assessment</b>
<b>Diagnostic Assessment</b>	<b>Standardized Assessment</b>



# Regents Exam

- Lengthy tests taken once at the end of a course.
- Standardized on a large regional sample of students.
- Can be used for:
  - Determining if the student has acquired content taught in the course.
  - Determining if a student can receive a high school diploma.



<b>High-Stakes Assessment</b>	<b>Low-Stakes Assessment</b>
<b>Summative Assessment</b>	<b>Formative Assessment</b>
<b>Diagnostic Assessment</b>	<b>Standardized Assessment</b>

# Phonics Inventory

- Individually-administered inventory of phonics rules that is easily administered by teachers.
- Directions are provided; phonics progression is research-based but has not been standardized on a normative sample.
- Can be used for:
  - Targeting phonics instruction for a small group or individual students in need of intervention.

Page 1	
<input type="radio"/> tade (made)	tad
<input type="radio"/> waig (plague)	wag
<input type="radio"/> fesp (nesp)	fish
<input checked="" type="radio"/> kur (fur)	
<input type="radio"/> nasp (clasp)	nap
<input type="radio"/> foat (goat)	foot

Page 2	
<input type="radio"/> seg (leg)	sing
<input type="radio"/> keal (heal)	kill
<input type="radio"/> sug (tug)	song
<input type="radio"/> pide (side)	
<input type="radio"/> leat (seat)	
<input type="radio"/> yad (bad)	

<b>High-Stakes Assessment</b>	<b>Low-Stakes Assessment</b>
<b>Summative Assessment</b>	<b>Formative Assessment</b>
<b>Diagnostic Assessment</b>	<b>Standardized Assessment</b>

# DIBELS Word Reading Fluency

- Quick and easy to administer (1 minute).
- Standardized on a large national sample of students.
- Can be used for:
  - Adjusting or targeting instruction for groups of students or individual students.
  - Screening large groups of students.
  - Evaluating overall word reading/decoding skill after instruction has occurred.
  - Evaluating progress toward word reading/decoding goals as often as twice per week.
  - Analyzing errors to provide more detailed skill information to inform instruction.

DIBELS 8 <sup>th</sup> Edition <i>Word Reading Fluency</i>				Benchmark WRF 1.Beginning	
Examiner script			Reminders		
Please read from this list of words (Point to the student form).			Start timer	When student says the first word.	
Start here (point to the first word) and go across the page (point across the page).			Prompts	Student hesitates: wait 3 seconds; give correct word; point to the next word, and say "Keep going"; mark the missed word as incorrect.	
When I say 'Begin', point to each word and read it the best you can. If you get stuck, I will tell you the word, so you can keep reading. Put your finger on the first word. Ready? Begin.			Discontinue	Student does not get any words correct within the first line (5 words): discontinue WRF; <u>do not administer ORF</u> .	
no	they	is	we	it	(5)
if	one	but	not	has	(10)
for	there	a	you	be	(15)
wall	help	father	call	black	(20)
alive	sports	meeting	above	island	(25)
came	stop	show	open	sky	(30)
further	front	story	always	feed	(35)
station	deep	across	paper	driver	(40)
powerful	double	still	often	top	(45)
first	note	count	none	against	(50)
shown	head	room	same	sure	(55)
off	nice	speak	distance	right	(60)
line	stay	allow	come	she	(65)
turn	peace	well	bank	hard	(70)
news	engine	race	heat	other	(75)
never	east	team	rose	when	(80)
party	share	complete	sea	high	(85)
switch	spent	job	listen	sick	(90)
getting	film	think	break	eat	(95)
huge	while	fear	wave	bit	(100)
morning	hole	safe	enter	picture	(105)
Total Correct _____					

©2011 University of Oregon. All rights reserved. Page 6

<b>High-Stakes Assessment</b>	<b>Low-Stakes Assessment</b>
<b>Summative Assessment</b>	<b>Formative Assessment</b>
<b>Diagnostic Assessment</b>	<b>Standardized Assessment</b>

# Assessment Data



What do the scores tell  
us?

# Norm-Referenced and Criterion-Referenced

## Interpreting Test Scores

Norm-Referenced Assessments	Criterion-Referenced Assessments
<p>Assessment tools that compare a student's score to those of a normative group to give a sense of where they fall in relation to others.</p> <p>Utilize standardized assessment procedures.</p>	<p>Assessment tools that compare a student's score to a specific performance level, or criterion that indicates skill or knowledge level and is predictive of later success.</p> <p>May or may not utilize standardized procedures.</p>
<p>Answers the question, "How does this student perform relative to other students?"</p>	<p>Answers the question, "How close is this student to a meaningful learning target?"</p>

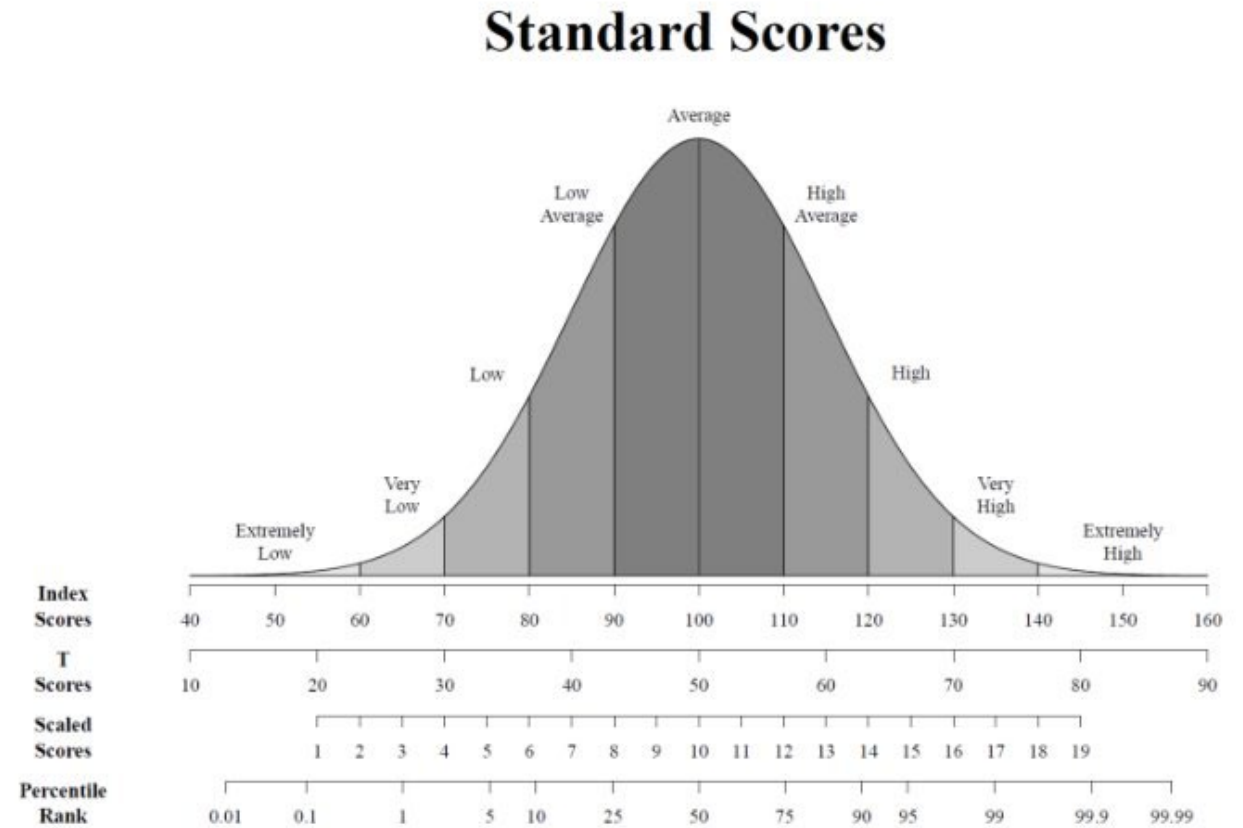


# Score Type

- Any measure can be made to be either normative or criterion-referenced.
- Certain statistics you see are specific to one or the other system.
- Normative systems may include standard scores, t-scores, percentiles, and grade/age equivalencies.
- Criterion-referenced systems may include odds ratios, grades, and predictive thresholds.

# Normal Curve with Scores

## Normative Scores



# Technical Adequacy

---

How do we determine the quality of assessment tools?



# Technical Adequacy Jigsaw



## Group Activity



1. Individually review your assigned handout.
2. With your group, provide the following:
  - Brief definition or key terms.
  - Simple example.
  - Connection or key take-away.

# Reliability

- What is it?
  - Reliability is the consistency of a set of scores that are designed to measure the same thing.
- Simple Example
  - Five different scales consistently provide the same weight for a watermelon at the grocery store.
  - Consistently hitting the same spot on a target.



# Validity

- What is it?
  - Validity is broadly defined as how well something measures what it is supposed to measure.
- Simple Example
  - A speedometer that shows you are traveling at 25 miles per hour, when your actual speed is 65, is not a valid measure of your car's speed.
  - A speedometer that shows you are traveling at a speed of 65 miles an hour, when that is in fact your actual speed, is providing a valid measure of your car's speed.



# Classification Accuracy

- What is it?
  - Provides a picture of how well scores on a screening assessment correctly identify students at risk.
- Simple Example
  - How accurate are the security scanners used by the Transportation Security Administration (TSA) in identifying unallowable material (true positives) versus allowable material that causes the scanners to go off (false positives)?



# Sample Representativeness

- What is it?
  - A representative sample is a group that closely matches the characteristics of its population as a whole.
- Simple Example
  - A baker's sample batch of 10 cookies includes 6 chocolate chip and 4 sugar cookies to match the entire batch's proportions of 100 cookies (60 chocolate chip and 40 sugar).





# Usability & Feasibility

## Additional Technical Adequacy Considerations

- Efficiency
  - How long will the assessment take to administer?
  - What qualifications are needed to administer the assessment?
  - Is the assessment administered individually or group administered?
- Cost
  - How much will the assessment cost to buy, learn, and continue to fund?
- Versatility
  - Can the assessment serve multiple purposes?
  - If relevant, does it come in multiple languages?
- Instructional Relevance
  - Will the scores offer me information that I can use to improve instruction?
- Others?

# Guiding EO's in Using Criteria to Select Assessments

---

Modeling How to Identify  
the Best Tool for Your  
Needs

# Putting It Together

How do I know what to use?

1. Identify Your Purpose and Priorities
2. Gather Aligned Assessment Tools
3. Evaluate Technical Adequacy and Usability/Feasibility
4. Make an Informed Choice
5. Review Your Decision Regularly

# Step 1. Identify Your Purpose and Priorities

## Selecting an Assessment

Step 1	Case Scenario
<p>Describe the purpose of the assessment tool you are searching for.</p> <ul style="list-style-type: none"><li>• Are you planning to engage in high-stakes or low-stakes assessment?</li><li>• Will this be for the purposes of summative assessment, formative assessment, or diagnostic assessment?</li><li>• Do you prefer the measure to be standardized, simple to administer, brief, etc.?</li></ul>	<ul style="list-style-type: none"><li>• A first-grade classroom teacher is providing a decoding intervention (Consonant-Vowel-Consonant words) to small groups.</li><li>• He needs a tool he can use to monitor the group's progress to make instructional changes.</li><li>• He also hopes to use the assessment to help him more precisely group students and start instruction.</li><li>• It needs to be quick to administer, easy to interpret, reliable, and valid.</li></ul>

# Step 2. Gather Aligned Assessment Tools

## Selecting an Assessment

Step 2	Case Scenario
<ul style="list-style-type: none"><li>• Gather a selection of assessment tools that may meet your needs and address priorities.</li><li>• Consult reliable sources, such as the National Center for Intensive Intervention's (NCII) Tools charts (<a href="https://intensiveintervention.org/tools-charts/overview">https://intensiveintervention.org/tools-charts/overview</a>) for ideas of high-quality assessments that may meet your needs.</li></ul>	<p>The teacher gathers the following:</p> <ul style="list-style-type: none"><li>• A decoding subtest from the school psychologist's standardized achievement battery.</li><li>• A set of decoding assessments from a published curriculum labeled "decoding progress monitoring measures".</li><li>• A set of progress monitoring measures called Nonsense Word Fluency (NWF) from Acadience Learning.</li></ul>

# Step 3. Evaluate Technical Adequacy and Usability/Feasibility

## Selecting an Assessment

Step 3	Case Scenario
<p>Gather and review information on reliability, validity, classification accuracy (if relevant), representativeness, and usability/feasibility.</p> <ul style="list-style-type: none"><li>Tools charts available from NCII provide a summary of this information for some assessment tools (<a href="https://intensiveintervention.org/tools-charts/overview">https://intensiveintervention.org/tools-charts/overview</a>), or you may need to seek this information from publishers or other locations.</li></ul>	<p>The teacher gathers information on the technical adequacy and usability features of the assessments by consulting the NCII Tools Charts for the Acadience measure, as well as asking the school psychologist to assist in finding necessary information on the additional measures he is comparing.</p>

# Step 3. Evaluate Technical Adequacy and Usability/Feasibility Continued

## Selecting an Assessment

Decoding Subtest	Progress Monitoring Measures from a Published Curriculum	NWF (e.g., Acadience, DIBELS, easyCBM)
<ul style="list-style-type: none"> <li>Standardized and norm-referenced</li> <li>Reliable and valid measure of word reading skill</li> <li>Takes 5-20 minutes to administer individually; requires additional training</li> <li>Should only be administered every 6-12 months</li> <li>Provides a standard score (100=50<sup>th</sup> percentile)</li> <li>Test booklets are expensive</li> </ul>	<ul style="list-style-type: none"> <li>Has not been standardized or researched with a norm group</li> <li>Reliability and validity are unknown</li> <li>Takes 1-3 minutes to administer, individually</li> <li>Provides a score of percent correct</li> </ul>	<ul style="list-style-type: none"> <li>Standardized and norm-referenced</li> <li>Reliable and valid as a screening and progress monitoring measure</li> <li>Developed to be administered as frequently as twice per week</li> <li>Takes 2 minutes to administer individually, can be scored on the computer</li> <li>Errors can be analyzed to provide diagnostic skill data</li> <li>Provides correct letter sounds per minute</li> <li>Available free of charge online</li> </ul>

# Step 4. Make an Informed Choice

## Selecting an Assessment

Step 4	Case Scenario
Weigh all the information you have gathered with your purpose and priorities to make the best choice of an assessment tool to use.	What choice would you recommend?



# Step 4. Make an Informed Choice

## Selecting an Assessment

Step 4	Case Scenario
<p>Weigh all the information you have gathered with your purpose and priorities to make the best choice of an assessment tool to use.</p>	<p>The teacher weighed the various aspects of the assessments and decided that the Acadience NWF measure best meets his needs.</p> <ul style="list-style-type: none"><li>• This measure has been specifically designed and researched for the purposes of progress monitoring.</li><li>• As a CBM that is standardized, it is reliable and valid, but also quick to administer, available for free online, and provides information that is easy to interpret.</li><li>• The measure can also be used as a universal screening tool and for diagnostic data.</li></ul>

# Step 5. Review Your Decision Regularly

## Selecting an assessment

Step 5	Case Scenario
<p>Engage in ongoing evaluation of the assessment tool you have selected.</p> <ul style="list-style-type: none"><li>• Does it meet your expectations and needs?</li><li>• Is it continuing to address the needs of your learner(s) as your instruction or students change?</li></ul>	<p>The teacher begins assessing students weekly using the Acadience NWF measure.</p> <ul style="list-style-type: none"><li>• He finds that it provides him with useful data that he uses to adjust the intensity of his instruction, adjust groups, and determine student progress.</li><li>• Several students achieve their decoding goals after several weeks of intervention, and the teacher decides it is no longer appropriate to monitor them with this measure.</li><li>• Given how well it has worked, he looks to more advanced measures of reading skill, such as word reading fluency and ORF, that Acadience and other CBM systems offer to begin using.</li></ul>

# Summary

- High-quality assessment leads to high-quality instruction and improved student learning.
- Selecting an evidence-based assessment aligned with instructional priorities and learning goals is essential.
- Utilize the correct score type, criterion or norm-referenced, based on how you intend to use the data.
- It is critical that assessment are technically adequate based on reliability, validity, classification accuracy, and sample representativeness.
- Educators should also consider the usability and feasibility of any assessment they wish to employ.

# Exit Ticket



- Let us know how this content can be improved to meet the needs of your work.
- Share a next step you would like to take with this content!
  - Is there a teacher who might benefit from using this process?
  - Was there a resource shared that you plan to learn more about?
  - Other ideas?

# Questions and Answers

