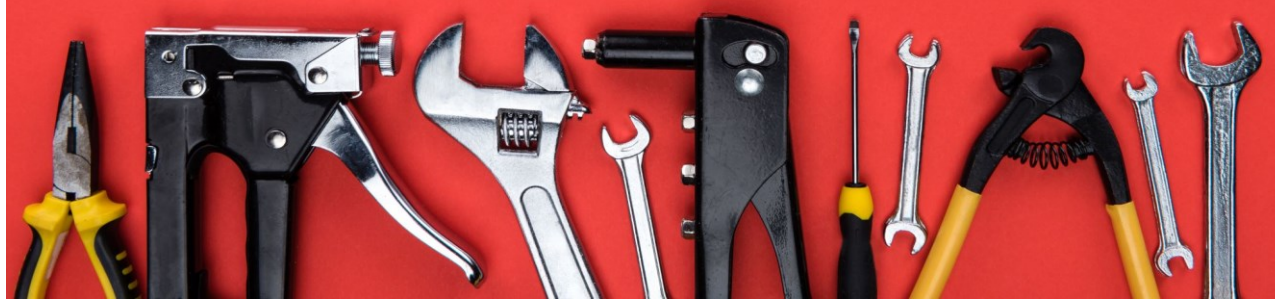
New Jersey Branch of the International Dyslexia Association
Tools of Assessment: Pathways to Literacy

Melissa Farrall, PhD
Affiliated Distinguished Scholar
Stern Center for Language and Learning
mfarrall@sterncenter.org
Williston, VT



1

The Plan for Today

- Big Picture
- What you need to know about tests and test scores
- Small Differences are Important
- Phonological Awareness: Test Roulette
- Screeners: Acceptable Risks
- Orthography In Vogue
- Fluency: Different Strokes
- Putting it Altogether

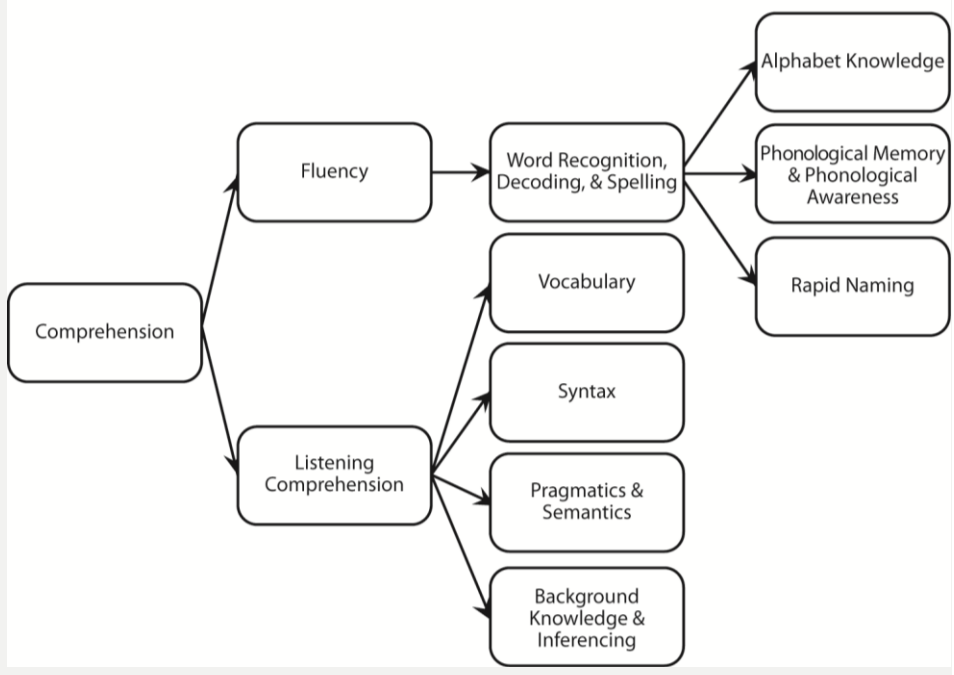
2

Simple View of Reading: Template for A Reading Evaluation

Gough & Tunmer, 1986; Hoover & Gough, 1990

3

Overall Structure of a Focused Reading Evaluation



4

Evaluation for a Specific Learning Disability

- Sometimes a focused reading assessment occurs in the context of a much larger evaluation.
- According to IDEIA 2004, a Specific Learning Disability is:
  - *a disorder in **one or more of the basic psychological processes** involved in **understanding or in using language, spoken or written**, which disorder may manifest itself in the **imperfect ability** to listen, think, speak, read, write, spell, or do mathematical calculations…*
- We might opt to assess other domains of intelligence:
  - *Verbal Comprehension,*
  - *Spatial Thinking,*
  - *Fluid Reasoning,*
  - *Working Memory, Processing Speed, and other aspects of Executive Function.*
- Do not forget vision and hearing.
- History of Instruction, Development, and Health

5

When we test our students, we want to know several things:

- *what skills they have,*
- *what skills they need to work on, and*
- *how they are performing with respect to other children of the same age or grade.*

6

Think of a test as the lens through which we peek into a student's head. We need the right tool for the job.

We have different types of tests at our disposal:

- Criterion Referenced Tests,
- Screeners/Benchmark Testing/Progress Monitoring Probes, and
- Standardized Norm-Referenced Tests.

7

Criterion-Referenced Tests

Criterion-referenced tests help determine whether a student has mastered a specific body of knowledge. They are typically designed and administered by classroom educators.

We can learn about large domains of expertise such as the events that led to WWII.

We can learn about very specific domains of expertise such as the rules for representing /k/, or the closed syllable pattern.

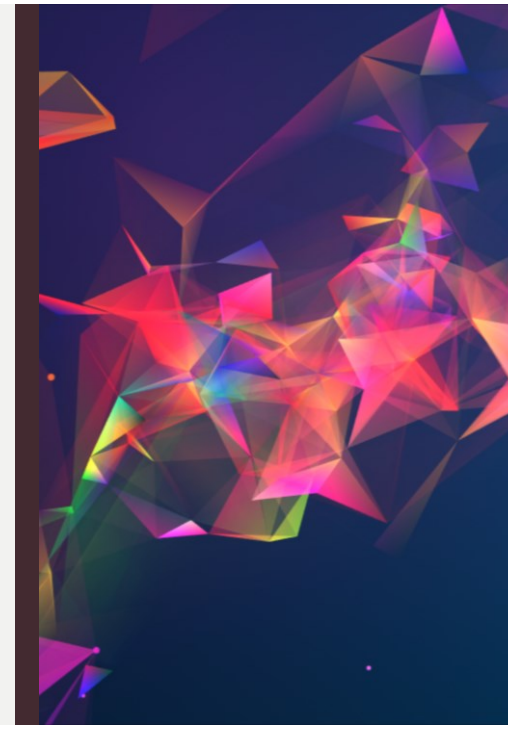
8

Criterion-Referenced Tests

Examples of summary statements based on criterion-referenced testing:

- *Sue earned a 93% on the unit test about volcanoes.*
- *Charlie identified 9/10 CVC words correctly and with automaticity.*
- *Ming read the 4th grade passage at a rate of 70 words correct per minute with 80 percent accuracy.*

- Well-designed criterion-referenced tests help us make decisions about mastery and how to pace our instruction. When we take data as part of a structured literacy lesson, we are essentially implementing a mini criterion-referenced test.



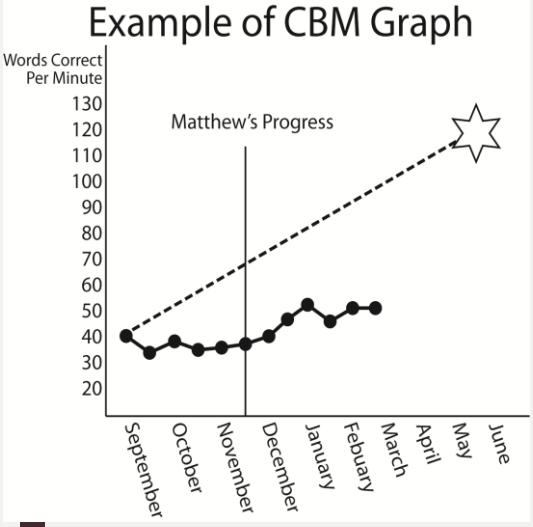
9

Sensitivity in Testing: Progress Monitoring Tools

- Progress monitoring tools were developed with several goals in mind:
  - *Permit teachers to document student progress in the regular classroom over the short term without having to rely on specialists;*
  - *Identify risk status and progress towards benchmarks; and*
  - *Establish whether an intervention is working or whether it would need to be changed.*
  - *They should also be low in cost and easy to score.*

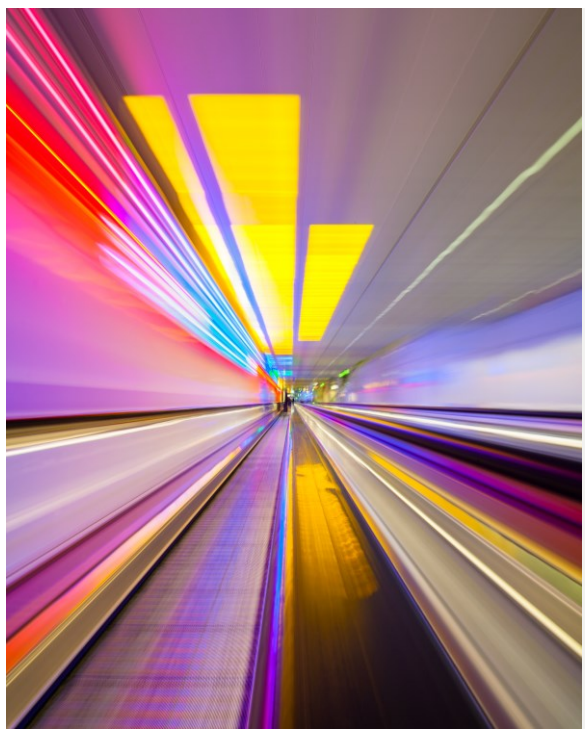
10

Making Decisions Using CBM Data



Screening/benchmarking: three times a year. Progress monitoring: up to two times a week for at-risk students	
Look at the last 3 data points. If the data points are:	
Close to the goal line; some above and some below	Your student is progressing appropriately. Continue your instruction as implemented.
All above the goal line	Your student is doing well. You might want to contemplate increasing your goal.
All below the goal line	Your student is not progressing as we hope and expect. Change your instruction.

11



Progress Monitoring Caveats

- Progress monitoring probes are designed to measure progress toward a benchmark. They are not a substitute for criterion-referenced or norm-referenced tests.
- Progress monitoring tools are not diagnostic in nature, and they will not provide specific information regarding skills that are mastered or those that are problematic.

12

Standardized Norm-Referenced Tests

Norm-referenced tests do not assess mastery but rather how a particular student compares to their peers by age or by grade (i.e., the norm group/sample). In this way, we can determine the severity of a weakness or the magnitude or a strength.

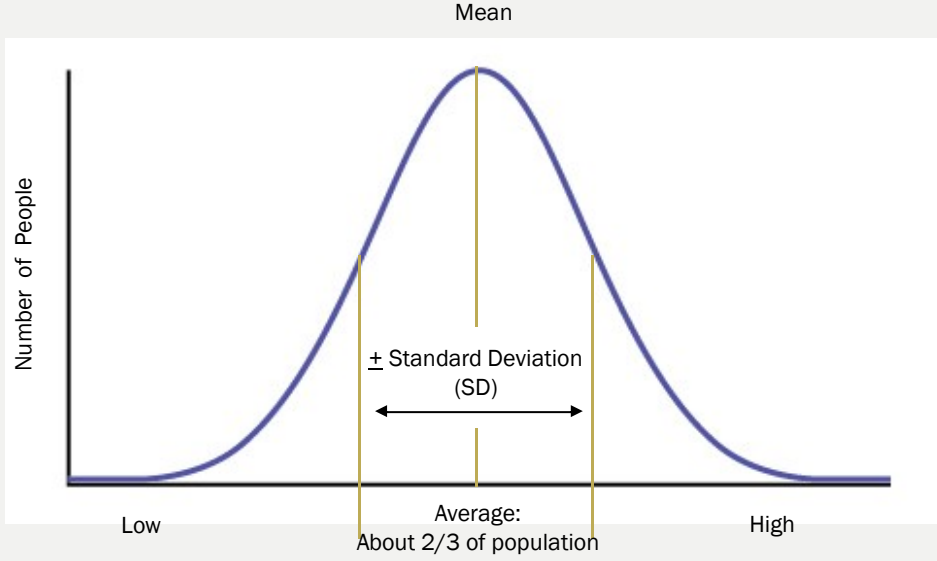
The norm sample is designed to reflect current U.S. Census data.

Because the scores are based on a comparison, it is important that each student experience the test in the exact same way (**standardization**).

Race and Cultural Identity
Geography
Intelligence
Gender
Age
Grade
Sociocultural
Acculturation of Parents

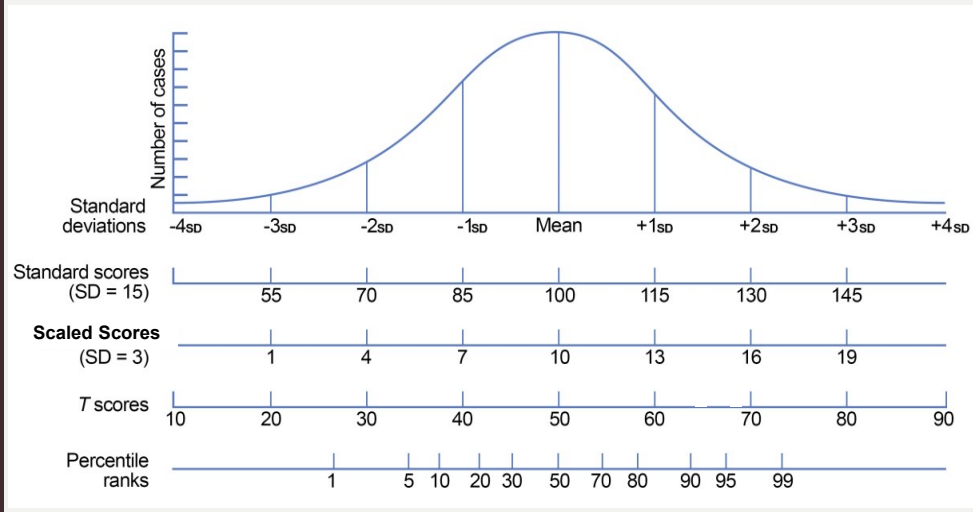
13

The Bell Curve



14

Normal Distribution, Standard Deviation and Scoring Systems:

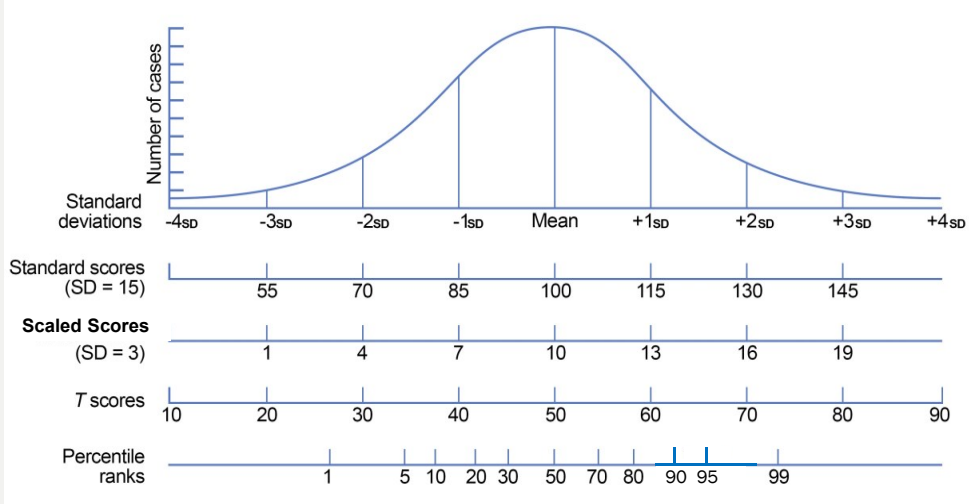


Raw scores (points for correct responses) are converted into a variety of different scoring systems.

Scaled Scores: Scoring System for the Day,

15

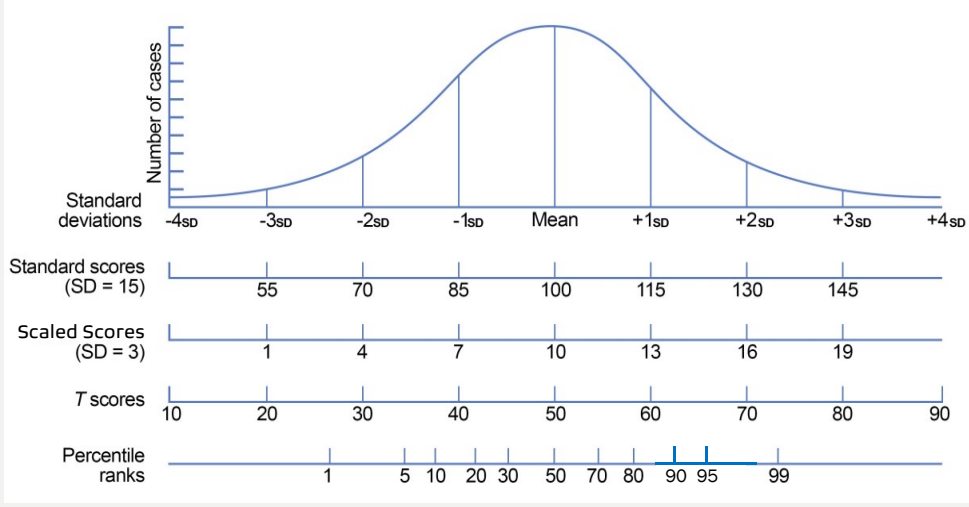
Percentile Ranks



Not Equidistant

16

What is Absent from This Picture?



17

Age/Grade Equivalents: Not What They Appear to Be

- Sasha earned a grade equivalent of 4.2 on the Anybody-Can-Do-It Test.
- Age/Grade Equivalents are misunderstood. They are **not the same** as the grade levels reported by criterion-referenced tests.
- On a **criterion-referenced test**, we might say that Sasha completed the fourth-grade level items, and we could draw the conclusion that Sasha demonstrated skill at the fourth-grade level.
- On **norm-referenced tests,** age/grade equivalents **do not** specify instructional levels. Age/Grade Equivalents provide a level that is based on the **average grade placement** of **all the students in the norming sample** who earned the same raw score.

18

How Age/Grade Equivalents are Calculated

ABC Test Raw Score Total	Actual Grade Placement	Grade Equivalent
1	1, 1, 2	1.3
2	1, 2, 2, 2	1.8
3	2, 2, 2	2.0
4	3, 3, 2, 2, 2	2.4
5	2, 2, 4, 8	4.0
6	3, 3, 5, 9	5.0

- A child with a raw score of 1 would receive a G.E. of 1.3.
- A child with a raw score of 2 would receive a G.E. of 1.8.....

19

The Ugly Truth

- Age/grade equivalents do not specify a particular grade placement or level of instruction. (See next slide.)
- Age/grade equivalents are not linked to standards for what is taught at any given point in a school year.
- Age/grade equivalents from different tests are not comparable.
- Age/grade equivalents are not equal units, and they cannot be subtracted or added. We cannot say that Adam made one year of progress in math when he moved from a grade equivalent of 3.2 to a grade equivalent of 4.2.
- Students with the same grade equivalent may have very different profiles.

20

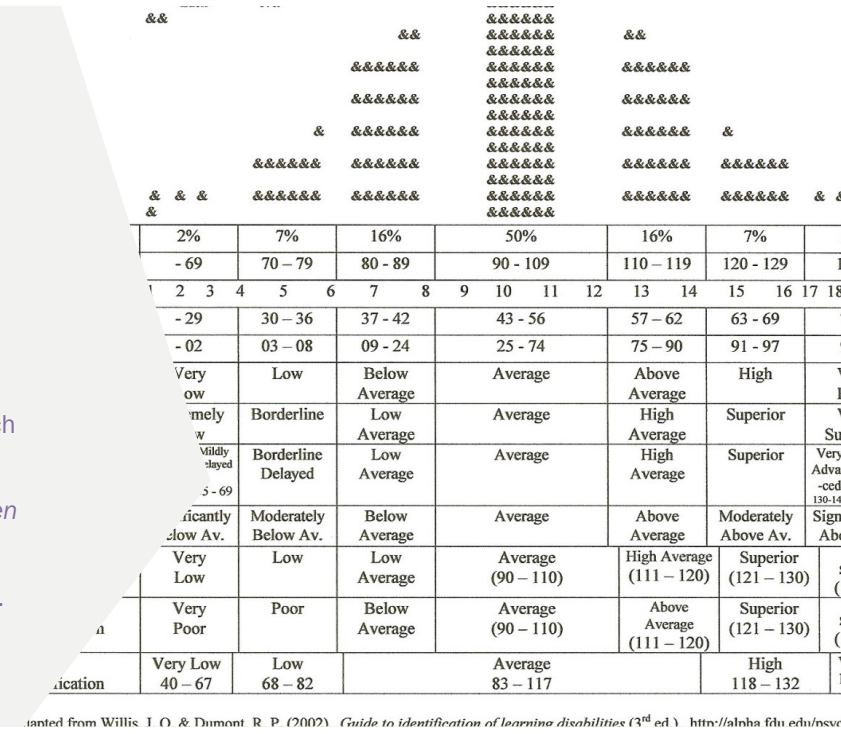
Age/Grade Equivalents: Two Students with the Same Raw Score

Test Item	Sasha	Pasha
1. Addition	1	1
2. Subtraction	1	0
3. Subtraction with Regrouping	1	0
4. Multiplication – Single Digit	1	1
5. Multiplication - Multidigit	0	1
6. Short Division	0	1
7. Long Division	0	0
Total Raw Score	4	4

21

"*Qualitative descriptors are only suggestions and are not evidence-based; alternate terms may be used as appropriate*" [emphasis in original].

Wechsler, D. (WISC-V Research Directors, S. E. Raiford & J. A. Holdnack) (2014). *Wechsler Intelligence scaled for Children* (5th ed.): *Technical and Interpretive Manual.* Bloomington, MN: Pearson, p. 152.]



	2%	7%	16%	50%	16%	7%	2
	- 69	70 – 79	80 - 89	90 - 109	110 – 119	120 - 129	13
1	2 3 4	5 6	7 8	9 10 11 12	13 14	15 16	17 18
	- 29	30 – 36	37 - 42	43 - 56	57 – 62	63 - 69	7
	- 02	03 – 08	09 - 24	25 - 74	75 – 90	91 - 97	9
	Very ow	Low	Below Average	Average	Above Average	High	Ve Hi
	mely w	Borderline	Low Average	Average	High Average	Superior	Ve Supe
	Mildly elayed 5 - 69	Borderline Delayed	Low Average	Average	High Average	Superior	Very Advan -ced 130-144
	ficantly elow Av.	Moderately Below Av.	Below Average	Average	Above Average	Moderately Above Av.	Signif Abov
	Very Low	Low	Low Average	Average (90 – 110)	High Average (111 – 120)	Superior (121 – 130)	Su (13
n	Very Poor	Poor	Below Average	Average (90 – 110)	Above Average (111 – 120)	Superior (121 – 130)	Su (13
	Very Low 40 – 67	Low 68 – 82	Average 83 – 117		High 118 – 132		Ve 13

anted from Willis, J. O. & Dumont, R. P. (2002). *Guide to identification of learning disabilities* (3rd ed.). http://alpha.fdu.edu/psych

22

There are 200 &s. each && = 1%.

110

	& & & &	& &&&&&& &&&&&&	&& &&&&&& &&&&&& &&&&&& &&&&&&	&& &&&&&& &&&&&& &&&&&& &&&&&& &&&&&& &&&&&&	&& &&&&&& &&&&&& &&&&&& &&&&&& &&&&&& &&&&&&	& &&&&&& &&&&&&	& & & &
Standard Scores	– 69	70 – 79	80 – 89	90 – 109	110 – 119	120 – 129	130 –
Percentile Ranks	– 02	03 – 08	09 – 24	25 – 74	75 – 90	91 – 97	98 –
WISC-V Classification	Extremely Low	Very Low	Low Average	Average	High Average	Very High	Extremely High
FAR & RIAS Classification	Significantly Below Av.	Moderately Below Av.	Below Average	Average	Above Average	Moderately Above Av.	Significantly Above Av.
Woodcock-Johnson Classif.	Very Low	Low	Low Average	Average (90 – 110)	High Average (111 – 120)	Superior (121 – 130)	Very Superior (131 –)
KTEA-3 15-pt. Classification	Very Low 40-54 / Low 55-69	Below Average 70 – 84		Average 85 – 115	Above Average 116 – 130		High 131-145 / Very High 146–160
KTEA-3 10-pt. Classification	Very Low – 69	Low 70 – 79	Below Average	Average (90 – 109)	Above Average	High 120 – 129	Very High 130 –
WIAT- 4 Classification	Very Low <55 / Low 55 – 69	Below Average 70 – 84		Average 85 – 115	Above Average 116 – 130		Super-ior 131-145 / Very Super-ior 146 –
Stanines	Very Low – 73	Low 74 – 81	Below Average 82 - 88	Low Average 89 – 96 / Average 97 – 103	High Average 104 - 111 / Above Average 112 – 118	High 119 – 126	Very High 127 –

Adapted from Willis, J. O. & Dumont, R. P., *Guide to Identification of Learning Disabilities* (3rd ed.) Peterborough, NH: Authors, 2002, pp. 39-40). Also available at http://www.myschoolpsychology.com/testing-information/sample-explanations-of-classification-labels/

23

Hallmarks of a Well-Designed Test

1. Our results are consistent over time, different forms, and different evaluators.

2. We measure skills that are important and that have been validated by research.

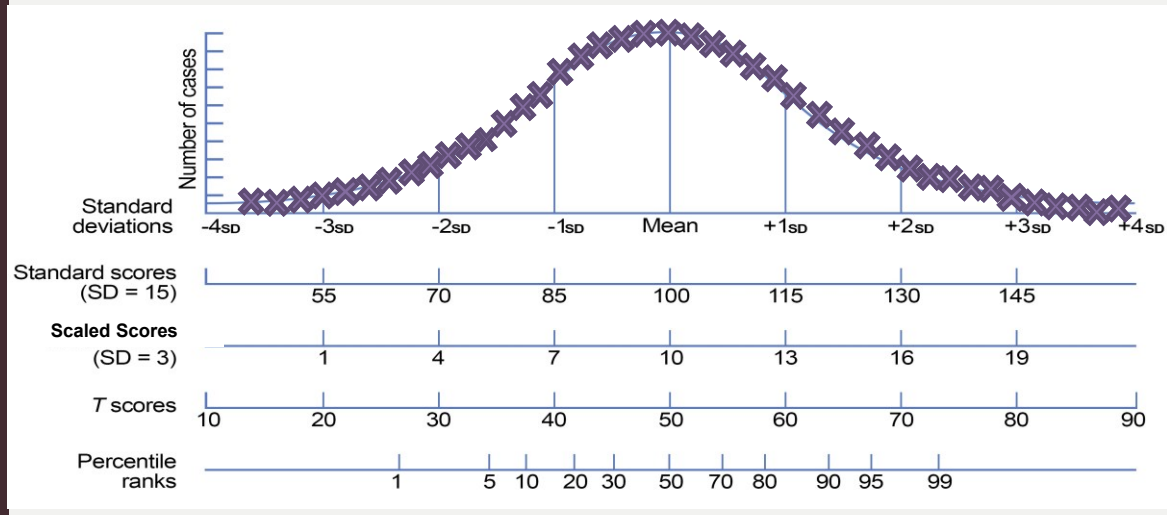
3. We follow the Goldilocks Principle: Subtests should not be too short or too long.

When subtests are too long, our students become understandably annoyed.

When subtests are too short, we do not get a good sample of skills and funny things start to happen with scores.

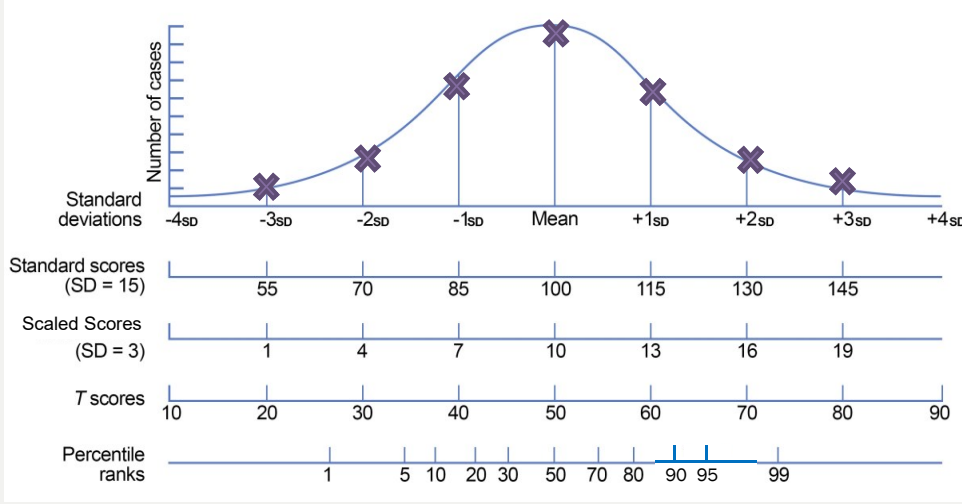
24

Item Gradients: Increments Between Scores



25

What happens when there are too few items in a subtest.



# Correct	Scaled Score
1	1
2	4
3	7
4	10
5	13
6	16
7	19

26

Phonological Awareness Test, Second Edition (NU norms) by Robertson & Salter (2018)

- The PAT-2:NU is a standardized assent of phonological awareness, phoneme-grapheme correspondence, and phonemic decoding skills.
- **The PA Index** (ages 5 thru 9).
  - Rhyming: Discrimination and Production
  - Segmentation: Sentences, Syllables, and Phonemes
  - Isolation: Initial, Final, and Medial
  - Deletion: Compound Words, Syllables, and Phonemes
  - Substitution with Manipulatives
  - Blending: Syllables and Phonemes
- **The Phoneme-Grapheme Index** (ages 6 thru 9).
  - Phoneme-Grapheme Correspondence: consonants, vowels, consonant blends, consonant digraphs, r-controlled vowels, vowel diagraphs, and diphthongs
  - Phonemic Decoding: nonsense words with VC, CVC, CCVC/CVCC, VV, VR, VCe, and diphthongs

27

Phonological Awareness Test, Second Edition (NU) Ages 5-0 thru 5-2

The Low End: The scaled scores and percentile ranks generated by raw scores for Ages 5-0 thru 5-2. Scaled Score Mean = 10, SD = ± 3, range 1 thru 19							
Percentile Rank	Rhyming	Segmentation	Isolation	Deletion	Substitution	Blending	**Scaled Score**
01	0						1
01							2
01			0			0	3
02		0					4
05							5
09				0			6
16							7
25		Sentences	initial	compound words	0	syllables	8

28

Phonological Awareness Test, Second Edition (NU)
Ages 9-6 thru 9-11

The High End: The scaled scores and percentile ranks generated by raw scores for ages 9-6 thru 9-11.
Scaled Score Mean = 10, SD = ± 3, range 1 thru 19

Percentile Rank	Rhyming	Segmentation	Isolation	Deletion	Substitution	Blending	Scaled Score
63	Max				Max	Max	11
75				Max			12
84			Max				13
91		Max					14
95th thru 99th							15-20

29

Phonological Awareness Testing Battles:

Evaluator A	Task	Scaled Score
WJ-IV Incomplete Words	Listening to a word with one or more phonemes missing and identifying the word.	11
WJ-IV Sound Blending	Listening to taped sounds & blending them into words.	10
TAPS-3 Word Discrimination	"Are these words the same words or different words?"	10
Evaluator B	**Task**	**Scaled Score**
CTOPP2 Blending	Listening to taped sounds & blending them into words.	10
CTOPP2 Elision	Saying a word without a part (word, syllable, phoneme)	5
Lindamood AC-3 Total Score	Tracking sound changes with colored blocks	5

30

The World According to Yopp (1988): Phonological and Phonemic Awareness Tasks in Terms of 3 Factors

Simple PA	Complex PA	Third Factor
Isolating initial & final sounds Blending sounds (5 sound sequences) Segmenting 4- & 5-phoneme words	Segmenting sounds in clusters Deletion Reversals Substitutions Pig Latin	Identifying words in compound words Identifying syllables in words Rhyming Recognition Rhyming Production

31

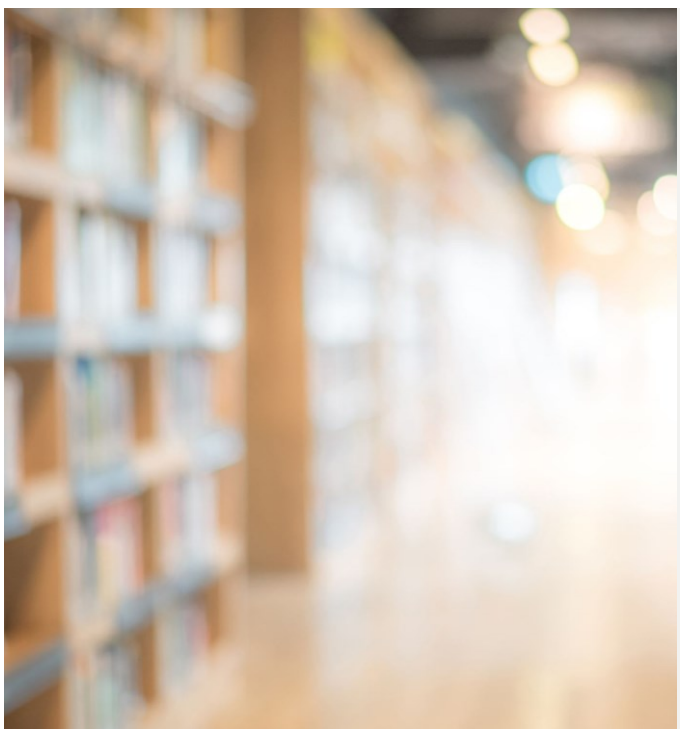
Phonological Awareness Testing Battles:

Evaluator A	Task	Scaled Score	Yopp's Factor
WJ-IV Incomplete Words	Listening to a word with one or more phonemes missing and identifying the word.	11	3rd
TAPS-3 Word Discrimination	"Are these words the same words or different words?"	10	3rd
WJ-IV Sound Blending	Listening to taped sounds & blending them into words.	10	Simple
Evaluator B	Task	Scaled Score	Yopp's Factor
CTOPP2 Blending	Listening to taped sounds & blending them into words.	10	Simple
CTOPP2 Elision	Saying a word w/out a part (word, syllable, phoneme)	5	Complex
Lindamood AC-3 Total Score	Tracking sound changes with colored blocks	5	Complex

32

Tests Measuring Aspects of Phonological Awareness			
	WIAT-4 (PK thru 12+)	KTEA-3 (PK thru 12+)	WJ-IV Oral Language
Phonological Awareness	Timed	Untimed	Untimed
Blending	No	9 items	33 items
Rhyming (Recog/Production)	No	8 items	24 items (Sound Awareness)
Sound Matching	No	5 items	No
Segmenting	No	15 items	37 items
Elision (deletion)	18 items	10 items	20 items (Sound Awareness)
Substitution	12 items	No	No
Reversal	8 items	No	No
Rapid Naming	No	Objects and Letters	Pictures
Simple		Complex		3rd Factor	

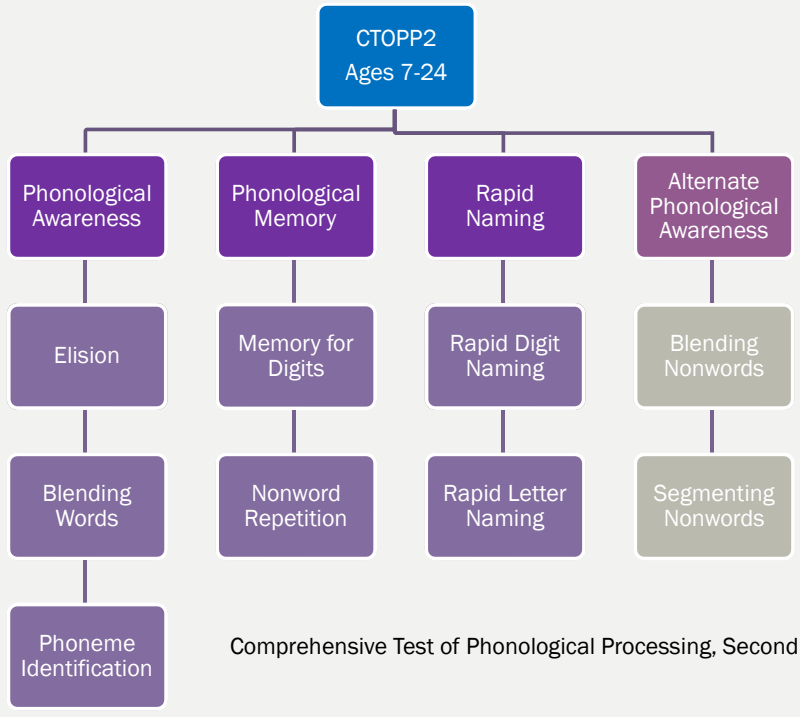
33



Kilpatrick and Phonemic Awareness

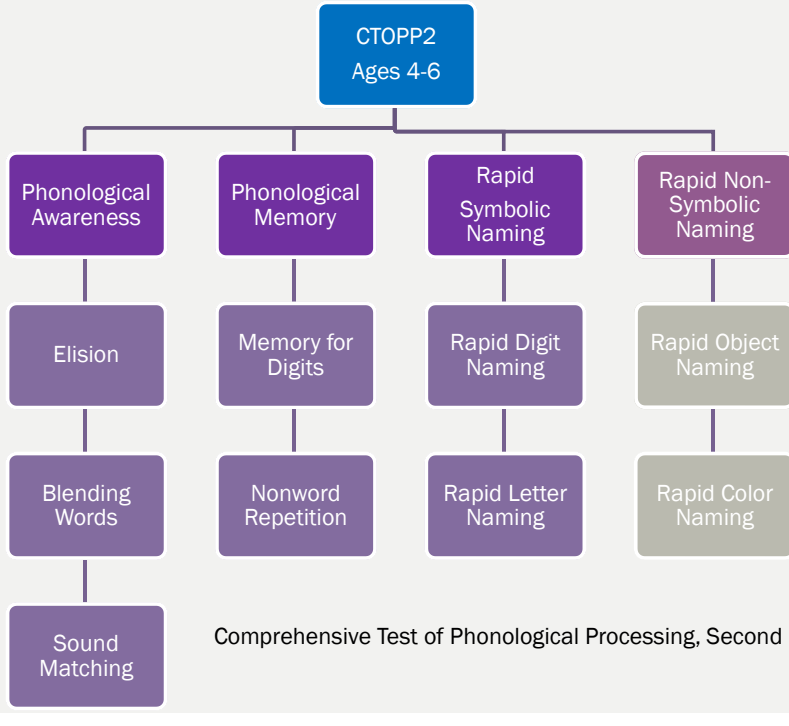
- Segmentation is used in assessment and instruction.
- Segmentation has a weaker correlation (relationship) with reading than tasks requiring students to manipulate speech sounds.
- Segmentation is necessary, but not sufficient.
- It is all about manipulation.
  - *See Equipped for Reading Success published in 2016.*

34



Comprehensive Test of Phonological Processing, Second Edition, Pearson, 2013

35



Letter & number symbols are predictive of academic difficulty.

RNSN used for children who do not yet know their letters & numbers.

Comprehensive Test of Phonological Processing, Second Edition, Pearson, 2013

36

A child who is asleep during the CTOPP2 PA subtests earns the following scaled scores:

Scaled Scores Earned While Asleep				
Ages	Elision	Blending Words	Sound Matching	Phonological Awareness Composite
4-0 thru 4-3	8	7	7	7
4-4 thru 4-7	7	6	6	6
4-8 thru 4-11	5	4	4	1
5-0 thru 5-5	5	4	4	1
5-6 thru 5-11	2	1	1	1
6-0 thru 6-5	1	1	1	1
Scaled Score M = 10 SD = ± 3, range 1 to 19				

37

Screeners

- A screening is a brief evaluation to identify the risk for performing below a certain threshold.
- They should be efficient and inexpensive.
- Developing an effective screener is tricky because there is an inherent tradeoff between correct and incorrect classifications from the screener. These decisions are value judgments based on the costs of false positives and negatives.
  - *A **false negative test** for a lethal, but curable, disease is disastrous.*
  - *A **false positive test** for a terrifying disease that requires a painful and debilitating treatment is not good either.*

	Probably Dyslexic	Probably Not Dyslexic
Positive Test Result	True Positive	False Positive
Negative Test Result	False Negative	True Negative
	Sensitivity	Specificity

38

Screeners for Dyslexia

Screeners can vary in their quality.

If the test uses a cut off score that is too high, we will identify more examinees as having dyslexia, and those numbers will include more falsely identified students (false positive). These students run the risk of being stigmatized. It can also be drain on resources that we then dedicate to assisting these students.

If the cut off score is too low, we will identify more examinees as not having dyslexia, and that will include more falsely non-identified students who actually are at significant risk (false negative). It could mean a lack of access to explicit reading instruction.

39

Dyslexia Indexes: (Screenings)

KTEA-3	**WIAT-4**
AUC (Area under the Curve) Combined Sensitivity and Specificity)	Values greater or equal to .90 are excellent. Values greater or equal to .80 are good.
Grades K – 1 (AUC - .90)	**Grades PK thru 3 (AUC = .95)**
Phonological Processing	Phonemic Proficiency
Letter & Word Recognition	Word Reading
Letter Naming Facility (RAN)	No RAN
Grades 2 thru 12+ (AUC = .89)	**Grades 4 thru 12+ (AUC = .92)**
Nonsense Word Decoding	Pseudoword Decoding
Spelling	Word Reading
Word Recognition Fluency (list format)	Orthographic Fluency (Word Recognition Fluency in disguise)

40

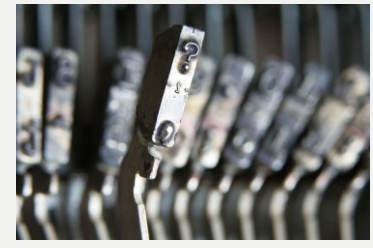
Norm-Referenced Tests: A Question of Bottom

	WJ IV Letter/ Word Recognition 2- 90+	KTEA-3 Letter & Word Identification PK-12+	WIAT-4 Word Reading PK- 12+
Total Number of items	78	100	110 total; 2 parts
Sound-Symbol Correspondence	10 upper/lower case	21 upper/lower case	35 lower case, blends, digraphs
CVC Words	6	5	4

41

41

Orthographic Processing



- The ability to visualize how language is represented on paper in the mind's eye: spelling, contractions, punctuation, capitalization, mathematical notation, and numbers.
- A sense of what is permissible and what is not.
- Distinct from spelling in that there is no motor component.
- Is not innate. Develops as the result of reading experience on a foundation of phonological awareness.

42

Orthographic Processing

- Orthographic weaknesses are presumed in those who have a reduced sight vocabulary, a slow reading rate, and spelling errors that are phonologically correct but do not follow the rules for English.
- Not a subtype of dyslexia. It is the result of insufficient exposure to written language, i.e., limitations in the environment, reading experience, and print exposure.

43

Orthographic Processing:

KTEA-3 OP Composite	WIAT-4 OP Composite
Letter Naming Facility (RAN)	No RAN
Word Recognition Fluency	Orthographic Fluency (Word Recognition Fluency in disguise)
Spelling	Orthographic Choice

Orthographic Choice: Only available on Q interactive. Designed to measure quality of the "orthographic lexicon." Examinees view three choices of letter strings and then touch the one that is spelled correctly. Regular and irregular words. Untimed.

According to the manual, weaknesses in this area may reflect lack of print exposure or a weakness in orthographic learning….

44

Automaticity: Word Lists

WIAT-4 Orthographic Fluency

- Grades 1-2: Set A 20 seconds for each of two trials
- Grades 3 – 12+: Set B 30 seconds for each of two trials
- Prompt when there is no response for 5 seconds (“Try the next one.”)

KTEA-3 Word Recog. Fluency

- Grades 1 & 2: Set A 15 seconds for each of two trials
- Grades 3+: Set B 15 seconds for each of two trials
- Prompt when there is no response for 5 seconds (“Go on to the next one.”)

Both tests also offer Decoding Fluency (nonsense words) subtests for grades 3 thru 12+.

45

Reading Fluency: An error is an error is an error.

- Automaticity presumes accuracy and a level of skill in which it must be easier to read the word than not.
- We are not capable of making this judgment by ear alone.
- All errors (repetition, self correction, synonyms) are the result of inaccuracies in decoding.
- We want to use tests that are sensitive to all errors and not just those that affect meaning.

46

Fluency Tests: Recognition of Deviations from Text

	WIAT-4 Errors	GORT-5 Errors	DIBELS-8 Errors
Repetitions	NO	YES	NO
Self-Corrections	NO	YES	NO (within 3 seconds)
Skipped Lines	NO	YES	YES
Contractions	NO	YES	YES
Insertions	YES	YES	NO
Omissions	YES	YES	YES
Substitutions	YES	YES	YES

Note: The KTEA-3 does not offer oral reading fluency with passages.

47

Wechsler Individual Achievement Test, Fourth Edition: Oral Reading Fluency

- Grades 1 through 12+
- Measures oral reading fluency in narrative and expository texts.
- 2 passages per grade level (grades 7/8, 9-12)
- Comprehension questions are designed to ensure focus on reading for meaning.
- Does not count repetitions, self corrections, skipped lines and contractions as errors.
- Vehicle for dropping back to lower levels (maximum of 3 drop backs)
- Special Warning: *"Estimation of the examinee's reading ability may be less precise on item sets that are far from the grade-appropriate item set.* Use clinical judgement to determine which item set offers a better estimate of the examinee's performance." (Manual, page 144).

48

What about Measures of Silent Reading Fluency?

- WJ-IV: Marking sentences as true/false
  - *Grass is orange.* *YES* *NO*
  - *Soda is dry.* *YES* *NO*
- KTEA-3: Answering YES/NO questions
  - *Do people walk on water?* *YES* *NO*
- Guessable, concrete, and low readability
- No diagnostic potential.

49

Slasher Tests

Itistimeforallgoodchildrentogotobed.

- Evidence that the techniques used in an evaluation are not necessarily like good teaching.
- Highly Efficient
- Can be administered in groups
- Not for children with graphomotor challenges.
- *Test of Silent Word Reading Fluency, Second Edition* (Mather, N., Hammill, D.D., Allen, E.A., Roberts, R., 2014)
- Test of Silent Contextual Reading Fluency, Second Edition (Hammill, D.D., Wiederholt, J.L., Allen, E.A., 2014).

50

Slasher Tests

~~Itistimeforallgoodchildrentogotobed.~~

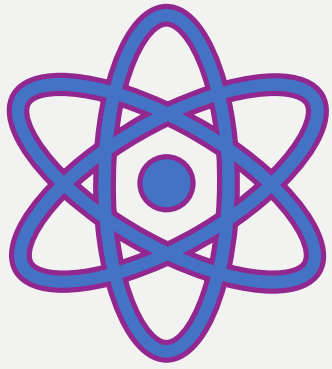
- Evidence that the techniques used in an evaluation are not necessarily like good teaching.

- Highly Efficient
- Can be administered in groups
- Not for children with graphomotor challenges

- *Test of Silent Word Reading Fluency, Second Edition* (Mather, N., Hammill, D.D., Allen, E.A., Roberts, R., 2014)
- Test of Silent Contextual Reading Fluency, Second Edition (Hammill, D.D., Wiederholt, J.L., Allen, E.A., 2014).

51

51



Reading Comprehension: The Questions We Use

In quantum physics, it is said that the act of looking at an object changes the object.

In reading assessment, it is said that the act of asking a question changes how a child thinks about a text…

Questions are the lens.

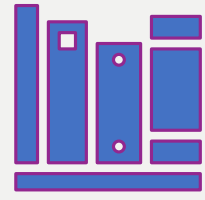
52

Q Type	Example	Skills Required/Demonstrated
Cloze Procedure	I gave the dog a _____.	Sentence level. Expressive language/word retrieval.
Mazes	John drank his glass of (sneak, gun, milk, smoke).	Sentence level. Adequate working memory.
True/False	The milk is wet. YES NO	Concrete at best.
Multiple-Choice	Why did Masha go to the store? A. To buy milk. B. To read a book. C. To play soccer. D. To see her friend.	Adequate working memory. No expressive language skill.
Open Ended	Why is it important to have breakfast?	Expressive language. Window into how a student thinks, as well as language usage and organizational skill.

53

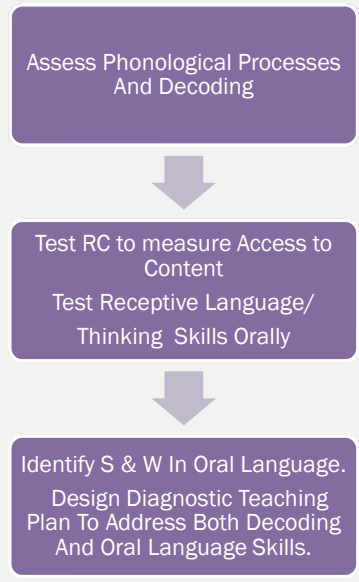
Poor Comprehension: Decoding, Language, or Both?

Sasha demonstrated Below Average skill on the Anybody-Can-Do-It Reading Comprehension test.



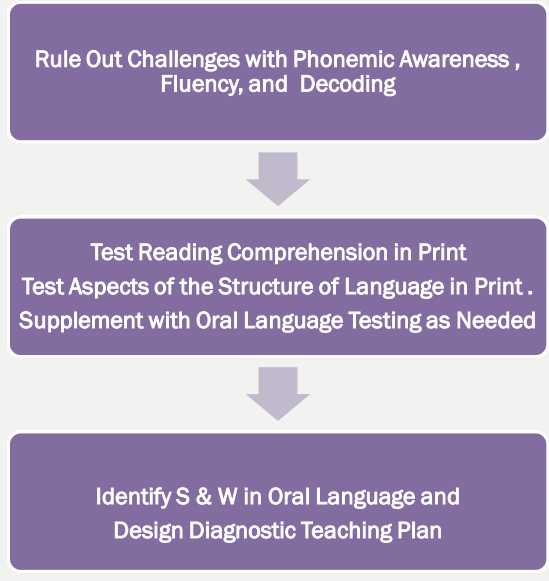
54

Testing Comprehension Skills in Poor Decoders



55

Testing Comprehension in Presumed Good Decoders



56

Poor comprehenders may all look alike on a reading comprehension test but poor comprehension due to…

Domain	Requires Instruction in:
Poor Decoding	Phonemic Awareness, Handwriting, Decoding & Spelling. Access to audio texts.
Poor Receptive Language	Structure Of Language: Vocabulary, Syntax, Abstract Language, Verbal Reasoning/Inferential Thinking
Limited Background Knowledge	Vocabulary And World Knowledge
Weak Memory And Organization	Strategies To Increase Recall And Organization

57

In Closing

- A well-designed evaluation presume a deep knowledge of:
  - *the science of the domain,*
  - *appropriate instructional methodologies, as well as*
  - *best practices in assessment, and*
  - *what tests measure.*
- Small differences in test design can have significant implications for how students perform.
- If we can use our tools appropriately, and think beyond test scores, we can strengthen the link between the data we take and effective, evidence-based recommendations.

58