Measurement and Interpretation of Standardized Reading Assessments for Professionals and Parents

PRESENTED BY **TATYANA ELLESEFF MA CCC-SLP**

SMART SPEECH THERAPY LLC

Overview

This webinar will provide a **basic** overview of popular reading tests and discuss their strengths and limitations with respect to their psychometric properties, testing components, as well as subtest interpretation.

Handouts: Can be accessed in the files section of SLPs for Evidence-Based Practice Facebook Group, under the webinar name after the presentation

An excellent resource for an in-depth understanding of psychometric properties of tests is a webinar entitled, "**Understanding Psychometric Properties of Standardized Tests".** It was presented by **Dr. Elena Plante** at the POWER UP Conference which was hosted by the Lavi Institute in conjunction with the SLPs for Evidence Based Practice Group.

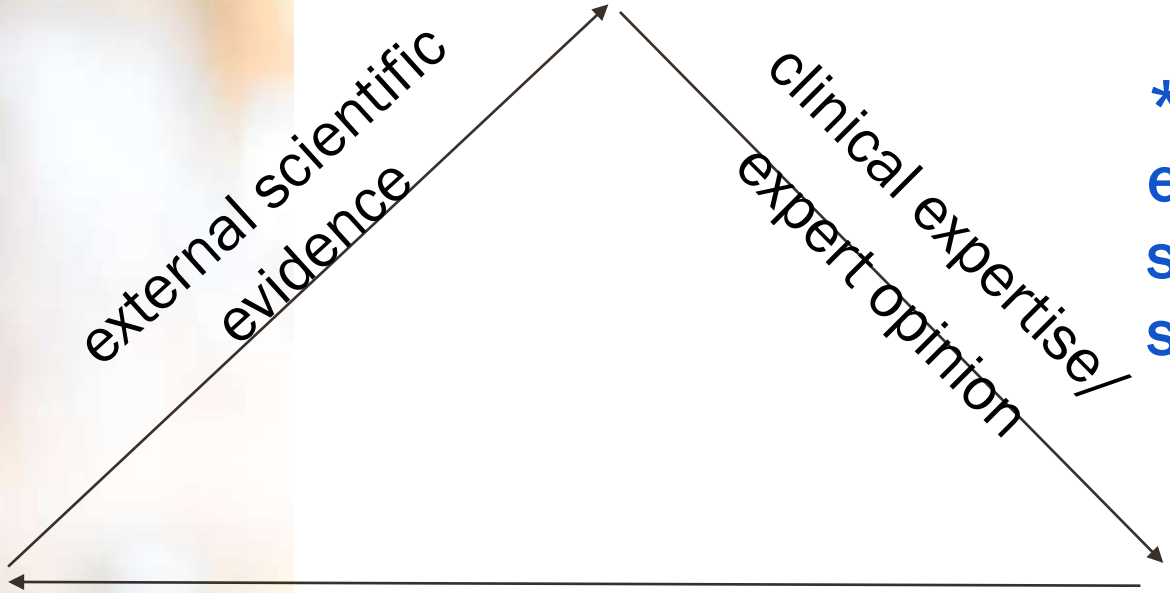
It can be accessed for FREE until 9/14/20: HERE

Learning Objectives

Be	By the end of this webinar participants will be able to:
List	List popular standardized reading tests
Discuss	Discuss discriminant accuracy of select standardized reading tests
Describe	Describe testing components of select, popularly used, standardized tests of reading
Explain	Explain how to interpret standardized testing results in order to understand the client's profile of reading strengths and limitations

Evidence-Based Practice

*Good quality studies based on solid scientific premises



**Practice-based evidence supported by scientific principles; scientifically defensible

***Influenced by clinician recommendations

How to evaluate standardized tests?

Standardization/Normative Sample - # participants arranged by age, sex, ethnicity, geographic region, and parent education level.

- **Were students with disability included in the sample? If so, what percentage?**

Reliability -the degree to which an assessment tool produces stable and consistent results

- **Test-Retest** –administration of the same test twice over a period of time (e.g., 3 weeks) to a group of individuals to see score stability (McCauley & Swisher, 1984)
- **Inter-rater** –scores remain stable if different examiners administer the test (McCauley & Swisher, 1984)
- **Inter-item** - assesses whether parts of an assessment are in fact measuring something similar to what the whole assessment claims to measure (Paul, 2007)

Validity -how well a test measures what it is purported to measure

- **Content** -how representative the test items are of the content that is being assessed (Paul, 2007). Determined by literature review, expert feedback, polls, studies, etc.
- **Construct** -assesses the extent to which a test can be used for as a specific purpose, such as to identify children with a reading disorder
- **Concurrent** - the extent to which a test agrees with other valid tests of the same measure (Paul, 2007)

Standard Error of Measurement (SEM) -"the degree of confidence that the child's 'true' score on a test is represented by the actual score the child received." (Betz, Eickhoff, and Sullivan, 2013, p.135) Provides an estimate of the amount of error in a student's observed test scores

Bias –linguistic, cultural, past experience, socio-economic, etc.

Discriminant Accuracy

Sensitivity and Specificity (Dollaghan, 2007)

- **Sensitivity** – does the assessment accurately identify those students who truly have a language/reading disorder as having a reading disorder
- **Specificity** - does the assessment accurately identify those students who truly do not have any disorders as typical
- Sensitivity and specificity determine the test's degree of **discriminant accuracy**, or the ability to distinguish the presence of a disorder

Vance and Plante (1994) established a criteria for accurate identification of a disorder (discriminant accuracy)

- 90% should be considered good discriminant accuracy
- 80% to 89% should be considered fair
- Below 80%, misidentifications occur at unacceptably high rates" and lead to "serious social consequences" of misidentified children. (p. 21)"

Most important information about the assessment

- If the test has low sensitivity and specificity or if that information is missing; NONE of the other psychometric properties matter

Cut scores

Numerical boundary between what is considered typical and disordered

- *The formula requires the mean and standard deviation of both a clinical and non-clinical sample, and estimates the score at which a subject has a greater probability of belonging to a clinical sample rather than a non-clinical sample.

Test Specific –vary from test to test

Age specific –differ depending on the child's age

Problem: Often applied arbitrarily without reference to how children actually score on the tests selected for use (Spaulding, Plante, & Farinella, 2006)

On the inclusion of students with disabilities in the normative samples (Pena & Plante, 2020 Facebook Group Discussion)

Test developers tend to use the same process as they do for psychological and educational tests. Namely to rank people to represent the full population. For the purpose of ranking, disordered children are used in the sample because it widens the normative range, allowing for more fine-grained divisions and better rank estimates of students who fall -1 SD. However, such tests are not meant for diagnostic purposes, or the determination if a child has a disorder.

Myth: **If a child with a disabling condition is represented in the normative sample than the test is appropriate for usage with that population** (e.g., ADHD, ASD, DLD, etc.)

Reality: For diagnostic purposes there should be no students with disorders included in the normative sample, since our goal is to diagnose impairment for intervention purposes.

Compromise: During the test development stage it is important to identify items that TD students pass and impaired kids fail for diagnostic accuracy purposes. But disordered students should not be included in the standardization norms because it lowers the mean, increases SD, thereby shifts the cut scores, which results in less likely identification of impaired students ("normalizes the disorder"). The overlap between disordered and typical becomes too great and its much harder to reliably identify those with an impairment.

Assessment Areas

Phonological and Phonemic Awareness Skills

- Phonological awareness assessment/intervention has predictive power until 2nd grade. After that it does not add information to the prediction of 4th-grade reading abilities (Hogan, Catts, & Little, 2005) unless the student continues to present with significant reading challenges as evident via sound blending deficits (Kilpatrick, 2012)

Orthographic Mapping Abilities

- Formation of letter-sound connections to bond the spellings, pronunciations, and meanings of specific words in memory
- Explains how children learn to read words by sight, to spell words from memory, and to acquire vocabulary words from print
  - Enabled by phonemic awareness and grapheme-phoneme knowledge (Ehri, 2014)

Rapid Naming Abilities

Rapid automatized naming (RAN) and not phonological awareness has been found to be a consistent predictor of reading fluency in all orthographies (Landerl, et al, 2019).

- Poor rapid automatized naming abilities (on alphanumeric and nonalphanumeric tasks) have been found to be a long-term and universal symptom of reading deficits (Araújo & Faísca, 2019)

Reading Fluency

Reading Comprehension

Language Related Assessment Areas

Literate Vocabulary Knowledge (Nippold, 2018)

- Difficult words that occur in academic contexts

Semantic Awareness (Taylor, Duff, Woollams, Monaghan, & Ricketts, 2015)

- Semantic processes are associated with word reading skills, namely children read words better when they know their meanings

Morphological Awareness (James, Currie, Xiuli Tong, & Cain, 2020)

- Plays a crucial role in supporting higher-level text processing
- It is partly mediated by vocabulary knowledge
- Becomes an increasingly important predictor of reading comprehension between 6 and 11 years
- Makes a unique contribution to reading comprehension ability beyond oral vocabulary and word reading skill

Assessment Overview

* Test of Integrated Language and Literacy (TILLS)

**Woodcock-Johnson IV Tests of Achievement* (WJ IV-ACH)

* *Woodcock-Johnson IV Tests of Oral Language* (WJ IV-OL)

* Kaufman Test of Educational Achievement Third Edition (KTEA-3)

* Wechsler Individual Achievement Test Fourth Edition (WIAT-4)

*Feifer Assessment of Reading (FAR)

Phonological Awareness Test-2: Normative Update (PAT-2: NU)

Comprehensive Test of Phonological Processing-2 (CTOPP-2)

Rapid Automatized Naming and Rapid Alternating Stimulus Test (RAN/RAS)

The Test of Silent Word Reading Fluency (TOSWFR-2)

Test of Silent Contextual Reading Fluency (TOSCRF-2)

Gray Oral Reading Tests- Fifth Edition (GORT-5)

Test of Reading Comprehension – Fourth Edition (TORC-4)

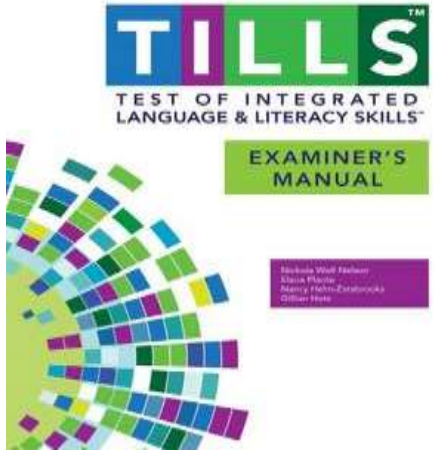
* Denotes a comprehensive test

Comprehensive Assessment: Language and Literacy

The Test of Integrated Language & Literacy Skills (TILLS) (2016) is an assessment of oral and written language abilities in students **6–18** years of age composed of 15 subtests.

Assesses literacy skills such as **reading fluency, reading comprehension, phonological awareness, spelling, as well as writing** in monolingual as well as simultaneously bilingual school age children.

Language Modality	Language Dimension	
	Sound/Word Level	Sentence/Discourse Level
Listening	1. Vocabulary Awareness 2. Phonemic Awareness	6. Listening Comprehension 8. Following Directions
Speaking	4. Nonword Repetition	3. Story Retelling 13. Social Communication
Reading	10. Nonword Reading 11. Reading Fluency	7. Reading Comprehension
Writing	5. Nonword Spelling 12a. Written Expression – Word Score	12b. Written Expression – Discourse Score 12c. Written Expression – Sentence Combining Score
Memory	14. Digit Span Forward 15. Digit Span Backward	9. Delayed Story Retelling



TILLS (cont.)

- Standardized to identify language and literacy disorders
- Excellent psychometric properties

Table 3.4. Sensitivity and specificity levels by age for all ages tested by the TILLS

Age groups	Sensitivity	Specificity
6-year-olds	84%	82%
7-year-olds	84%	86%
8-year-olds	97%	100%
9-year-olds	83%	81%
10-year-olds	81%	81%
11-year-olds	86%	82%
12-year-olds	83%	100%
13-year-olds	84%	86%
14- to 18-year-olds	87%	87%

TILLS (cont.)

Subtests Sensitivity to Language and Literacy Impairments Based on Age Groups

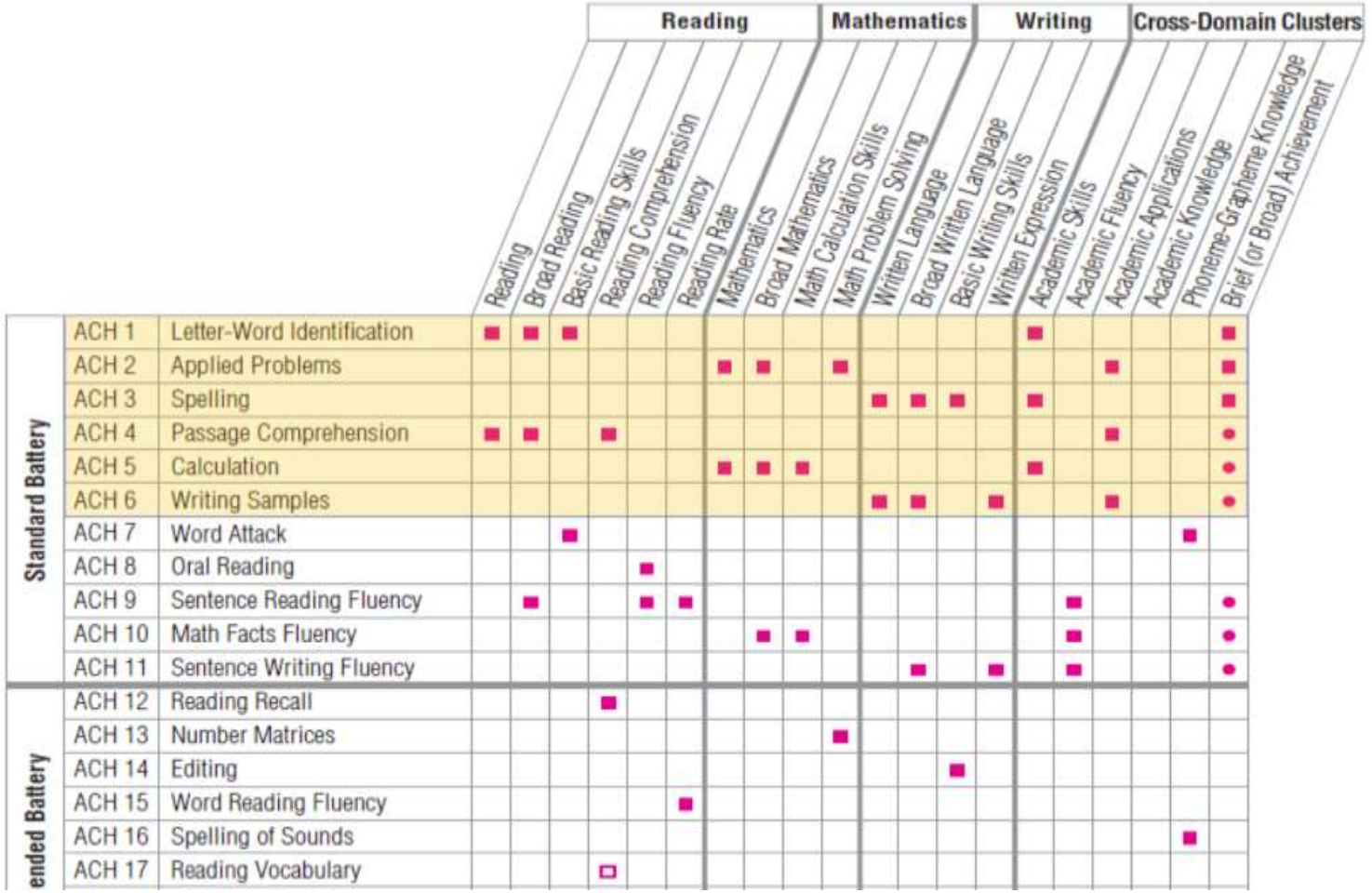
Table 2.2. TILLS subtests that support diagnosis of language and literacy disorders at different ages

Age range (years)	Identification Core[a]	Sensitivity	Specificity	Cut score[b]
6;0–7;11	1. Vocabulary Awareness (VA) 2. Phonemic Awareness (PA) 4. Nonword Repetition (NWRep)	84	84	24
8;0–11;11	1. Vocabulary Awareness (VA) 5. Nonword Spelling (NWSpell) 10. Nonword Reading (NWRead) 12. Written Expression–Discourse Score (WE-Disc)	88	85	34
12;0–18;11	2. Phonemic Awareness (PA) 5. Nonword Spelling (NWSpell) 7. Reading Comprehension (RC)[c] 11. Reading Fluency (RF) 12. Written Expression–Word Score (WE-Word)	86	90	42

Comprehensive Assessments: Educational

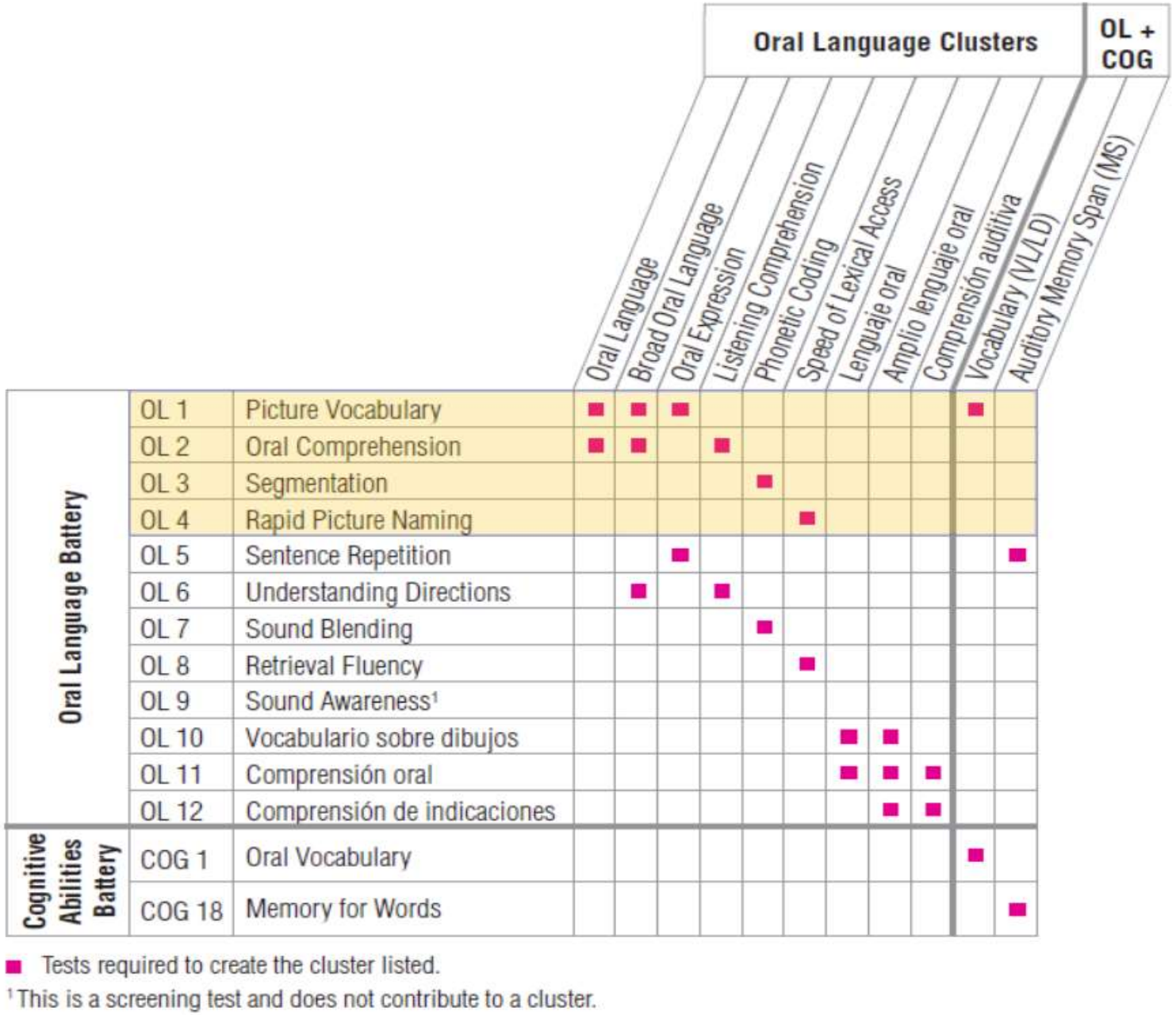
- ***Woodcock-Johnson® IV Tests of Achievement*** **(WJ IV™ ACH)** for screening, diagnosing, and monitoring progress in reading, writing (and math) (2014) (Ages: 2-90+)
  - 20 tests for measuring four broad academic domains: reading, written language, mathematics, and academic knowledge
  - Detailed subtest descriptions of both standard and extended batteries of the WJ IV™ ACH can be found HERE and HERE to understand what it does and does not test
- ***Woodcock-Johnson® IV Tests of Oral Language*** **(WJ IV™ OL)**)(2014) (Ages: 2-90+) can be used to determine and describe an individual's strengths and weaknesses with regard to expressive language
  - 12-test battery consists of nine English tests and three Spanish tests
  - Detailed subtest descriptions of both standard and extended batteries of the WJ IV™ ACH can be found HERE to understand what it does and does not test

WJ-IV:ACH



			Reading						Mathematics			Writing				Cross-Domain Clusters				
			Reading	Broad Reading	Basic Reading Skills	Reading Comprehension	Reading Fluency	Reading Rate	Mathematics	Broad Mathematics	Math Calculation Skills	Math Problem Solving	Written Language	Broad Written Language	Basic Writing Skills	Written Expression	Academic Skills	Academic Fluency	Academic Applications	Academic Knowledge	Phoneme-Grapheme Knowledge	Brief (or Broad) Achievement
Standard Battery	ACH 1	Letter-Word Identification	■	■	■											■						■
	ACH 2	Applied Problems						■	■		■							■				■
	ACH 3	Spelling										■	■	■		■						■
	ACH 4	Passage Comprehension	■	■		■												■				●
	ACH 5	Calculation						■	■	■						■						●
	ACH 6	Writing Samples										■	■		■			■				●
	ACH 7	Word Attack			■																■	
	ACH 8	Oral Reading					■															
	ACH 9	Sentence Reading Fluency		■			■	■									■					●
	ACH 10	Math Facts Fluency							■	■							■					●
	ACH 11	Sentence Writing Fluency											■		■		■					●
ended Battery	ACH 12	Reading Recall				■																
	ACH 13	Number Matrices									■											
	ACH 14	Editing												■								
	ACH 15	Word Reading Fluency					■															
	ACH 16	Spelling of Sounds																			■	
	ACH 17	Reading Vocabulary				□																

WJ-IV:OL



			Oral Language	Broad Oral Language	Oral Expression	Listening Comprehension	Phonetic Coding	Speed of Lexical Access	Lenguaje oral	Amplio lenguaje oral	Comprensión auditiva	Vocabulary (VL/LD)	Auditory Memory Span (MS)
			Oral Language Clusters										OL + COG
Oral Language Battery	OL 1	Picture Vocabulary	■	■	■							■	
	OL 2	Oral Comprehension	■	■		■							
	OL 3	Segmentation					■						
	OL 4	Rapid Picture Naming						■					
	OL 5	Sentence Repetition			■								■
	OL 6	Understanding Directions		■		■							
	OL 7	Sound Blending					■						
	OL 8	Retrieval Fluency						■					
	OL 9	Sound Awareness[1]											
	OL 10	Vocabulario sobre dibujos							■	■			
	OL 11	Comprensión oral							■	■	■		
	OL 12	Comprensión de indicaciones								■	■		
Cognitive Abilities Battery	COG 1	Oral Vocabulary										■	
	COG 18	Memory for Words											■

■ Tests required to create the cluster listed.

[1] This is a screening test and does not contribute to a cluster.

WJ IV™ ACH and OL

Developed to rank children within the range of the general population

No mention of sensitivity and specificity in the technical manual for either test (https://www.wjscore.com/Files/WJIVTechnicalManual.PDF)

- Discriminant accuracy for the purpose of disorder identification is **unknown**
  - **Do quite well on this test and be reading, writing or oral language impaired**

Comprehensive Assessments: Educational (cont.)

Kaufman Test of Educational Achievement Third Edition (KTEA-3) (2014) is a measure of academic achievement for individuals ages 4:0 through 25:11

2014

Feifer Assessment of Reading (FAR) (2015) a comprehensive assessment of reading and related processes, helps determine an individual's specific subtype of reading impairment and inform intervention planning ages 4:0 through 21:11

2015

Wechsler Individual Achievement Test - Fourth Edition (WIAT-4) (Fall 2020) achievement test for use in a variety of clinical, education, and research settings for individuals ages 4:0 through 50:11

2020

KTEA-3

Subtests

- Phonological Processing (PP)
- Math Concepts & Applications (MCA)
- Letter & Word Recognition (LWR)
- Math Computation (MC)
- Nonsense Word Decoding (NWD)
- Writing Fluency (WF)
- Silent Reading Fluency (SRF)
- Math Fluency (MF)
- Reading Comprehension (RC)
- Written Expression (WE)
- Associational Fluency (AF)
- Spelling (SP)
- Object Naming Facility (ONF)
- Reading Vocabulary (RV)
- Letter Naming Facility (LNF)
- Listening Comprehension (LC)
- Word Recognition Fluency (WRF)
- Oral Expression (OE)
- Decoding Fluency (DF)

Sample Reports

- Parent (Includes subtest descriptions)
- Scores (Includes summary profiles and explanation of scores)

KTEA-3

Developed to rank children within the range of the general population

No mention of sensitivity and specificity in the technical manual for this test

- Discriminant accuracy for the purpose of disorder identification is **unknown**
  - **Clinical Observations: More closely aligned scores wise to the TILLS as compared to several other achievement tests**

KTEA-3 vs. TILLS (Age 10)

ACQUIRED KNOWLEDGE/ACADEMIC FUNCTIONING

Acquired Knowledge-Reading Achievement			
Basic Reading Skills			
KTEA-3	STDS	%ile	Classification
Letter & Word Recognition	88	21	Slightly Below Expected
Reading Comprehension			
KTEA-3	STDS	%ile	Classification
Reading Comprehension	77	6	Below Expected

Acquired Knowledge-Written Language Achievement			
KTEA-3	STDS	%ile	Classification
Written Expression	77	6	Below Expected
Spelling	67	1	Well Below Expected

	Subtest Scores and TILLS Total		
Subtest	Raw Score	Standard Score and TILLS Total	Percentile Rank
1 VA	17	3	4
2 PA	5	0	1
3 SR	22	9	31
4 NWRep	20	9	14
5 NWSpell	0	2	0
6 LC	7	0	0
7 RC	8	3	2
8 FD	6	4	2
9 DSR	13	6	8
10 NWRead	4	1	0
11 RF	79	3	3
12a WE-Disc	95	11	45
12b WE-Sent	.95	5	0
12c WE-Word	85	0	0

WIAT-4* (New)

Sample Reports

- Parent (Includes subtest descriptions)
- Scores (Includes summary profiles and explanation of scores)

Core Academic Composites	Grade Range	Subtests
Reading	K–12+	Word Reading
		Reading Comprehension
Written Expression	K–12+	Spelling
		Alphabet Writing Fluency
		Sentence Composition
		Essay Composition
Mathematics	K–12+	Numerical Operations
		Math Problem Solving

Processing Composites	Grade Range	Subtests
NEW! Orthographic Processing	1–12+	Spelling
		NEW! Orthographic Fluency
		NEW! Orthographic Choice (Q-interactive® only; provides an Orthographic Processing Extended composite score)
NEW! Phonological Processing	1–12+	NEW! Phonemic Proficiency
		Pseudoword Decoding

Supplemental Composites	Grade Range	Subtests
Oral Language	PreK–12+	Listening Comprehension
		Oral Expression
NEW! Basic Reading	1–12+	Word Reading
		Pseudoword Decoding
		NEW! Phonemic Proficiency
Decoding	1–12+	Word Reading
		Pseudoword Decoding
NEW! Reading Fluency	3–12+	Oral Reading Fluency
		NEW! Orthographic Fluency
		NEW! Decoding Fluency
NEW! Writing Fluency	1–4	Alphabet Writing Fluency
		NEW! Sentence Writing Fluency
Math Fluency	1–12+	Math Fluency: Addition
		Math Fluency: Subtraction
		Math Fluency: Multiplication

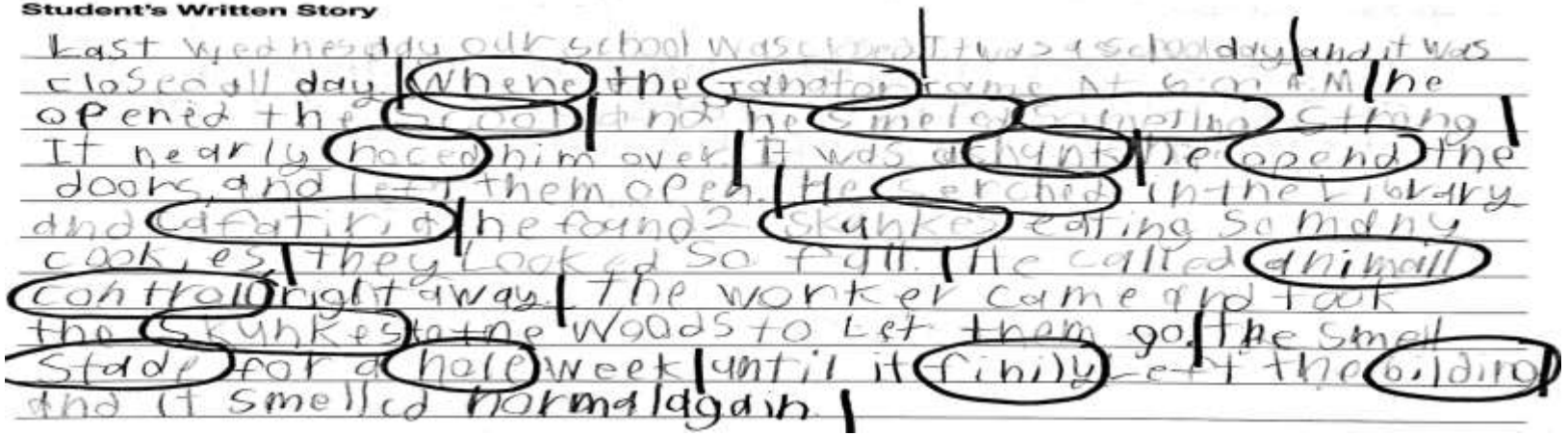
WIAT-3 vs. TILLS (Age:11)

Subtest Score Summary

Subtest	Raw Score	Standard Score	90% Confidence Interval	Percentile Rank	Normal Curve Equiv.	Stanine	Grade Equiv.	Age Equiv.	Growth Score
Word Reading	54	108	104-112	70	61	6	7.2	12:8	592
Pseudoword Decoding	27	96	91-101	39	44	4	4.2	9:0	510
Oral Reading Fluency	138[1]	109	102-116	73	63	6	6.7	12:4	540
Spelling	21	86	81-91	18	30	3	3.7	9:0	516

SUBTEST 12: Written Expression (WE)

Student's Written Story



Subtest	Subtest Scores and TILLS Total		
	Raw Score	Standard Score and TILLS Total	Percentile Rank
1 VA	39	11	53
2 PA	5	0	0
3 SR	27	11	57
4 NWRep	18	6	4
5 NWSpell	7	7	14
6 LC	19	9	36
7 RC	15	8	16
8 FD	10	7	12
9 DSR	29	13	79
10 NWRead	11	5	6
11 RF	122	4	5
12a WE-Disc	94	11	51
12b WE-Sent	1.71	9	32
12c WE-Word	84	0	0
13 SC			
14 DSF	6	7	14
15 DSB	2	6	5

FAR Subtest and Index Score Structure

Phonological Index subtests	Fluency Index subtests	Comprehension Index subtests
Phonemic Awareness	Rapid Automatic Naming	Semantic Concepts
Nonsense Word Decoding	Verbal Fluency	Word Recall
Isolated Word Reading Fluency	Visual Perception	Print Knowledge
Oral Reading Fluency	Irregular Word Reading Fluency	Morphological Processing
Positioning Sounds	Orthographical Processing	Silent Reading Fluency: Comprehension

Note: The **Mixed Index** includes the subtests from the Phonological and Fluency Indexes; the **Total Index** includes all subtests.

FAR (2015) Ages: 4:00-21:11

Addresses four specific subtypes of dyslexia: dysphonetic dyslexia, surface dyslexia, mixed dyslexia, and reading comprehension deficits.*

Comprises 15 individual subtests measuring various aspects of vocabulary, phonological awareness, decoding skills, rapid automatic naming, orthographical processing, morphological processing, word memory, reading fluency (word and story; silent and oral), and comprehension skills.

Dyslexia Subtypes

Appealing because of a belief that subtypes will guide treatment practices

Poor readers can present with a variety of permutations of reading difficulties

Poor research evidence to support them (Zoubrinetzky, Bielle & Valdois, 2014)

The number of symptoms of dyslexia described in the literature exceeds the number of subtypes, and underlying relations remain unclear (Tamboer, Vorst, & Oort, 2016)

Multiple cognitive deficit model of dyslexia is supported, whereas the existence of subtypes remains unclear (Tamboer, Vorst, & Oort, 2016; Kornilov, & Grigorenko, 2018)

FAR (cont.)

Useful resources

Comprehensive Score Report

Interpretative Report

Sensitivity .67 (unacceptable) and Specificity .98 (excellent) BUT not as relevant as actual sensitivity

FAR Validity

Classification Accuracy of FAR Total Index Score for Identifying Specific Learning Disability

Data from Table 6.20, *Feifer Assessment of Reading Professional Manual*

		Classification Predicted by FAR Total Index Score	
Actual Group	***n***	**Learning Disability**	**No Learning Disability**
Learning Disability	59	40 (67.8%)	19 (32.2%)
No Learning Disability	1078	19 (1.8%)	1,059 (98.2%)

Sensitivity = 0.678 Positive Predictive Power = 0.678 False Positive = 0.322

Specificity = 0.982 Negative Predictive Power = 0.982 False Negative = 0.018

The overall percentage of cases correctly classified = 96.7%. Wilk's lambda = .754, $p \leq .001$.

Table 6.21 of the *FAR Professional Manual* shows the effects on these statistics of using different FAR Total Index Score cut-offs.

FAR vs. TILLS (Age 7)

Subtest	Raw Score	Standard score	Index standard score	Percentile rank
Phonemic Awareness (PA)	89	126		96
Nonsense Word Decoding (NWD)	n/a	--		--
Isolated Word Reading Fluency (ISO)	33	96		39
Oral Reading Fluency (ORF)	0.81	86		18
Positioning Sounds (PS)	29	113		81
Phonological Index (PI)		421	107	68

	Subtest Scores and TILLS Total		
Subtest	Raw Score	Standard Score and TILLS Total	Percentile Rank
1 VA	11	7	13
2 PA	2	4	4
3 SR			
4 NWRep	23	13	79
5 NWSpell	5	10	49
6 LC			
7 RC			
8 FD			
9 DSR			
10 NWRead	2	5	6
11 RF	58	9	11

PAT-2: NU (2018) Ages 5:00-9:11

A standardized assessment of phonological awareness, phoneme-grapheme correspondence, and phonemic decoding skills

- **Rhyming**: Discrimination and Production-identify rhyming pairs and provide a rhyming word
- **Segmentation**: Sentences, Syllables, and Phonemes-divide by words, syllables, and phonemes
- **Isolation**: Initial, Final, and Medial-identity sound position in words
- **Deletion**: Compound Words, Syllables, and Phonemes-manipulate root words, syllables, and phonemes in words
- **Substitution with Manipulatives**: isolate a phoneme in a word, then change in to another phoneme to form a new word
- **Blending**: Syllables and Phonemes blend units of sound to form new words
- ***Phoneme-Grapheme Correspondence**: assesses knowledge of sound/symbol correspondence for consonants, vowels, consonant blends, consonant digraphs, r-controlled vowels, vowel diagraphs, and diphthongs
- ***Phonemic Decoding**: assesses general knowledge of sound/symbol correspondence to blend sounds into nonsense words

PAT-2: NU (cont.)

Sensitivity and Specificity are NOT reported in the manual BUT:

15% of children in the normative sample presented with a disability (Manual, pg. 21)

- 3% language impairment
- 12% special education

Why does this matter?

- According to Peña, Spaulding and Plante (2006),"by including such children [with disabilities] in the normative sample, we may be "shooting ourselves in the foot" in terms of testing for the purpose of identifying disorders." (pg. 248)
  - Adversely impacts discriminant accuracy (differentiation between typical and disordered)
  - Lowers the mean score or essentially normalizes the disorder (e.g., children with mild disabilities will not be flagged)

CTOPP-2 (2013) Ages 4:00-24:11

- Phonological Segmentation
- Blending Words
- Sound Matching
  - Initial, Medial and Final Phoneme Isolation
- **Blending Nonwords**
- **Segmenting Nonwords**
- Memory for Digits
- **Nonword Repetition**
- **Rapid Digit Naming**
- **Rapid Letter Naming**
- **Rapid Color Naming**
- **Rapid Object Naming**

CTOPP-2 Subtests

1. Elision measures the ability to remove phonological segments from spoken words to form other words.
2. Blending Words measures the ability to synthesize sounds to form words.
3. Sound Matching measures the ability to select words with the same initial and final sounds.
4. Phoneme Isolation measures the ability to isolate individual sounds within words.
5. Blending Nonwords measures the ability to synthesize sounds to form nonwords.
6. Segmenting Nonwords measures the ability to segment nonwords into phonemes.
7. Memory for Digits measures the ability to repeat numbers accurately.
8. Nonword Repetition measures the ability to repeat nonwords accurately.
9. Rapid Digit Naming measures the ability to rapidly name numbers.
10. Rapid Letter Naming measures the ability to rapidly name letters.
11. Rapid Color Naming measures the ability to rapidly name colors.
12. Rapid Object Naming measures the ability to rapidly name objects.

CTOPP-2 (cont.)

Sensitivity and Specificity are NOT reported in the manual BUT:

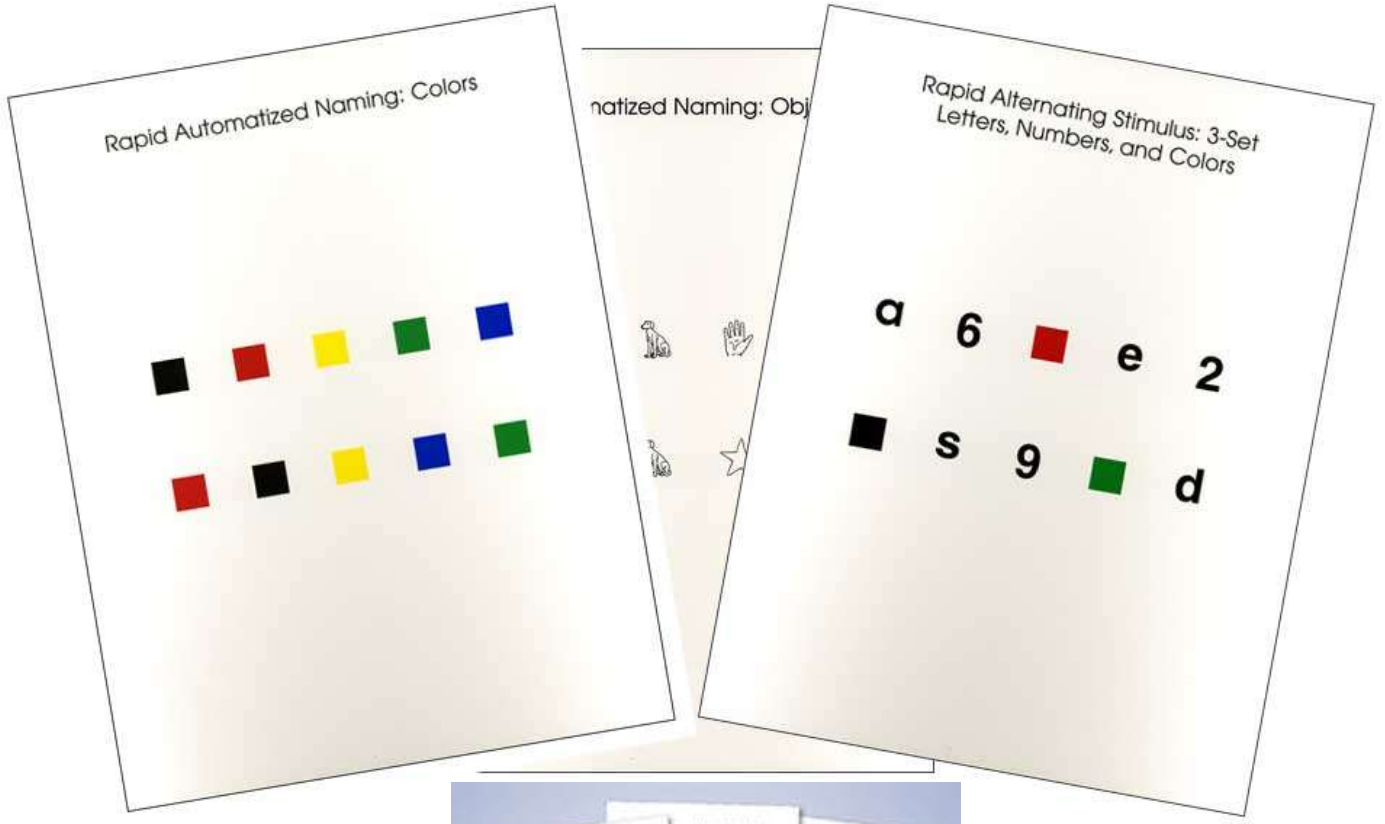
<7% of children in the normative sample presented with a disability (Manual, pg. 44)

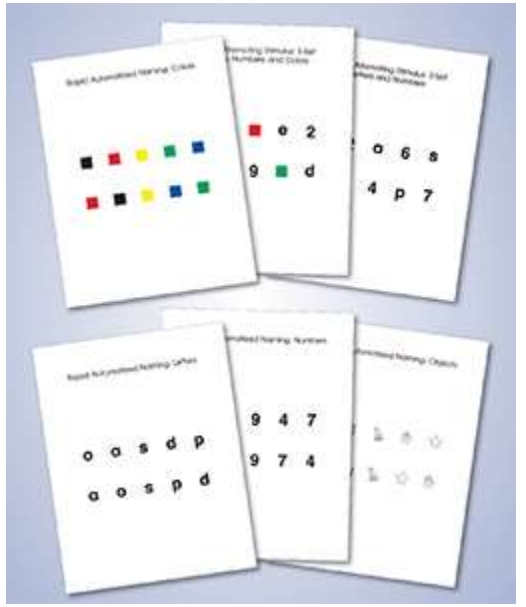
- SLI
- ID (formerly MR)
- HI
- OHI
- ADHD
- Other

RAN/RAS (2005) 5:00-18:11

On all tests the participants are asked to name visual symbols (letters, numbers, objects, and colors) as quickly and accurately as possible (scores are time based).

- The main tests are made up of five high-frequency stimuli that are repeated randomly 10 times in an array of five rows for a total of fifty stimulus items.
- Additionally there are two rapid alternating stimulus tests (2-Set Letters and Numbers; 3-Set Letters, Numbers, and Colors) which are made up of 10 and 15, respectively, high-frequency stimuli that are randomly repeated in an array of five rows for a total of 50 stimulus items.





RAN/RAS (cont.)

Normative data provided by Wolf, Bally, and Morris (1986) were used for the computation of standard scores for latency within each category. Raw scores were used for the number of errors within each category.

Sensitivity and Specificity are NOT reported in the manual

Rapid Naming: CTOPP-2 vs. RAN/RAS

RAN-RAS Tests and CTOPP rapid naming subtests share similarities, but the two measures differ in format, reflecting different theoretical viewpoints… The RAN-RAS tests treat rapid naming as a cognitive ability that includes phonology but also other linguistic and visual processes… The CTOPP was designed on the basis of a model of overall phonological processing that includes phonological awareness, phonological memory, and rapid naming as related subcomponents." (Norton & Wolf, 2012, p. 435)

Controversy exists regarding whether rapid naming should be considered a subskill related to phonological processing or whether RAN is a separate process (**Norton & Wolf, 2012, p. 4357-438**)

1. RAN and phonological processing are not strongly correlated
2. Regression and structural equation models consistently report that RAN and PA account for unique variance in reading ability (e.g., Cutting & Denckla 2001, Katzir et al. 2006)
   i. Different underlying factors support RAN and PA (Powell et al. 2007)
3. Genetic and neuroimaging studies find different biological bases for RAN and PA abilities

TOSWFR-2 (2014) Ages: 6:3-24:11

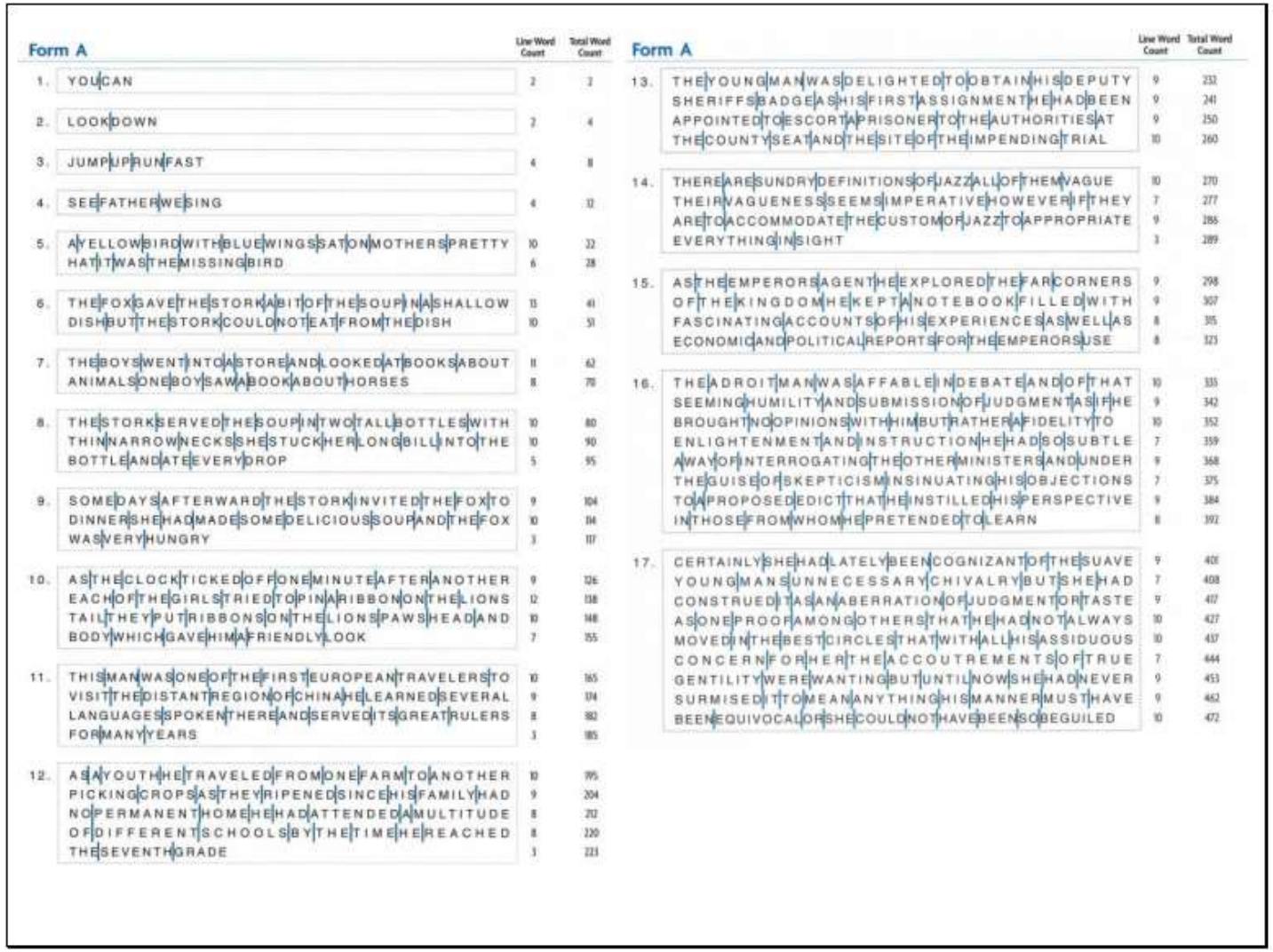
Assesses the students' ability to recognize printed words accurately and efficiently, Students need to identify increasingly difficult words that have no spaces between them by drawing lines between the boundaries of as many words as possible within the time limit of 3 minutes.

Unacceptable Sensitivity .75 and Specificity .75 Access detailed info HERE (pgs. 1 and 6)

TOSCRF-2 (2014) Ages: *7:0 – 24:11*

Companion test to TOSWFR-2
Students read connected text of increasingly difficult graded passages. Measures a student's essential contextual reading abilities (i.e., word identification, word meaning, word building, sentence structure, comprehension, and fluency)

Unacceptable Sensitivity <80
Access info HERE **(pg. 32)**



Form A

		Line Word Count	Total Word Count
1.	YOU CAN	2	2
2.	LOOK DOWN	2	4
3.	JUMP UP RUN FAST	4	8
4.	SEE FATHER WE SING	4	12
5.	A YELLOW BIRD WITH BLUE WINGS SAT ON MOTHERS PRETTY HAT IT WAS THE MISSING BIRD	10 6	22 28
6.	THE FOX GAVE THE STORK A BIT OF THE SOUP IN A SHALLOW DISH BUT THE STORK COULD NOT EAT FROM THE DISH	13 10	41 51
7.	THE BOYS WENT INTO A STORE AND LOOKED AT BOOKS ABOUT ANIMALS ONE BOY SAW A BOOK ABOUT HORSES	11 8	62 70
8.	THE STORK SERVED THE SOUP IN TWO TALL BOTTLES WITH THIN NARROW NECKS SHE STUCK HER LONG BILL INTO THE BOTTLE AND ATE EVERY DROP	10 10 5	80 90 95
9.	SOME DAYS AFTERWARD THE STORK INVITED THE FOX TO DINNER SHE HAD MADE SOME DELICIOUS SOUP AND THE FOX WAS VERY HUNGRY	9 10 3	104 114 117
10.	AS THE CLOCK TICKED OFF ONE MINUTE AFTER ANOTHER EACH OF THE GIRLS TRIED TO PIN A RIBBON ON THE LIONS TAIL THEY PUT RIBBONS ON THE LIONS PAWS HEAD AND BODY WHICH GAVE HIM A FRIENDLY LOOK	9 12 10 7	126 138 148 155
11.	THIS MAN WAS ONE OF THE FIRST EUROPEAN TRAVELERS TO VISIT THE DISTANT REGION OF CHINA HE LEARNED SEVERAL LANGUAGES SPOKEN THERE AND SERVED ITS GREAT RULERS FOR MANY YEARS	10 9 8 3	165 174 182 185
12.	AS A YOUTH HE TRAVELED FROM ONE FARM TO ANOTHER PICKING CROPS AS THEY RIPENED SINCE HIS FAMILY HAD NO PERMANENT HOME HE HAD ATTENDED A MULTITUDE OF DIFFERENT SCHOOLS BY THE TIME HE REACHED THE SEVENTH GRADE	10 9 8 8 3	195 204 212 220 223

Form A

		Line Word Count	Total Word Count
13.	THE YOUNG MAN WAS DELIGHTED TO OBTAIN HIS DEPUTY SHERIFFS BADGE AS HIS FIRST ASSIGNMENT HE HAD BEEN APPOINTED TO ESCORT A PRISONER TO THE AUTHORITIES AT THE COUNTY SEAT AND THE SITE OF THE IMPENDING TRIAL	9 9 9 10	232 241 250 260
14.	THERE ARE SUNDRY DEFINITIONS OF JAZZ ALL OF THEM VAGUE THEIR VAGUENESS SEEMS IMPERATIVE HOWEVER IF THEY ARE TO ACCOMMODATE THE CUSTOM OF JAZZ TO APPROPRIATE EVERYTHING IN SIGHT	10 7 9 3	270 277 286 289
15.	AS THE EMPERORS AGENT HE EXPLORED THE FAR CORNERS OF THE KINGDOM HE KEPT A NOTEBOOK FILLED WITH FASCINATING ACCOUNTS OF HIS EXPERIENCES AS WELL AS ECONOMIC AND POLITICAL REPORTS FOR THE EMPEROR'S USE	9 9 8 8	298 307 315 323
16.	THE ADROIT MAN WAS AFFABLE IN DEBATE AND OF THAT SEEMING HUMILITY AND SUBMISSION OF JUDGMENT AS IF HE BROUGHT NO OPINIONS WITH HIM BUT RATHER A FIDELITY TO ENLIGHTENMENT AND INSTRUCTION HE HAD SO SUBTLE A WAY OF INTERROGATING THE OTHER MINISTERS AND UNDER THE GUISE OF SKEPTICISM INSINUATING HIS OBJECTIONS TO A PROPOSED EDICT THAT HE INSTILLED HIS PERSPECTIVE IN THOSE FROM WHOM HE PRETENDED TO LEARN	10 9 10 7 9 8	335 342 352 359 368 375 384 392
17.	CERTAINLY SHE HAD LATELY BEEN COGNIZANT OF THE SUAVE YOUNG MAN'S UNNECESSARY CHIVALRY BUT SHE HAD CONSTRUED IT AS AN ABERRATION OF JUDGMENT OR TASTE AS ONE PROOF AMONG OTHERS THAT HE HAD NOT ALWAYS MOVED IN THE BEST CIRCLES THAT WITH ALL HIS ASSIDUOUS CONCERN FOR HER THE ACCOUTREMENTS OF TRUE GENTILITY WERE WANTING BUT UNTIL NOW SHE HAD NEVER SURMISED IT TO MEAN ANYTHING HIS MANNER MUST HAVE BEEN EQUIVOCAL OR SHE COULD NOT HAVE BEEN SO BEGUILED	9 7 9 10 7 9 9 10	401 408 417 427 437 444 453 462 472

GORT-5

Rate - the amount of time in seconds taken by a student to read a story aloud

Accuracy - the number of words the student pronounces correctly when reading the passage

Fluency - a combination of the student's Rate and Accuracy Scores

Comprehension - the number of open-ended questions about the stories that the student answers correctly

Oral Reading Index - formed by combining students' Fluency and Comprehension scaled scores

GORT-5 (cont.)

Sensitivity .82 and Specificity .86

Cut score = 90

However, 15% of children in the normative sample presented with a disability (Manual, pg. 37)

- ADHD
- SLD
- DLD
- ID
- DD
- Other

Oral Fluency Norms

COMPILED ORF NORMS

Hasbrouck & Tindal (2017)

From Hasbrouck, J. & Tindal, G. (2017). *An update to compiled ORF norms* (Technical Report No. 1702). Eugene, OR. Behavioral Research and Teaching, University of Oregon.

Grade	Percentile	Fall WCPM*	Winter WCPM*	Spring WCPM*
1	90		97	116
	75		59	91
	50		29	60
	25		16	34
	10		9	18
2	90	111	131	148
	75	84	109	124
	50	50	84	100
	25	36	59	72
	10	23	35	43
3	90	134	161	166
	75	104	137	139
	50	83	97	112
	25	59	79	91
	10	40	62	63

Grade	Percentile	Fall WCPM*	Winter WCPM*	Spring WCPM*
4	90	153	168	184
	75	125	143	160
	50	94	120	133
	25	75	95	105
	10	60	71	83
5	90	179	183	195
	75	153	160	169
	50	121	133	146
	25	87	109	119
	10	64	84	102
6	90	185	195	204
	75	159	166	173
	50	132	145	146
	25	112	116	122
	10	89	91	91

*WCPM = Words Correct Per Minute

TORC-4 (2009) Ages: 7:00-17:11

Relational Vocabulary - From the Student Question Booklet, the student reads a set of three words that are in some way related to each other. The student is to then silently read another four words and choose two words that are related to the first set of three words.

Sentence Completion - From the Student Question Booklet, the student silently reads a sentence that is missing two words. The student then silently reads a list of word pairs and chooses the word pair that best completes the sentence.

Paragraph Construction - After silently reading a list of sentences that are not in logical order, the student must then rearrange the sentences to form a coherent paragraph.

Text Comprehension - Students silently read a short passage and then answer five multiple-choice questions relative to the passage.

Contextual Fluency - This subtest measures how many individual words students can recognize, in 3 minutes, in a series of passages taken from the Text Comprehension Subtest. Each passage, printed in uppercase letters without punctuation or spaces between words, becomes progressively more difficult in content, vocabulary, and grammar. This is a format pioneered by J.P. Guilford to represent reading in his Structure of Intellect model. As students read the segments, they draw a line between as many words as they can in the time allotted. (E.g., THE LITTLE DOG JUMPED HIGH)

TORC-4 (cont.)

Sensitivity and Specificity are provided in context of criterion comparisons (with other existing tests)

*Important to understand sensitivity and specificity of other tests in order know if this comparison is appropriate.

16% of children in the normative sample presented with a disability (Manual, pg. 27)

- Speech/language - 5%
- ADHD -5%
- SLD -4%
- Emotional Disturbance- 1%
- Other -1%

Table 6.8
Predictive Outcome Analyses Demonstrating TORC-4's Ability to Predict Criterion Measures of Literacy

			Predictive outcome analysis			
Criterion measure	*N*	Level of acceptability	Sensitivity index	Specificity index	PPV	Percent agreement
ROS	1,150	I-A	.71	.70	.39	.69
WJ III	34	I-A	.80	.93	.67	.91
WLOS	190	—	.65	.77	.50	.74
TOWL-4	312	I-A	.82	.78	.54	.79
Global Literacy Ability	1,686	I-A	.72	.72	.43	.72

Note. PPV = positive predictive value; ROS = *Reading Observation Scale* (Hammill, Wiederholt, & Brown, 2009); WJ III = *Woodcock-Johnson III Achievement Tests* (Woodcock, McGrew, & Mather, 2001); WLOS = *Written Language Observation Scale* (Hammill & Larsen, 2009b); TOWL-4 = *Test of Written Language–Fourth Edition* (Hammill & Larsen, 2009a). Level of acceptability: I-A = sensitivity and specificity ≥ .70; I-B sensitivity and positive predictive value ≥ .70; II = sensitivity, specificity, and positive predictive value ≥ .70; III = sensitivity, specificity, and positive predictive value ≥ .75.

Manual pg. 53

A Note on Reading Comprehension

Many children with reading difficulties can read and comprehend short paragraphs containing factual information of decreased complexity but not **longer, more complex, and increasingly abstract age-level text**.

GORT-5 contains reading comprehension passages, which the students need to answer after the stimuli booklet has been removed from them. The passages are **far more simplistic then the academic texts** so the students may do well on this test **yet still continue to present with significant comprehension deficits**

TORC-4 *Test Comprehension* subtest contains reading comprehension passages, which the students need to answer via a multiple-choice format. The passages are **far more simplistic then the academic texts** so the students may do well on this test **yet present with significant comprehension deficits**

WJ-IV *Passage Comprehension* subtest gives the students sentences with a missing word, and the students are asked to orally provide the word. **However, filling-in a missing word does not adequately assess comprehension.**

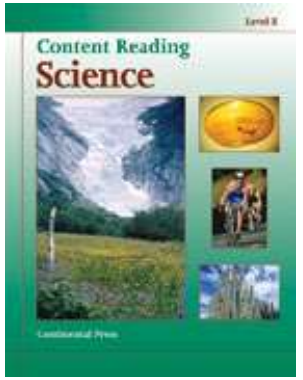
WIAT-4 *Reading Comprehension* subtest requires the student to read a passage and answer questions by referring back to the text. **Just because a student can look up the answers in text does not mean that they understand the text.**

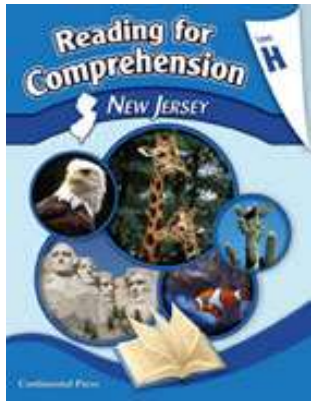
Clinical Reading Assessment

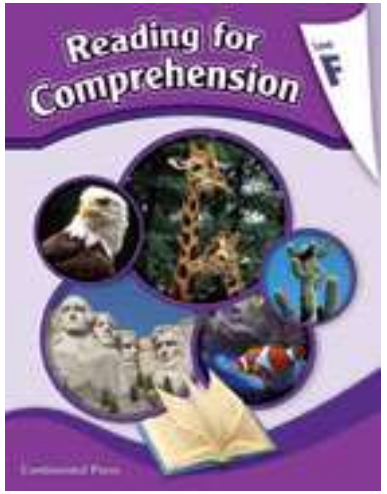
Comprehension Plus (Grades 1-6)
- Continental Press (HERE/HERE)
  - Reading for Comprehension (Grades 1-8)
    - General/Specific States (FL,IL, NJ, NY)
  - Content Reading (Grades 2-8)
    - Science
    - Social Science
    - Geography

Select grade level text

Ask student to read it
- Calculate error rate
- Calculate reading fluency sample (1 minute)
- Perform an in-depth error analysis
- Ask the main idea of text
- Ask to summarize key text points
- Ask abstract comprehension questions pertaining to the text
- Define text embedded words







Links to Examples:
Elementary
Adolescent

Takeaway

All standardized comprehensive and specialized reading tests (even the ones with excellent sensitivity and specificity) have limitations

All standardized test administration needs to be supplemented by clinical assessments of reading with a focus on advanced (age/grade level) basic reading attainment

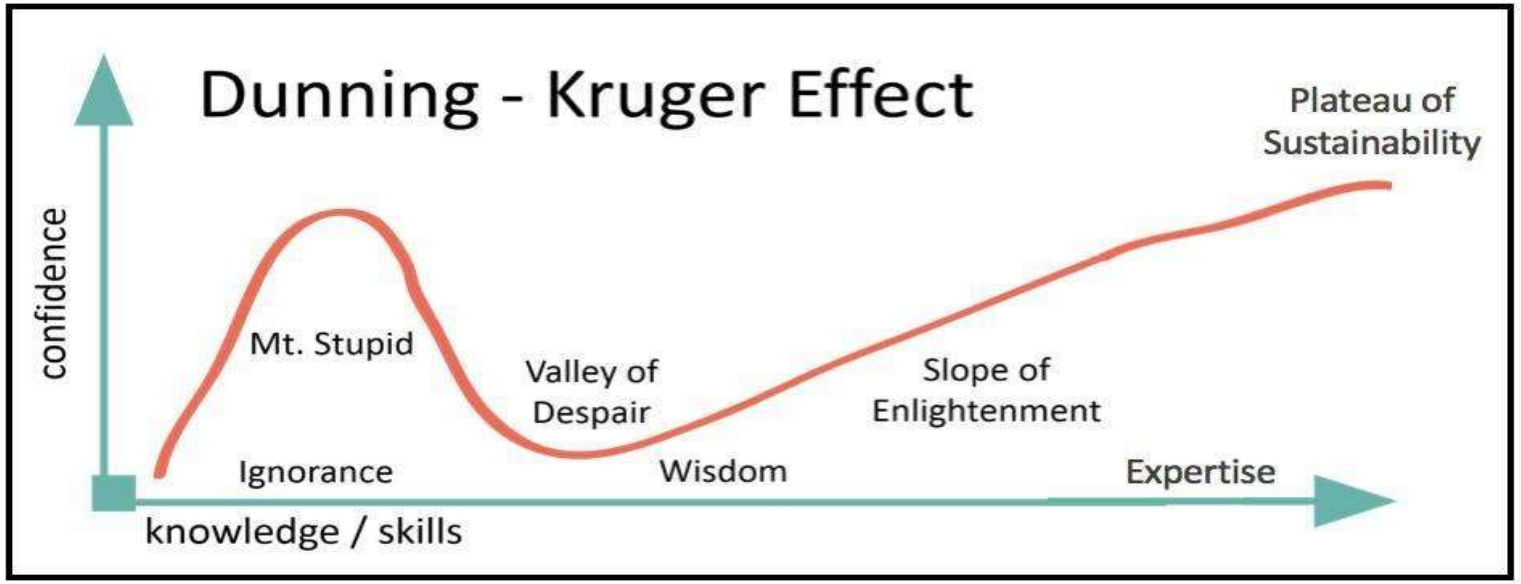
Standardized test presentation needs to be balanced with **scientifically defensible** clinical experience

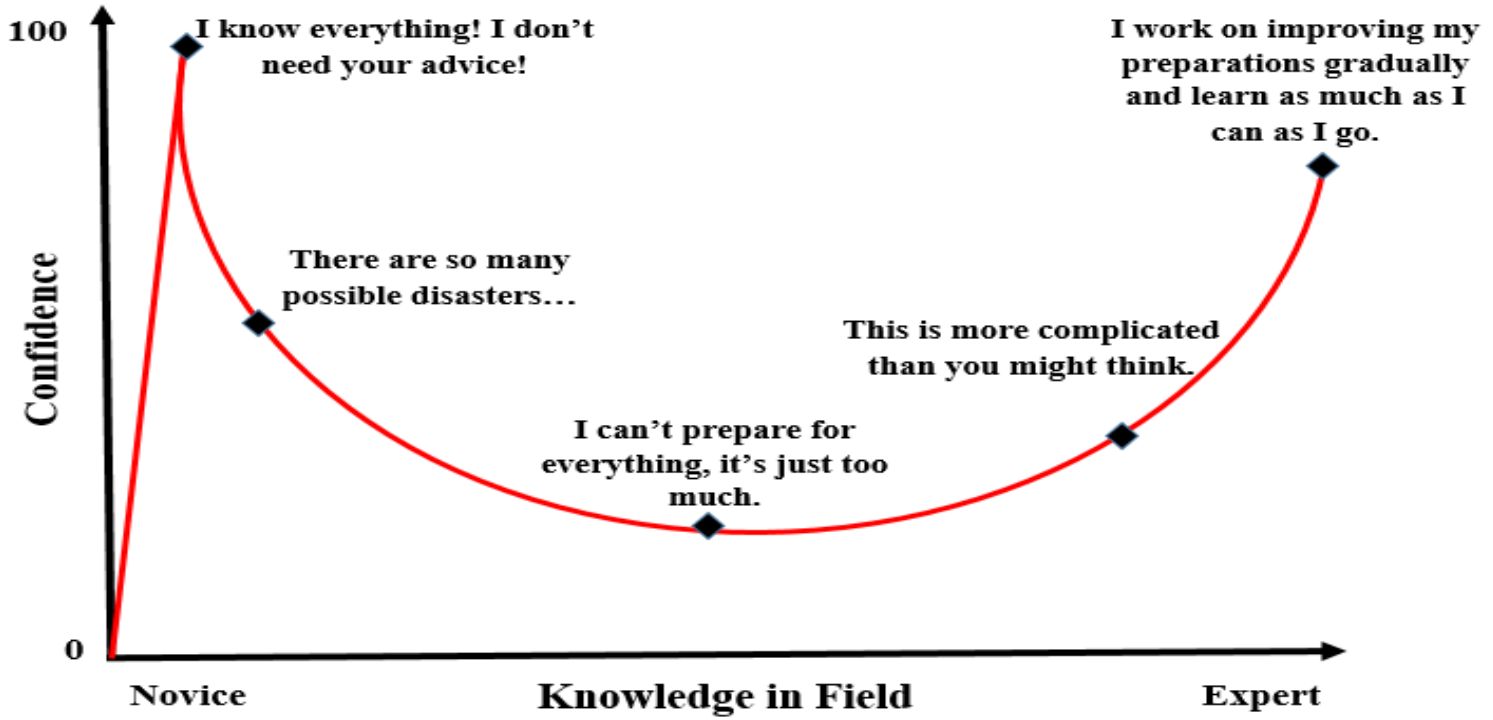
- This does not mean doing what you have seen works or doing whatever you like. This means following the science of reading (SOR) and consuming good quality reading research.

Recommendations for intervention need to be guided by informed assessment practices and not current reading related fads

Responses to scientifically informed recommendations

Be Mindful of the Dunning–Kruger Effect!





Conclusion

Because students with reading deficits continue to be underserved in the schools it is highly important to use both psychometrically-sound standardized assessments and clinical assessments (of relevant areas) in order to use targeted tasks **which adequately reflect the learner's difficulties in the "real world".**

It is important to ensure that assessments **yield diagnostic information** needed to formulate treatment goals for the student in question

All students need to receive fair and appropriate assessments which will result in targeted and relevant therapeutic services

Anything less is a denial of Free Appropriate Public Education (FAPE) to which **all students** are entitled to

It is NEVER too late to help students of any age, including adults!

Contact Information:
Tatyana Elleseff MA CCC-SLP

*Group: https://www.facebook.com/groups/EBPSLPs/

Blog: www.smartspeechtherapy.com/blog/

Business Page: www.facebook.com/SmartSpeechTherapyLlc

Twitter: https://twitter.com/SmartSPTherapy

Email: tatyana.elleseff@gmail.com

Shop: http://www.smartspeechtherapy.com/shop/