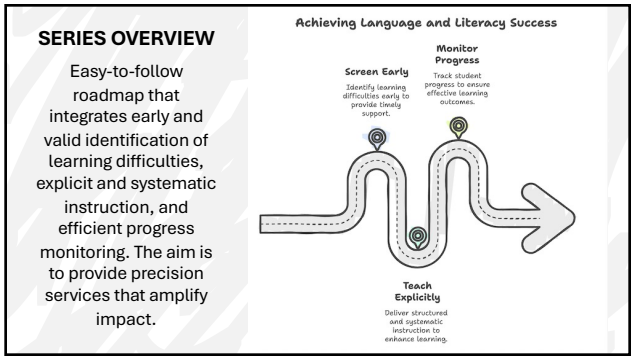

A Roadmap for Language and Literacy Success: Screen Early, Teach Explicitly, and Monitor Progress
Trina D. Spencer, PhD, BCBA-D
Douglas B. Petersen, PhD, BCS-CL
LEARNING DIFFICULTIES AUSTRALIA
60 years

1

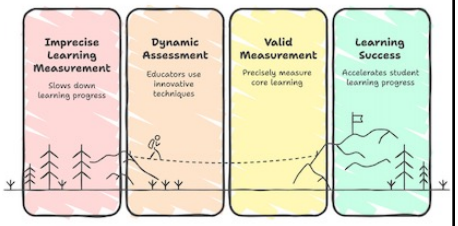

SERIES OVERVIEW
Easy-to-follow roadmap that integrates early and valid identification of learning difficulties, explicit and systematic instruction, and efficient progress monitoring. The aim is to provide precision services that amplify impact.
Achieving Language and Literacy Success
Screen Early
Identify learning difficulties early to provide timely support.
Monitor Progress
Track student progress to ensure effective learning outcomes.
Teach Explicitly
Deliver structured and systematic instruction to enhance learning.

2


Get on the Right Track by Measuring What Matters:
VALID Tier-1 Screening and Diagnostic Assessments for DLD and Dyslexia
Tuesday 26th August 10 am-12 pm (AEST)
Accelerate Learning with Multi-tiered, Explicit, and Systematic Language Instruction
SYSTEM-WIDE Explicit, Systematic Academic Language Instruction and Intervention
Tuesday 2nd September 10 am-12 pm (AEST)
Follow the Compass with Progress Monitoring
SYSTEM-WIDE *Valid* Benchmark Assessments and Progress Monitoring of Decoding and Academic Language
Tuesday 9th September 10 am – 12 pm (AEST)

3

Get on the Right Track by Measuring What Matters:
VALID Tier-1 Screening and Diagnostic Assessments for DLD and Dyslexia



4

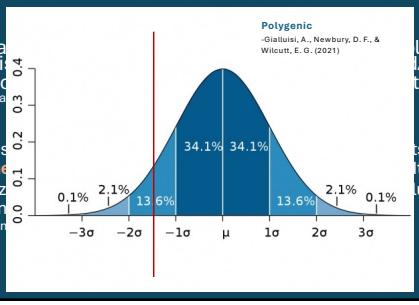
What is dyslexia?

"Dyslexia
origin. It i
word rec
–Internationa

logical in
/or fluent
ties."

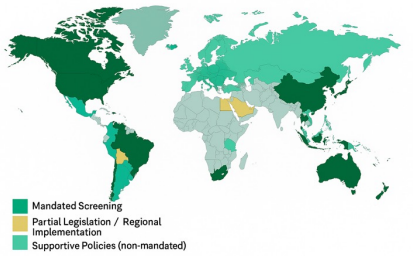
"Dyslexia is
developme
characteriz
recognition
-Snowling, Hulm

ts the
t is
luent word



5

Countries with Dyslexia Screening Legislation



6

Are Dyslexia Screeners Accurate?

	Sick	Healthy
	TRUE POSITIVES (TP)	FALSE POSITIVES (FP)
	FALSE NEGATIVES (FN)	TRUE NEGATIVES (TN)

$PVV = \frac{TP}{TP + FP}$

$NPV = \frac{TN}{TN + FN}$

Sensitivity $\frac{TP}{TP + FN}$

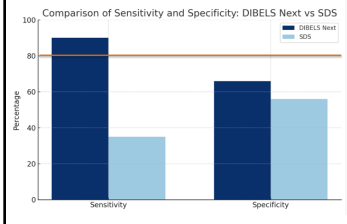
Specificity $\frac{TN}{TN + FP}$

Sensitivity: The percent of people with the disorder who are correctly identified as having the disorder.

Specificity: The percent of people without the disorder who are correctly identified as *not* having the disorder.

7

Are Dyslexia Screeners Accurate?



Burns et al. (2022)

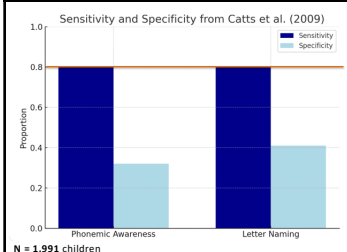
"Shaywitz Dyslexia Screen (SDS) and Dynamic Indicators of Basic Early Literacy Skills Next (DIBELS Next) were compared for 115 K-3 students with specific reading deficits using the Phonological Awareness Composite of the Comprehensive Test of Phonological Processing as the criterion."

"Results suggested that the decision accuracy for DIBELS Next (78%) was better than SDS (45%), and both sensitivity (**DIBELS Next = 90%, SDS = 35%**) and positive post-test probability (DIBELS Next = 71%, SDS = 42%) favored DIBELS Next."

...Specificity of 56% for the SDS and 66% for DIBELS Next.

8

Are Dyslexia Screeners Accurate?



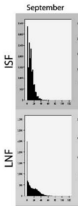
N = 1,991 children

Grade level at screening: Kindergarten

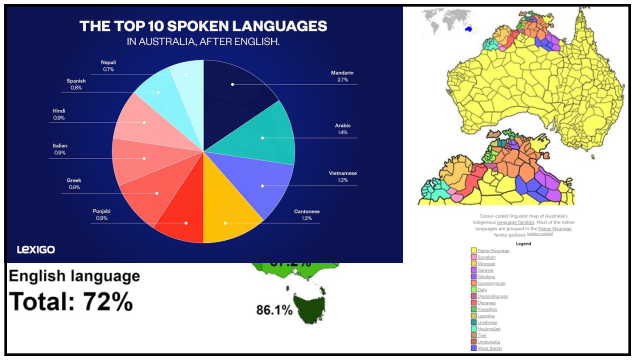
Phonemic Awareness:
80% sensitivity
32% Specificity

Letter Naming:
80% sensitivity
41% specificity

"...screening instruments designed to identify children at risk for reading disabilities continue to have limited predictive validity."



9

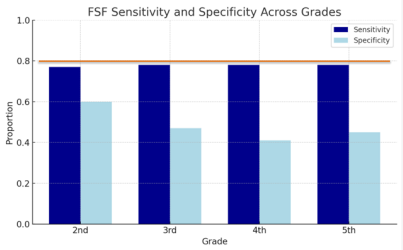


10

Are Dyslexia Screeners Accurate for Culturally and Linguistically Diverse Children?

-600 kindergarten students followed longitudinally to end of first grade, 300 followed to the end of fifth grade.

-50% Hispanic (primarily) Multilingual Students:
-Petersen, Gragg, & Spencer 2015

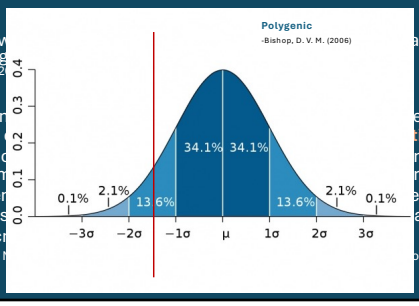


11

What is developmental language disorder (DLD)?

"People w… anding, and using …
-McGregor (2…

"Developr… ental condition … ties with language … ned by other bion… r, hearing impairmen… ese difficulties … ally persist acr…
-Bishop, D. V. M… ortium. (2017)



12

abc 123

Children with DLD are **4 times more likely to have math disabilities** and **6 times more likely to have reading disabilities.**

THE DLD PROJECT

1 IN 14 people have Developmental Language Disorder (DLD), an invisible, lifelong disability.

DLD is the accepted term in English-speaking countries replacing specific language impairment, language disorder and language delay.

13

Countries with DLD Screening Legislation

14

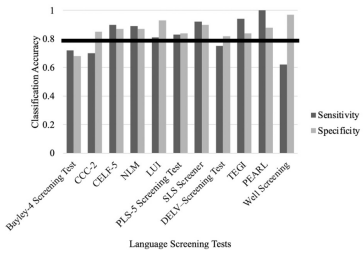
Are Current DLD Screeners Accurate?

Review Article

A Review of Screeners to Identify Developmental Language Disorde

Xue Bao,* Rouzana Komesidou,* and Tiffany P. Hogan*

*School of Health and Rehabilitation Sciences, MGH Institute of Health Professions, Boston



15

Test Accuracy When Administered to Diverse Students

2024

EVALUATION OF CULTURAL-LINGUISTIC FACT
LANGUAGE TESTING WITH ENGLISH LEARNE

Mayra Alejandra Reyes Ruiz

LSHSS

YOU HAVE ACCESS | Clinical Forum | January 2003

Alternative Assessment of Language and Literac
Diverse Populations

Authors Sandra P. Laing and Alan Kamhi AUTHORS INFO & AFFILIATIONS

Publication Language, Speech, and Hearing Services in Schools • Volume 34, Number 1 • Pag

Performance of ELL Referred Students on WISC-V, CTOPP-2, and TAPS-4

		Scale Scores		Standard Scores	
	n	M	SD	M	SD
WISC-V					
Similarities	38	6.18	3.00	80.92	14.97
Vocabulary	38	6.84	2.52	84.21	12.60
Block Design	39	8.44	2.26	92.18	11.29
Visual Puzzles	41	8.85	2.57	94.27	12.87
Matrix Reasoning	41	8.12	3.01	90.61	15.05
Figure Weights	41	8.85	2.60	94.27	13.02
Digit Span	40	5.55	2.28	77.75	11.38
Picture Span	41	7.46	2.25	87.32	11.24
Coding	40	7.90	3.02	89.50	15.10
Symbol Search	37	8.46	2.85	92.30	14.27
CTOPP-2					
Elision	59	5.25	2.07	76.27	10.36
Blending Words	59	6.12	2.44	80.51	12.24
Phoneme Isolation	44	5.68	1.74	78.41	8.679
Memory for Digits	50	5.66	2.16	78.30	10.81
Nonword Repetition	51	6.35	3.26	81.96	16.10
Rapid Digit Naming	52	7.13	2.24	85.67	11.20
Rapid Letter Naming	47	6.77	2.43	83.83	12.17

16

Are Static Dyslexia and DLD Screeners Accurate When Administered to Young Children?

Are Static Dyslexia and DLD Screeners Accurate When Administered to Culturally and Linguistically Diverse Students?

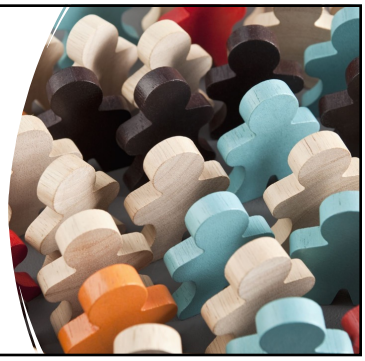
The results from most currently available static screeners cannot be validly interpreted when administered to culturally and linguistically diverse students.

17

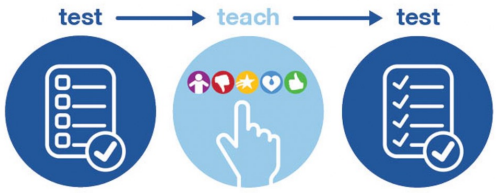
Assessment Bias

Sensitivity and specificity is often lower for culturally and linguistically diverse students

- High potential for Content and Linguistic Bias
- Assessment in L1 does not guarantee valid results
- Static assessment cannot control for cultural, linguistic, and historical factors



18



Dynamic Assessment

19

Feuerstein's Theory of Structural Cognitive Modifiability

"What if intelligence can be taught and was in fact the ability to learn?"
Feuerstein et al. , 1979)

1. Structural Cognitive Modifiability (SCM)
Intelligence is not fixed, instead, it is modifiable.

2. Mediated Learning Experience (MLE)
Learning is most effective when an adult mediates the environment.

3. Learning Potential Instead of Static Scores
Dynamic assessment provides insight into learning potential

20

How does dynamic assessment compare to static assessment?

Static Assessment
Measure a person's current skills and performance

Assumes equal access to learning

Dynamic Assessment
Measures a person's ability or potential to learn skills

Acknowledges unequal access to learning

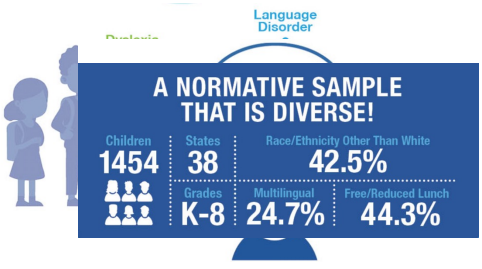
21

What are the Results of Research on Dynamic Assessment (and why are they so strong)?

	Sensitivity	Specificity	Aspect of Language Measured
Forbush Romero et al., 2021	100%	100%	Narratives
Kramer et al., 2009	92%	100%	Narratives
Laurie & Pesco, 2023	85.78%	100%	Narratives
Peña et al., 2006	100%	100%	Narratives
Peña et al., 2007	93.3%	92%	Narratives
Peña et al., 2014	89%	89%	Narratives
Petersen et al., 2017	100%	100%	Narratives
Petersen et al., 2020	100%	95.2%	Vocabulary

22

23



24

DYMOND™

Diagnostic Assessment for DLD and Dyslexia

- Identify students who have DLD and/or Dyslexia
- Help determine eligibility for special education
- Differentiate difference from disorder
- Identify strengths, weaknesses, and intervention targets

WHO: Students (K-8) referred for special education eligibility

EXAMINERS: Individuals who have specialized training in assessment, evaluation, and diagnosis of language and/or reading disorders.

Tier 1 Screening for DLD and Dyslexia

- Determine risk for DLD and Dyslexia among all students, regardless of cultural and linguistic background
- Fulfill legislative requirements
- Reduce over-identification

WHO: All students (K-8)

EXAMINERS: Educators trained to administer the DYMOND (e.g., general education teacher)

25

DYMOND Subtests Administration Time

Diagnosis/Screening

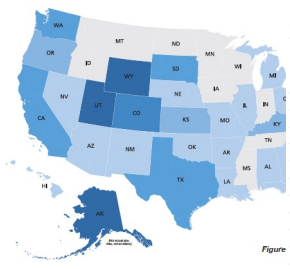
- Dynamic Assessment of Narrative Discourse (DAND)
  - **6-8 minutes**
- Dynamic Assessment of Decoding (DAD)
  - **5-7 minutes**

Identifying Strengths, Weaknesses and Informing Instruction

- Dynamic Assessment of Inferential Word Learning (DAIWL)
  - 2 minutes
- Rapid Automatized Naming (RAN)
  - 2 minutes

26

1454 typically developing children tested in 80 sites across 36 U.S. states



Figure

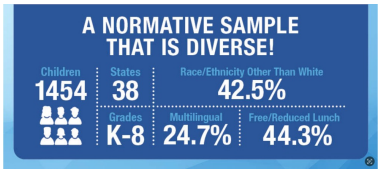
Table 4.1. Demographic Characteristics of the Normative Sample.

Characteristic	N=1454	% of Sample	US Census %
GENDER			
Female	740	50.9%	50.4%
Male	714	49.1%	49.6%
RACE/ETHNICITY			
American Indian/Alaska Native	45	3.1%	1.3%
Asian	39	2.7%	6.3%
Black/African American	120	8.3%	13.6%
Hispanic	376	25.9%	19.1%
Multiracial/Other	30	2.1%	3.0%
Native Hawaiian/Pacific Islander	8	0.6%	0.3%
White	836	57.5%	58.9%
MULTILINGUAL	359	24.7%	20.6%

27

The majority of norm-referenced tests do not include multilingual children who have limited English proficiency in the normative sample. This could potentially eliminate representation of 1 in 4 children across the U.S.

28



359 (24.7%) of the DYMOND normative sample included multilingual children.

Children who received a score of 1 on the WIDA assessment (WIDA, 2020) were still eligible to be included in the normative sample.

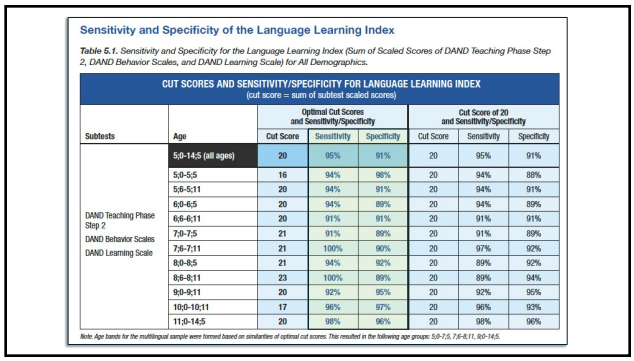
29

It is the test developers' responsibility to adequately make the case that the test can be used for its intended purposes (validity)

Purpose 1 and 2 of the DYMOND: Identifying Language Learning Disorder and Decoding Learning Disorder

The Results of the DYMOND *can be inferred to indicate a probable language learning disorder and decoding learning disorder for the majority of children across the U.S., including those who are culturally and linguistically diverse*

30



Sensitivity and Specificity of the Language Learning Index

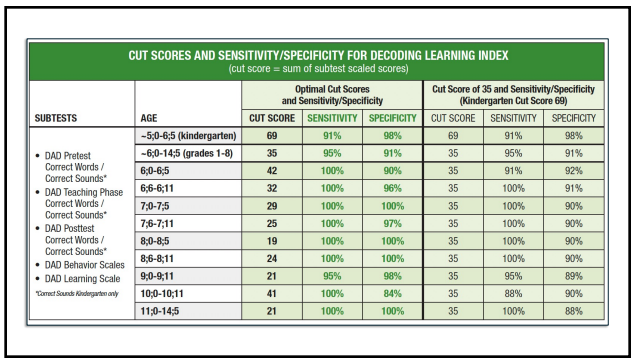
Table 5.1. Sensitivity and Specificity for the Language Learning Index (Sum of Scaled Scores of DAND Teaching Phase Step 2, DAND Behavior Scales, and DAND Learning Scale) for All Demographics.

CUT SCORES AND SENSITIVITY/SPECIFICITY FOR LANGUAGE LEARNING INDEX
(cut score = sum of subtest scaled scores)

		Optimal Cut Scores and Sensitivity/Specificity			Cut Score of 20 and Sensitivity/Specificity		
Subtests	Age	Cut Score	Sensitivity	Specificity	Cut Score	Sensitivity	Specificity
DAND Teaching Phase Step 2 DAND Behavior Scales DAND Learning Scale	5;0-14;5 (all ages)	20	95%	91%	20	95%	91%
	5;0-5;5	16	94%	98%	20	94%	88%
	5;6-5;11	20	94%	91%	20	94%	91%
	6;0-6;5	20	94%	89%	20	94%	89%
	6;6-6;11	20	91%	91%	20	91%	91%
	7;0-7;5	21	91%	89%	20	91%	89%
	7;6-7;11	21	100%	90%	20	97%	92%
	8;0-8;5	21	94%	92%	20	89%	92%
	8;6-8;11	23	100%	89%	20	89%	94%
	9;0-9;11	20	92%	95%	20	92%	95%
	10;0-10;11	17	96%	97%	20	96%	93%
	11;0-14;5	20	98%	96%	20	98%	96%

Note. Age bands for the multilingual sample were formed based on similarities of optimal cut scores. This resulted in the following age groups: 5;0-7;5, 7;6-8;11, 9;0-14;5.

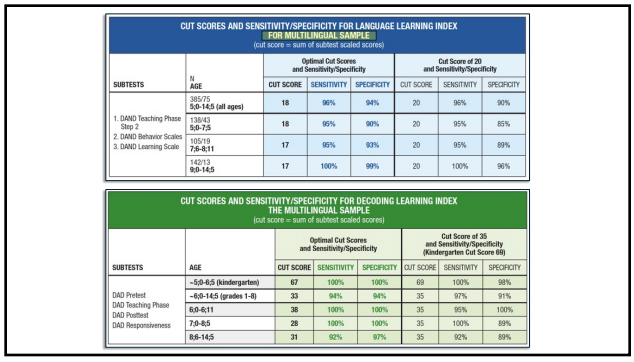
31



CUT SCORES AND SENSITIVITY/SPECIFICITY FOR DECODING LEARNING INDEX
(cut score = sum of subtest scaled scores)

		Optimal Cut Scores and Sensitivity/Specificity			Cut Score of 35 and Sensitivity/Specificity (Kindergarten Cut Score 69)		
SUBTESTS	AGE	CUT SCORE	SENSITIVITY	SPECIFICITY	CUT SCORE	SENSITIVITY	SPECIFICITY
• DAD Pretest Correct Words / Correct Sounds* • DAD Teaching Phase Correct Words / Correct Sounds* • DAD Posttest Correct Words / Correct Sounds* • DAD Behavior Scales • DAD Learning Scale *Correct Sounds Kindergarten only	~5;0-6;5 (kindergarten)	69	91%	98%	69	91%	98%
	~6;0-14;5 (grades 1-8)	35	95%	91%	35	95%	91%
	6;0-6;5	42	100%	90%	35	91%	92%
	6;6-6;11	32	100%	96%	35	100%	91%
	7;0-7;5	29	100%	100%	35	100%	90%
	7;6-7;11	25	100%	97%	35	100%	90%
	8;0-8;5	19	100%	100%	35	100%	90%
	8;6-8;11	24	100%	100%	35	100%	90%
	9;0-9;11	21	95%	98%	35	95%	89%
	10;0-10;11	41	100%	84%	35	88%	90%
	11;0-14;5	21	100%	100%	35	100%	88%

32



CUT SCORES AND SENSITIVITY/SPECIFICITY FOR LANGUAGE LEARNING INDEX
FOR MULTILINGUAL SAMPLE
(cut score = sum of subtest scaled scores)

			Optimal Cut Scores and Sensitivity/Specificity			Cut Score of 20 and Sensitivity/Specificity		
SUBTESTS	N AGE		CUT SCORE	SENSITIVITY	SPECIFICITY	CUT SCORE	SENSITIVITY	SPECIFICITY
1. DAND Teaching Phase Step 2 2. DAND Behavior Scales 3. DAND Learning Scale	385/75 5;0-14;5 (all ages)		18	96%	94%	20	96%	90%
	138/43 5;0-7;5		18	95%	90%	20	95%	85%
	105/19 7;6-8;11		17	95%	93%	20	95%	89%
	142/13 9;0-14;5		17	100%	99%	20	100%	96%

CUT SCORES AND SENSITIVITY/SPECIFICITY FOR DECODING LEARNING INDEX
THE MULTILINGUAL SAMPLE
(cut score = sum of subtest scaled scores)

		Optimal Cut Scores and Sensitivity/Specificity			Cut Score of 35 and Sensitivity/Specificity (Kindergarten Cut Score 69)		
SUBTESTS	AGE	CUT SCORE	SENSITIVITY	SPECIFICITY	CUT SCORE	SENSITIVITY	SPECIFICITY
DAD Pretest DAD Teaching Phase DAD Posttest DAD Responsiveness	~5;0-6;5 (kindergarten)	67	100%	100%	69	100%	98%
	~6;0-14;5 (grades 1-8)	33	94%	94%	35	97%	91%
	6;0-6;11	38	100%	100%	35	95%	100%
	7;0-8;5	28	100%	100%	35	100%	89%
	8;6-14;5	31	92%	97%	35	92%	89%

33

CUT SCORES AND SENSITIVITY/SPECIFICITY FOR LANGUAGE LEARNING INDEX BY RACE/ETHNICITY FOR ALL AGES
(cut score = sum of subtest scaled scores)

SUBTESTS	Race/Ethnicity	Optimal Cut Scores and Sensitivity/Specificity			Cut Score of 20 and Sensitivity/Specificity		
		CUT SCORE	SENSITIVITY	SPECIFICITY	CUT SCORE	SENSITIVITY	SPECIFICITY
1. DAND Teaching Phase Step 2 2. DAND Behavior Scales 3. DAND Learning Scale	White	20	98%	94%	20	96%	90%
	Hispanic	21	92%	91%	20	95%	85%
	Black	19	95%	88%	20	95%	89%
		16	93%	96%			
	Asian	20	100%	98%	20	100%	96%
	American Indian, Native Alaskan, Hawaiian/Pacific Islander, and Multiracial	19	88%	89%	20	88%	82%

CUT SCORES AND SENSITIVITY/SPECIFICITY FOR DECODING LEARNING INDEX BY RACE/ETHNICITY FOR GRADES 1-8
(cut score = sum of subtest scaled scores)

SUBTESTS	Race/Ethnicity	Optimal Cut Scores and Sensitivity/Specificity			Cut Score of 35 and Sensitivity/Specificity		
		CUT SCORE	SENSITIVITY	SPECIFICITY	CUT SCORE	SENSITIVITY	SPECIFICITY
DAD Pretest DAD Teaching Phase DAD Posttest DAD Responsiveness	White	33	90%	92%	35	97%	89%
	Hispanic	31	100%	97%	35	100%	93%
	Black	35	92%	91%	35	92%	91%
	Asian	32	100%	100%	35	100%	96%
	American Indian, Native Alaskan, Hawaiian/Pacific Islander, and Multiracial	33	100%	87%	35	100%	84%

34

CUT SCORES AND SENSITIVITY/SPECIFICITY FOR LANGUAGE LEARNING INDEX BY SOCIOECONOMIC STATUS (SES; FREE/REDUCED LUNCH) FOR ALL AGES
(cut score = sum of subtest scaled scores)

SUBTESTS	SES; FREE/REDUCED LUNCH	Optimal Cut Scores and Sensitivity/Specificity			Cut Score of 20 and Sensitivity/Specificity		
		CUT SCORE	SENSITIVITY	SPECIFICITY	CUT SCORE	SENSITIVITY	SPECIFICITY
1. DAND Teaching Phase Step 2 2. DAND Behavior Scales 3. DAND Learning Scale	Free/Reduced Lunch	20	92%	90%	20	92%	90%
	Paid Lunch	19	95%	95%	20	96%	94%
		21	97%	93%			

CUT SCORES AND SENSITIVITY/SPECIFICITY FOR DECODING LEARNING INDEX BY SOCIOECONOMIC STATUS (SES; FREE/REDUCED LUNCH) FOR GRADES 1-8
(cut score = sum of subtest scaled scores)

SUBTESTS	SES; FREE/REDUCED LUNCH	Optimal Cut Scores and Sensitivity/Specificity			Cut Score of 35 and Sensitivity/Specificity		
		CUT SCORE	SENSITIVITY	SPECIFICITY	CUT SCORE	SENSITIVITY	SPECIFICITY
DAD Pretest DAD Teaching Phase DAD Posttest DAD Responsiveness	Free/Reduced Lunch	32	92%	90%	35	96%	86%
		29	91%	95%			
	Paid Lunch	33	100%	94%	35	100%	91%

35

Testing the Limits

- Administered to multilingual Japanese/English-speaking school-age children in Japan.
- 149 kindergarten through sixth-grade students from three different sites across Japan.
- All participants were multilingual, with home languages including Japanese, Mandarin, Cantonese, Korean, and Russian, and varied degrees of English language proficiency.

36

Testing the Limits

- Despite the varying levels of English proficiency, nearly all children were able to complete the DYMOND. Only children who could not understand the basic English instructions necessary to complete the tasks were discontinued from the study.
- The results indicated 100% sensitivity and 94% specificity.

37

LANGUAGE DISORDER

DYNAMIC ASSESSMENT OF NARRATIVE DISCOURSE

1. TEACHING PHASE Model Story with Pictures & Icons
2. TEACHING PHASE Step 1: Pictures & Icons
3. TEACHING PHASE Step 2: Icons only
4. RESPONSIVENESS
5. POSTTEST

38

Example of how to use Overcorrection procedure with Level 1 and Level 2 Prompts

- it didn't work
- it still tasted like mustard

"What happened after he tried to scrape off the mustard?"

"After he took a bite, it still tasked like disgusting mustard. You say that."

"Go back to the attempt. Remember to tell me the consequence."

Level 1 + Overcorrection

Child: *(omits consequence in independent retell)* ✗

Examiner Provides Level 1 Prompt: *"What happened after he tried to scrape off the mustard?"*

Child: *"It still tasted gross."* ✓

Examiner Provides Overcorrection procedure: *"Good job! Go back to the attempt (point to attempt icon) and remember to tell me the consequence."*

Level 1 + Level 2 + Overcorrection

Child: *(omits consequence in independent retell)* ✗

Examiner Provides Level 1 Prompt: *"What happened after he tried to scrape off the mustard?"*

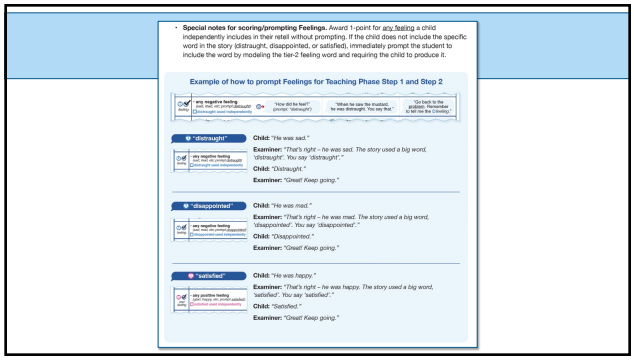
Child: *"I can't remember."* ✗

Examiner Provides Level 2 Prompt: *"After he took a bite, it still tasted like disgusting mustard. You say that."*

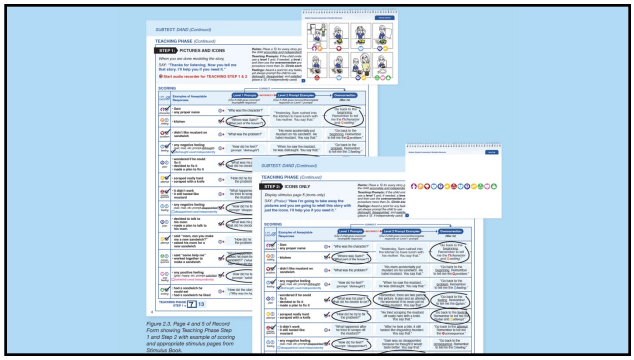
Child: *"He took a bite. It still tasted gross."*

Examiner Provides Overcorrection procedure: *"Good job! Now go back to the attempt (point to attempt icon) and remember to tell me the consequence."*

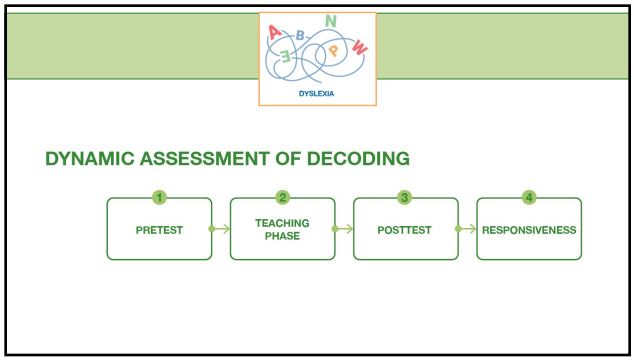
39


Special notes for scoring/prompting Feelings. Award 1-point for any feeling a child independently includes in their retell without prompting. If the child does not include the specific word in the story (distraught, disappointed, or satisfied), immediately prompt the student to include the word by modeling the tier-2 feeling word and requiring the child to produce it.
Example of how to prompt Feelings for Teaching Phase Step 1 and Step 2
"distraught"
Child: "He was sad."
Examiner: "That's right – he was sad. The story used a big word, 'distraught'. You say 'distraught'."
Child: "Distraught."
Examiner: "Great! Keep going."
"disappointed"
Child: "He was mad."
Examiner: "That's right – he was mad. The story used a big word, 'disappointed'. You say 'disappointed'."
Child: "Disappointed."
Examiner: "Great! Keep going."
"satisfied"
Child: "He was happy."
Examiner: "That's right – he was happy. The story used a big word, 'satisfied'. You say 'satisfied'."
Child: "Satisfied."
Examiner: "Great! Keep going."

40



41


DYSLEXIA
DYNAMIC ASSESSMENT OF DECODING
1 PRETEST
2 TEACHING PHASE
3 POSTTEST
4 RESPONSIVENESS

42


Language Dynamics GROUP
ASSESSMENT TOOLS
CUBED
PEARL
DYMOND
SKILLED READING
DECODING
LANGUAGE
INSTRUCTION TOOLS
INSTRUCTION TOOLS

43


The Possible Reality

44