Processing Strengths and Weaknesses for LD Identification: Reliability and Validity of Proposed Methods

Presentation for the International Dyslexia Association

October 22, 2015

Jeremy Miciak, Ph.D.

University of Houston

There are no conflicts of interest.

TEXAS CENTER for LEARNING DISABILITIES

Reading for success

The Texas Center for Learning Disabilities (TCLD) investigates the classification, early intervention, and remediation of learning disabilities.

Present Controversy: RTI…then what?

- **LDA White Paper Consensus**

  1. Maintain SLD definition and strength statutory requirements

  2. Present methods of LD identification are not sufficient

  3. A PSW method makes the most empirical and clinical sense

  4. RTI for prevention, but should be followed by comprehensive evaluation for LD

  5. Assessment of cognitive processes informs LD identification and intervention

CRITICAL ISSUES IN RESPONSE-TO-INTERVENTION, COMPREHENSIVE EVALUATION, AND SPECIFIC LEARNING DISABILITIES IDENTIFICATION AND INTERVENTION: AN EXPERT WHITE PAPER CONSENSUS

J. Hale, V. Alfonso, V. Berninger, B. Bracken, C. Christo, E. Clark, M. Cohen, A. Davis, S. Decker, M. Denckla, R. Dumont, C. Elliott, S. Feifer, C. Fiorello, D. Flanagan, E. Fletcher-Janzen, D. Geary, M. Gerber, M. Gerner, S. Goldstein, N. Gregg, R. Hagin, L. Jaffe, A. Kaufman, N. Kaufman, T. Keith, F. Kline, C. Kochhar-Bryant, J. Lerner, G. Marshall, J. Mascolo, N. Mather, M. Mazzocco, G. McCloskey, K. McGrew, D. Miller, J. Miller, M. Mostert, J. Naglieri, S. Ortiz, L. Phelps, B. Podhajski, L. Reddy, C. Reynolds, C. Riccio, F. Schrank, E. Schultz, M. Semrud-Clikeman, S. Shaywitz, J. Simon, L. Silver, L. Swanson, A. Urso, T. Wasserman, J. Willis, D. Wodrich, P. Wright, & J. Yalof[1]

Abstract. Developed in concert with the Learning Disabilities Association of America (LDA), this White Paper regarding specific learning disabilities identification and intervention represents the expert consensus of 58 accomplished scholars in education, psychology, medicine, and the law. Survey responses and empirical evidence suggest that five conclusions are warranted: 1) The SLD definition should be maintained and the statutory requirements in SLD identification procedures should be strengthened; 2) neither ability-achievement discrepancy analysis nor failure to respond to intervention alone is sufficient for SLD identification; 3) a "third method" approach that identifies a pattern of psychological processing strengths and weaknesses, and achievement deficits consistent with this pattern of processing weaknesses, makes the most empirical and clinical sense; 4) an empirically-validated RTI model could be used to prevent learning problems, but comprehensive evaluations should occur for SLD identification purposes, and children with SLD need individualized interventions based on specific learning needs, not merely more intense interventions; and 5) assessment of cognitive and neuropsychological processes should be used for both SLD identification and intervention purposes.

LDA White Paper Argues:

- RTI is insufficient for LD identification

    - There is no consensus on what type of RTI to use for LD identification

    - There is no consensus on a measurement model for defining responsiveness

    - There is no "true positive" in an RTI model
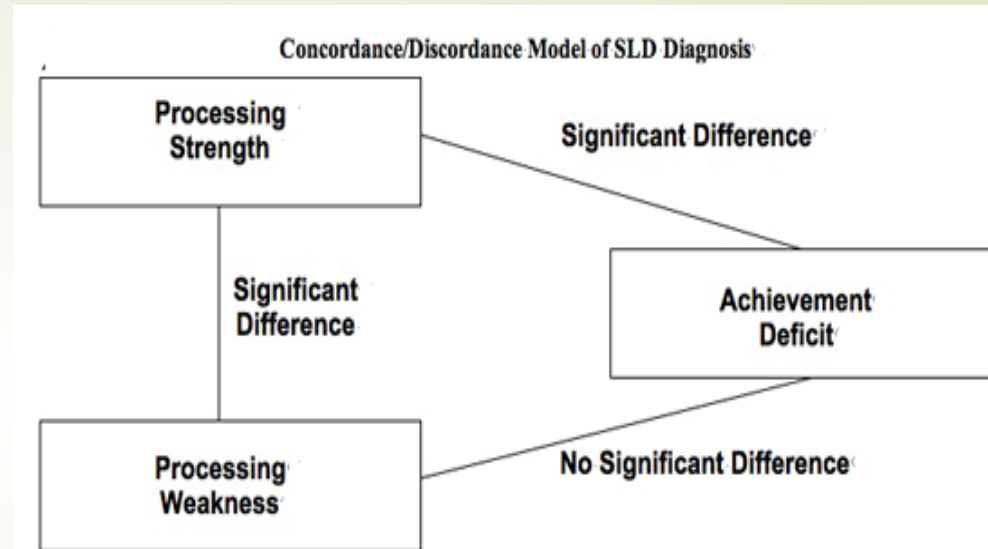
What are PSW methods for LD identification?

- Methods aimed at identifying a (necessary) intra-individual pattern of cognitive processing strengths and weaknesses that may explain academic under acheivement.

- Differentiates LD from garden variety low achievement

- Three proposed methods to operationalize:

  Concordance/ Discordance Model (Hale & Fiorello, 2004)

  Cross Battery Assessment approach (Flanagan et al., 2007)

  Discrepancy/ Consistency Method (Naglieri, 1999)

PSW Methods: Concordance/ Discordance Model (C/DM)

1. Concordance between academic deficit and theoretically related cognitive processing skill

2. Discordance between academic deficit and theoretically unrelated cognitive processing skill

3. Discordance between cognitive processing weakness and strength (Hale & Fiorello, 2004)

Concordance/Discordance Model of SLD Diagnosis

Processing Strength

Significant Difference

Significant Difference

Achievement Deficit

Processing Weakness

No Significant Difference

Source: Hale, J. B., & Fiorello, C. A. (2004). *School Neuropsychology: A Practitioner's Handbook*. New York, NY: Guilford Press.

PSW Methods: Cross Battery Assessment (XBA)

1. Normative deficit in academic achievement area

2. Normative deficit in a cognitive processing skill

3. Theoretical link between deficits in academic and cognitive areas

4. Otherwise "normal" profile of cognitive abilities

5. Exclusionary criteria are met

Limited Research on PSW Models

- Proponents cite:

  - Improved psychometric methods (Hale et al., 2006; Flanagan et al., 2007; Naglieri, 1999)

  - Relation between cognitive processing skills and academic deficits (Johnson et al., 2010)

- No empirical investigation of

  - identification rates of proposed PSW methods

  - agreement between proposed methods

  - academic profiles and demographic characteristics of proposed groups

Study 1: Technical Adequacy of PSW Methods

Since we know the relations between cognitive processes and achievement, as well as the reliabilities of the tests, how reliably could PSW methods measure "true LD"?

Simulation of PSW Methods (Stuebing et al., SPR, 2012)

- Created data sets where LD status of child is known; asked how well 3 PSW methods identified those children known to demonstrate LD at the observed level.

- Based on the idea that cognitive assessments should occur after Tier 2

- For all 3 methods, number of children identified as LD low (about 2-3% depending on size of discrepancy)

not LD," highly accurate (high specificity and few false negatives), but if "yes LD", many false positives (low PPV)

Of 10,000 assessments:

- CDM: 1,558 identified as LD (8,436 as not LD); 25 correct, so 1,533 are false positives and get the wrong treatment

- DCM: 362 identified as LD (9,638 not LD); 89 correct, so 273 are false positives and get the wrong treatment

- XBA: 678 would be identified as LD (9,322 not LD); 353 correct, 325 are false positives and get the wrong treatment

Study 2: Equivalence and validity of C/DM and XBA PSW Methods

If we apply 2 PSW methods to a pool of inadequate responders, what happens? ID rates? Do they agree on ID decisions? Do resulting groups differ?

Participants

- 7 middle schools

- 139 7th and 8th grade students who:

   1. Demonstrated inadequate response to Tier 2 intervention

   2. Completed comprehensive assessment battery

- Inadequate response to intervention defined as SS < 90 on:

   - WJ-III Basic Reading

   - TOWRE, or

   - WJ-III Passage Comp

Table 1.

Demographics statistics for both adequate and inadeaquate responders.

Variable	Inadequate Responders N = 139	Adequate Responders N = 89
Age		
M	11.92	11.44
SD	0.75	0.51
% Male	51.8	47.19
% F/R Lunch	83.46	75.32
% ESL	13.53	4.49
Race/Ethnicity		
% Black	43.88	46.07
% White	5.04	17.98
% Hispanic	48.92	34.83
% Other	2.16	1.12

Measures

Academic Measures

- ## GRADE

- ## WJ-III
  - Letter Word ID*
  - Word Attack*
  - Reading Comprehension*

- ## TOWRE*

- ## TOSREC

- ## WJ-III Spelling

*Used to determine adequate response

Cognitive Processing Measures

- CTOPP
  - Blending (G_a)
  - Elision (G_a)
  - Rapid Naming (G_{lr})
- KBIT-2
  - Matrices (G_f)
  - Verbal Knowledge (G_c)
- Underlining Test (G_s)

- Test of Spatial Working Memory (G_{sm})
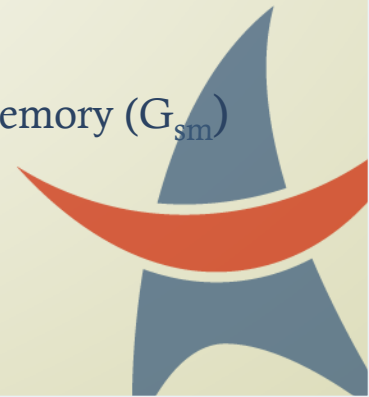
Identification Rates and Area of Deficit

Academic deficit area(s) of students meeting and not meeting LD criteria according to PSW methods

Area of Deficit						
Basic reading	Reading fluency	Reading comprehension	C/DM < 85 N (%)	C/DM < 90 N (%)	XBA < 85 N (%)	XBA < 90 N (%)
yes	yes	yes	1 (0.7)	3 (2.1)	0	0
yes	yes	no	4 (2.8)	4 (2.8)	1 (0.7)	2 (1.4)
yes	no	yes	1 (0.7)	3 (2.1)	2 (1.4)	2 (1.4)
yes	no	no	5 (3.6)	7 (5)	6 (4.3)	8 (5.8)
no	yes	yes	0	3 (2.1)	0	0
no	no	yes	19 (13.7)	32 (23)	5 (3.6)	12 (8.6)
no	yes	no	11 (7.9)	14 (10.1)	10 (7.1)	10 (7.1)
no	no	no	98 (70.5)	73 (52.5)	115 (82.7)	105 (75.5)
Total meeting criteria:			41 (29.5)	66 (47.5)	24 (17.3)	34 (24.5)

Agreement on LD Identification

*Agreement on LD identification between the C/DM and XBA methods at different low achievement cut points*

| Approach: | Approach | | | |
| | Table 5. | | | |
	C/DM < 85	C/DM < 90	XBA < 85	XBA < 90
C/DM < 85	-	62.1	30.0	13.6
C/DM < 90	-	-	20.0	20.5
XBA < 85	0.31	0.11	-	23.4
XBA < 90	-0.04	0.03	-	-

Below diagonal = kappa; above diagonal = percentage overlap (total identified by both approaches/ total identified).

Performance on External Variables
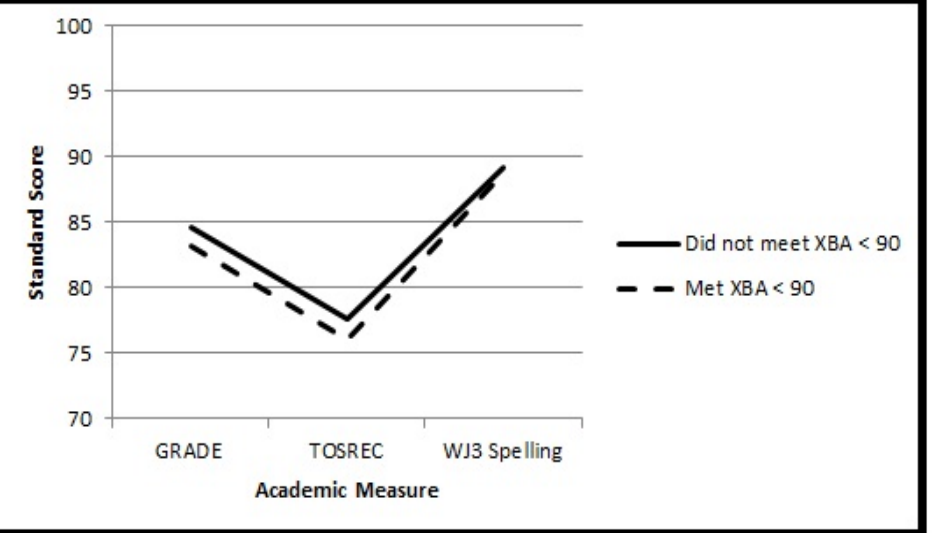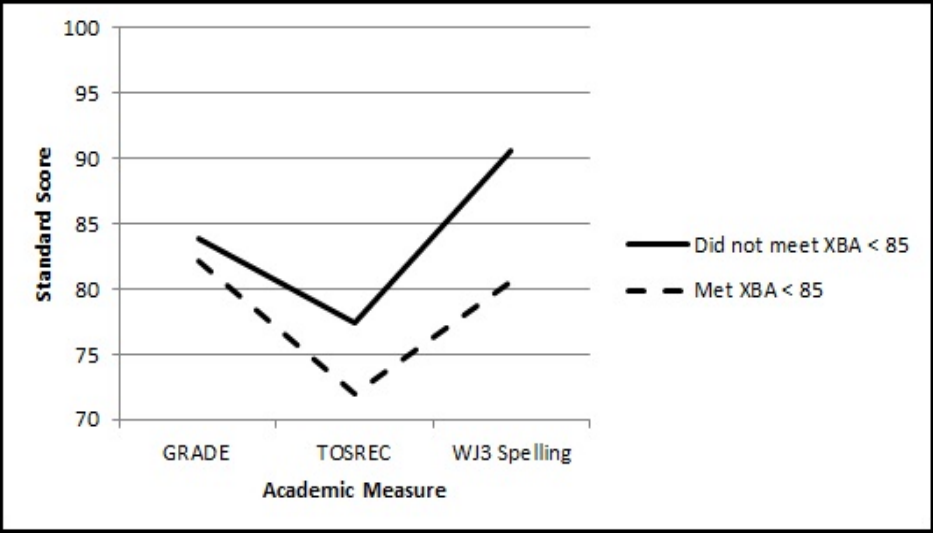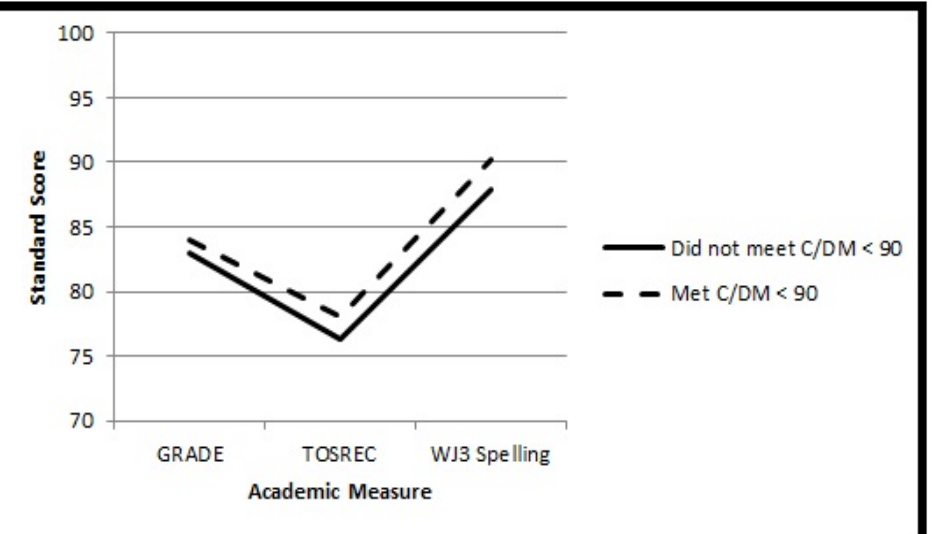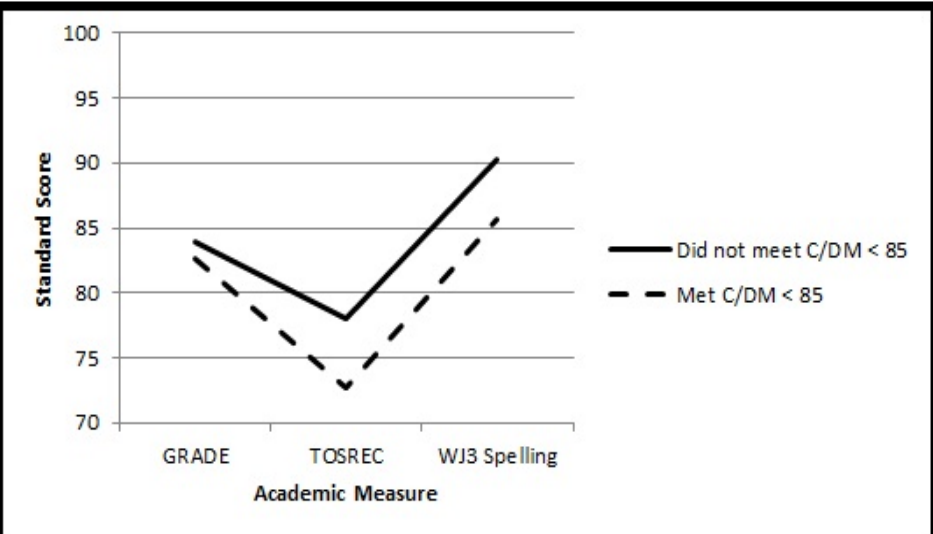
Study 3: The Effect of Measures Selection on PSW Identification

Do the measures we use to identify LD using a PSW method impact resulting decisions?
What happens if we change the measures we use?

Methods

- Participants: 177 2nd grade students who did not respond to school-based Tier 2 intervention

- Assessed in fall, prior to Tier 3 intervention

- All students empirically classified as meeting or not meeting C/DM LD criteria according to 2 different, but theoretically equivalent assessment batteries

Assessment Batteries

| Reading Domain | Academic Measures | | Related Cognitive Processing Measures[1] |
	Assessment Battery 1	Assessment Battery 2	
Basic Reading	WJ3 Letter/Word ID	WJ3 Word Attack	CTOPP Phonological Awareness
Reading Fluency	TOWRE Phonemic Decoding	TOWRE Sight Words	CTOPP Rapid Letter Naming
Reading Comprehension	WJ3 Passage Comp	Gates MacGinitie Passage Comp	KBIT-2 Verbal Knowledge

Agreement with a cut-point of 85

Academic Battery 2	Academic Battery 1		
	Non LD	LD	Total
Non LD	92	23	115
LD	26	36	62
Total	118	59	177

Kappa = .38; Percentage overlap = 41%
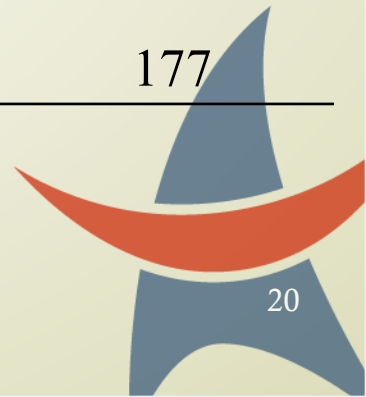
Agreement with a cut-point of 90

| | Academic Battery 1 | | |
Academic Battery 2	Non LD	LD	Total
Non LD	69	34	103
LD	28	46	74
Total	97	80	177

Kappa = .28; Percentage Overlap = 43%
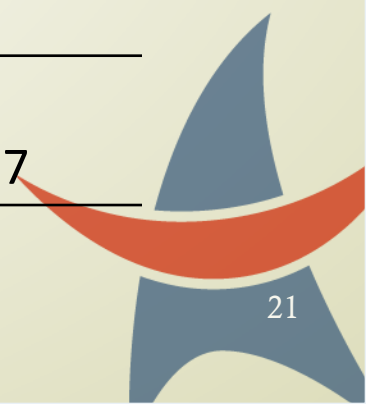
Study 4: Simulation Evaluating the Effect of Measures Selection on PSW Identification

Can different measurement batteries ever reliably identify LD using a PSW method?

Are these results specific to the sample and measures in Miciak, Taylor et al.?

- Simulated > 70,000 latent correlations between a cognitive strength, cognitive weakness, and academic weakness

- Generated observed values with varying degrees of reliability

- Used alternate indicators of academic achievement (with varying correlation and reliability)

- Compared agreement for battery 1 and battery 2

- Mean positive agreement of .42

- Taylor et al. (in review)

Agreement Tables

Achievement Value				Test 2	
				LD	No
< 85	Test 1	LD		46.7 (40.4)	52.4 (33.5)
		No		52.4 (33.5)	3519.1 (170.7)
< 90		LD		74.4 (60.7)	74.7 (45.0)
		No		74.7 (45.1)	3446.7 (196.7)

Note: Values in parentheses are standard deviations.

Cool Graphic

Positive agreement with achievement < 85 by correlation for 3 levels of reliability.
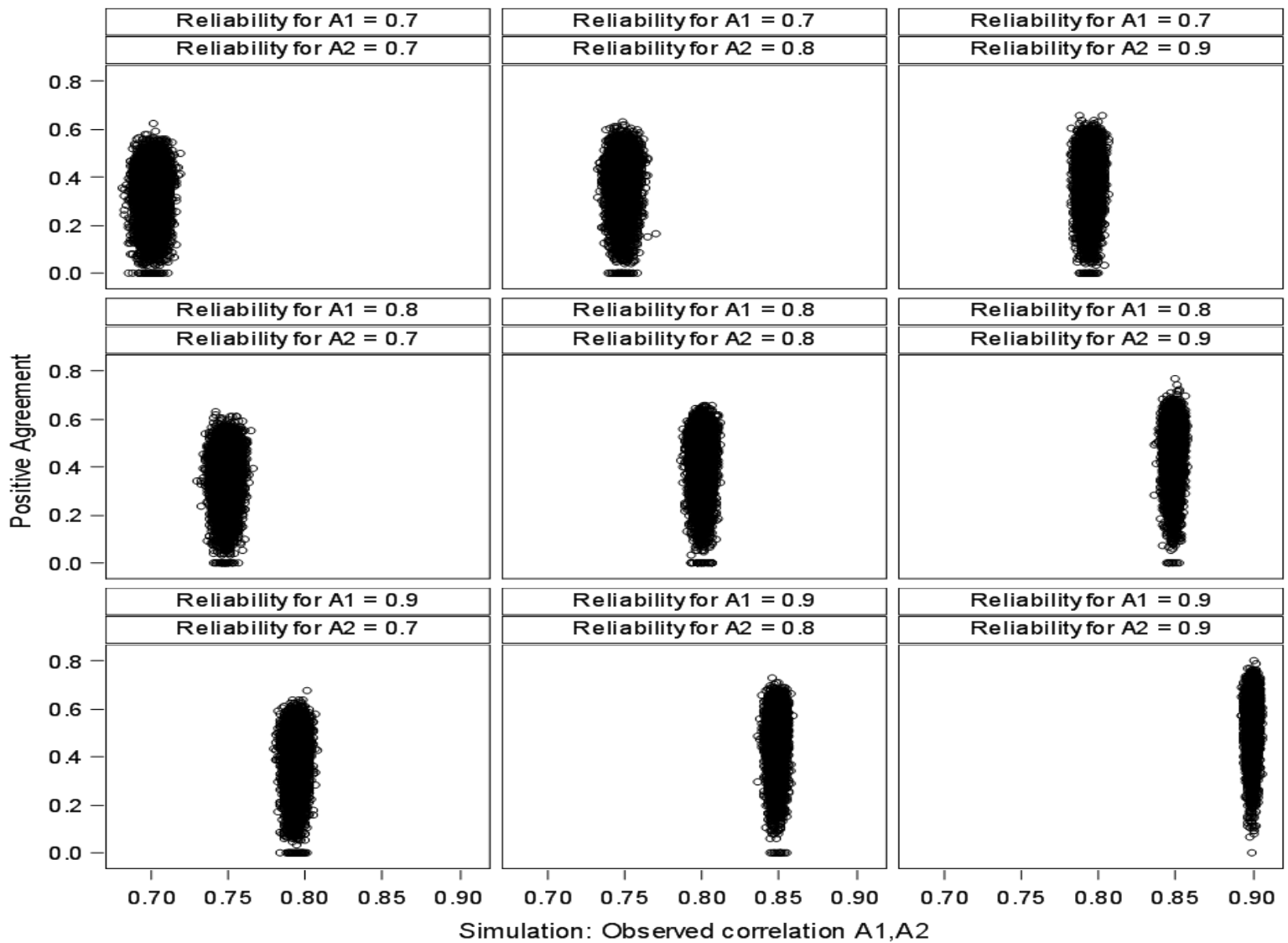
Study 5: PSW Status and Treatment Response

Do students who meet PSW LD criteria according to XBA and C/DM methods respond differently to reading intervention?

Study 3: Miciak et al., 2015

Identify students as "LD" or "not LD" by C/DM and XBA

Intensive Intervention in Reading

Evaluate Posttest Performance

Hypothesis: To the extent PSW status is educationally meaningful, students should respond differently to the same intervention.

But first, replication

- Do the XBA and C/DM Methods identify the same students as LD?

Table 3			
Agreement for LD identification decisions for the XBA and C/DM methods for LD Identification			
	C/DM		
XBA Method	LD	Not LD	Total
LD	59	31	90
Not LD	64	52	116
Total	123	83	206
Kappa = -.10; XBA = Cross Battery Assessment Method (Flanagan et al. 2007); C/DM = Concordance Discordance Model (Hale & Fiorello, 2004);			
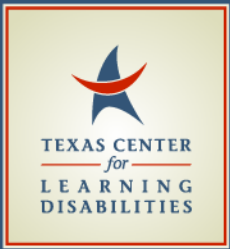
Results

- Conducted 39 contrasts to evaluate whether LD status or inclusionary criteria specified by the C/DM & XBA methods were significant predictors of intervention response.

- Among 39 contrasts, 4 were statistically significant.

- Eta squared statistics for significant contrasts ranged from .008 - .018

- Only one contrast (Gc => Word Reading) predicted > 1% of the variance at posttest. Pretest predicted between 53% and 69% of the variance at posttest.
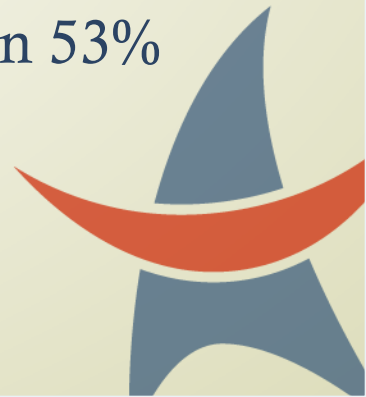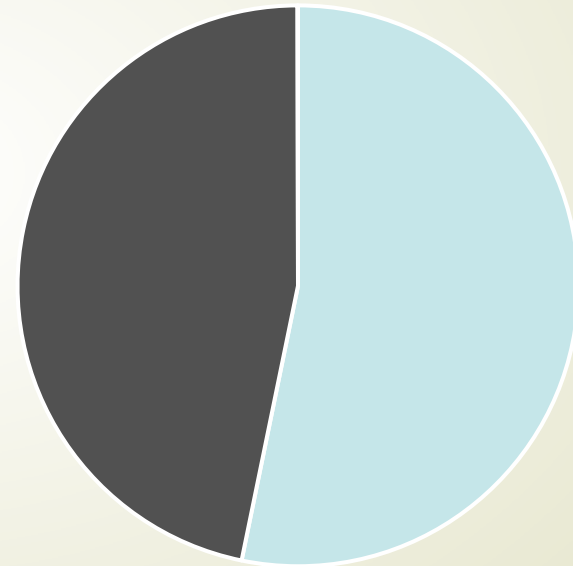
Reading Comprehension at Posttest

Variability Explained in Passage
Comprehension at Posttest

Variability Explained in Passage
Comprehension at Posttest

■ Pretest ■ Error ■ C/DM LD

■ Pretest ■ Error ■ XBA LD
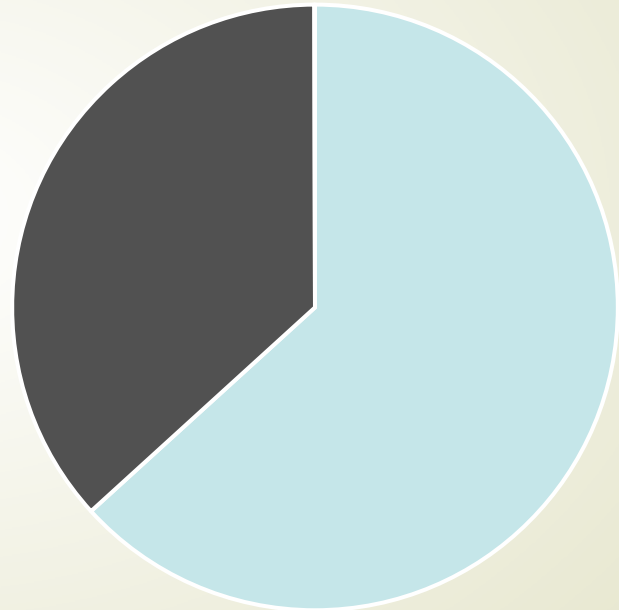
Reading Fluency at Posttest

Variability Explained in Reading
Fluency at Posttest
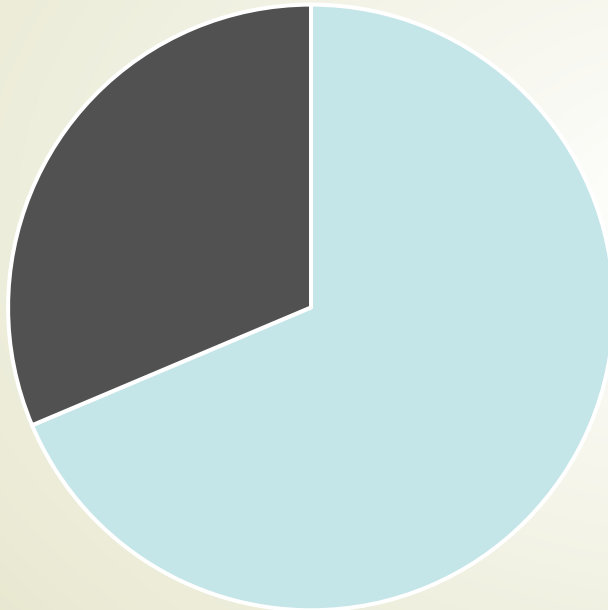
Variability Explained in Reading
Fluency at Posttest

■ Pretest ■ Error ■ C/DM LD
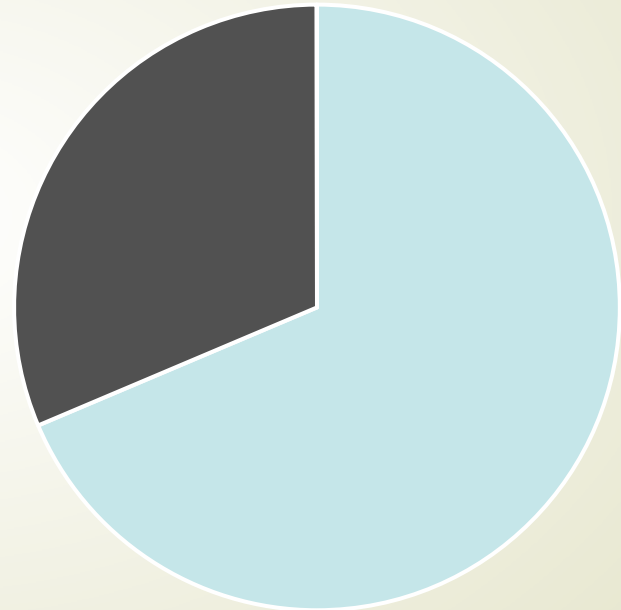
■ Pretest ■ Error ■ XBA LD

Word Reading at Posttest

Variability Explained in Word Reading at Posttest

Pretest   Error   C/DM LD
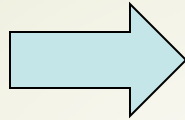
Variability Explained in Word Reading at Posttest

Pretest   Error   XBA LD

How much better can we predict responders?
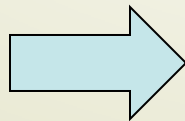
Pretest only

Cross tabulation of predictions based on r^2 = .828 and a cut point for pass/fail of z < --.66		
	Pass	Fail
Pass	670	76
Fail	76	178
Total number of misclassifications = 152		

Pretest + Gc Status

| Cross tabulation of predictions based on r^2 = .838 and cut point for pass/fail of z < -.66		
	Pass	Fail
Pass	672	73
Fail	74	181
Total number of misclassifications = 147		

Discussion
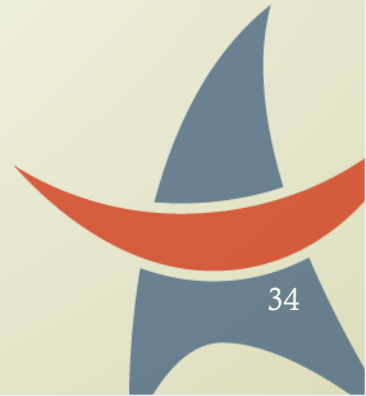
PSW Reliability Summary

- Generally, PSW Methods identify few students. Hours of testing for every 1 student identified with LD.

- PSW Methods are not interchangeable. Different methods will identify different students as LD.

- PSW Methods do not overcome problems of poor reliability at the individual level.

All methods are unreliable, right?

- Cognitive discrepancy methods like the C/DM exacerbate reliability problems.

  - Increased complexity

  - More factors measured = more opportunity for error

  - Difference scores are typically less reliable than their constituent scores; they contain error from two measures

82   94

88   99

93   105
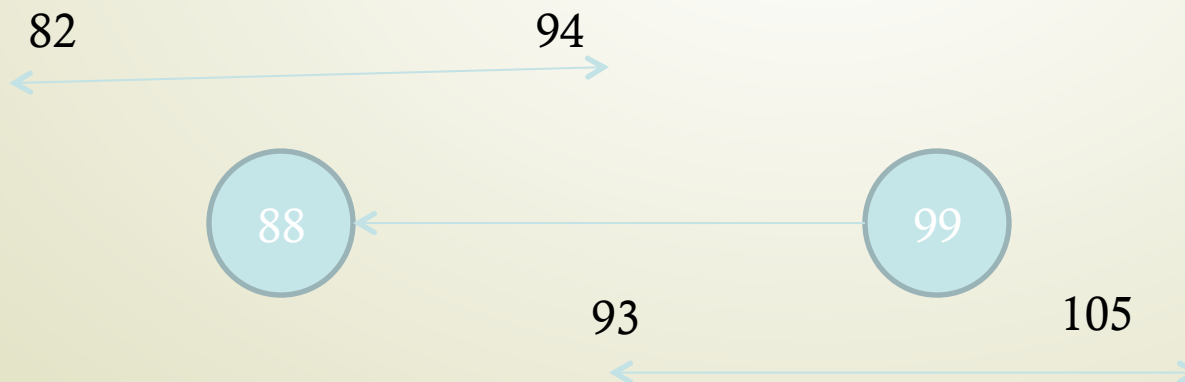
Within 95% CI
Difference score range:
-1 - 23

What do cognitive assessments add?

- **No evidence that PSW pattern causes differential treatment response.**

- Processing subtypes weakly related to intervention outcomes; little evidence that knowledge of cognitive strengths and weaknesses facilitates intervention (Kearn & Fuchs, 2013; Pashler et al., 2010)

- No additional information not found in achievement profiles.

- Delays treatment.

THANK YOU

jeremy.miciak@times.uh.edu