

The Prevention and Identification of Reading Disability

Douglas Fuchs, Donald L. Compton,
Lynn S. Fuchs, and Joan D. Bryant

National Research Center on Learning Disabilities
(NRCLD)

Vanderbilt University

OSEP Grant #H324U010004

Purpose and Method of NRCLD Experimental RTI Studies in Reading and Math

- 3 purposes across reading and math studies
 1. Examine efficacy of 1st-grade preventive tutoring
 2. Assess RD and MD prevalence and severity as a function of classification method with and without tutoring
 3. Explore pretreatment cognitive abilities associated with development of reading and math skills
- Random assignment to 1st-grade study conditions
- Longitudinal follow up to assess stability of disability (RD: 1st-4th grade; MD: 1st-3rd grade) as a function of treatment and methods of disability classification
- Reading and math studies initiated in consecutive years so samples do not overlap

Study Purpose

Using a 2-tier model (not including special education) in 1st grade, explore RTI as a means of preventing and identifying RD

In This Presentation ...

- Study overview
- How does prevention affect RD prevalence?
- How do different RTI measures and classification methods affect RD prevalence and severity?
- How many data points are necessary to achieve a reliable slope?

Study Overview

Overview

- In Fall of first grade, we identified low performers in 42 classes in 16 schools in Middle TN.
- We randomly assigned them to 3 conditions:
 - Fall Tutoring
 - Spring Tutoring (if unresponsive to general education during the Fall)
 - Control.
- We provided small-group instruction to children assigned to Fall Tutoring.
- We used Fall progress monitoring to identify children unresponsive to Tier 1 general education.
- We provided small-group instruction to Spring Tutoring students who were unresponsive to general education.

Overview

- We collected weekly WIF data: 9 waves in Fall and 9 waves in Spring
- We administered a battery of standardized reading tests at Fall, mid-year, end of grade 1, end of grade 2
- To address RTI prevention, we contrasted Spring Tutoring and Control groups at Fall, mid-year (when tutoring was initiated for unresponsive students in the Spring Tutoring group), and end of grades 1 and 2.
- To address RTI identification, we compared classification rates for Spring Tutoring and Control groups at end of grades 1 and 2 (will follow through grade 4).

Districts, Schools, and Teachers

- 2 school districts in Tennessee (urban Metro-Nashville and suburban Williamson County)
- 8 Title 1 and 8 non-Title 1 elementary schools
- 42 first-grade teachers assigned randomly within schools to PALS ($n = 21$) and No-PALS ($n = 21$); in this presentation, we collapse PALS and No-PALS classes

Identifying “At-Risk” Students

- In the 42 classes, all students screened on:
 - RLN (CTOPP)
 - CBM Word Identification Fluency
 - Teacher judgment
- The 6 lowest students per class on both measures, also judged as “lowest” by the teacher, were designated “low study entry.”

Study Conditions

- In each class, the 6 “low study entry” were rank ordered and split into top and bottom strata.
- Within each stratum, children were randomly assigned to:
 - *Fall Tutoring* ($n = 84$)
 - *Spring Tutoring* ($n = 84$) -- if unresponsive to general education
 - *Control* ($n = 84$)
- For prevention questions, we use spring tutoring and control students; for identification question, we use students from all 3 conditions in the database through end of grade 2.

Study Conditions (Cont'd)

- We used “dual discrepancy” on Fall WIF slope and level to identify the subset of students in spring tutoring and control groups unresponsive ($.75 \text{ SD} < \text{normative mean}$) to Tier 1 general education.
 - Spring Tutored: $n = 40$
 - Control: $n = 24$

Tutoring

- Groups of 2-4 students
- Validated treatment protocol
 - Letter-sound correspondence, decoding words, sight word recognition, fluency-building, and partner reading, with point system for motivation
 - 9 wks, 4x per wk, 35-45 min per session
- Fidelity
 - All sessions audiotaped
 - Tapes of sessions #14 and #28 checked for all tutors against a 79-item checklist
 - Inter-rater agreement on coding of tapes was 96% across sessions and tutors
 - > 95% tutor fidelity across sessions and tutors

Measures

- *Screening*
 - CTOPP Rapid Letter Naming, CBM-WIF
- *Progress Monitoring*
 - Weekly CBM-WIF
- *Fall*
 - CTOPP (Elision, Memory for Digits); 4-subtest WASI; WRMT-R (WI, WA); Woodcock Diagnostic Reading Battery (Listening Comprehension); comprehensive cognitive battery also administered.
- *Mid-Year*
 - WRMT-R (WID, WA); TOWRE (Sight Word Reading, Phonemic Decoding)
- *End-Year and End Grade 2*
 - WJ (Passage Comprehension); WRMT-R (WID, WA), TOWRE (Sight Word Reading, Phonemic Decoding); Woodcock Diagnostic Reading Battery (Listening Comprehension); Social Skills Rating System (SSRS; short form); Teacher Rating of Reading Effort

Prevention Question

- Can prevention affect RD prevalence?

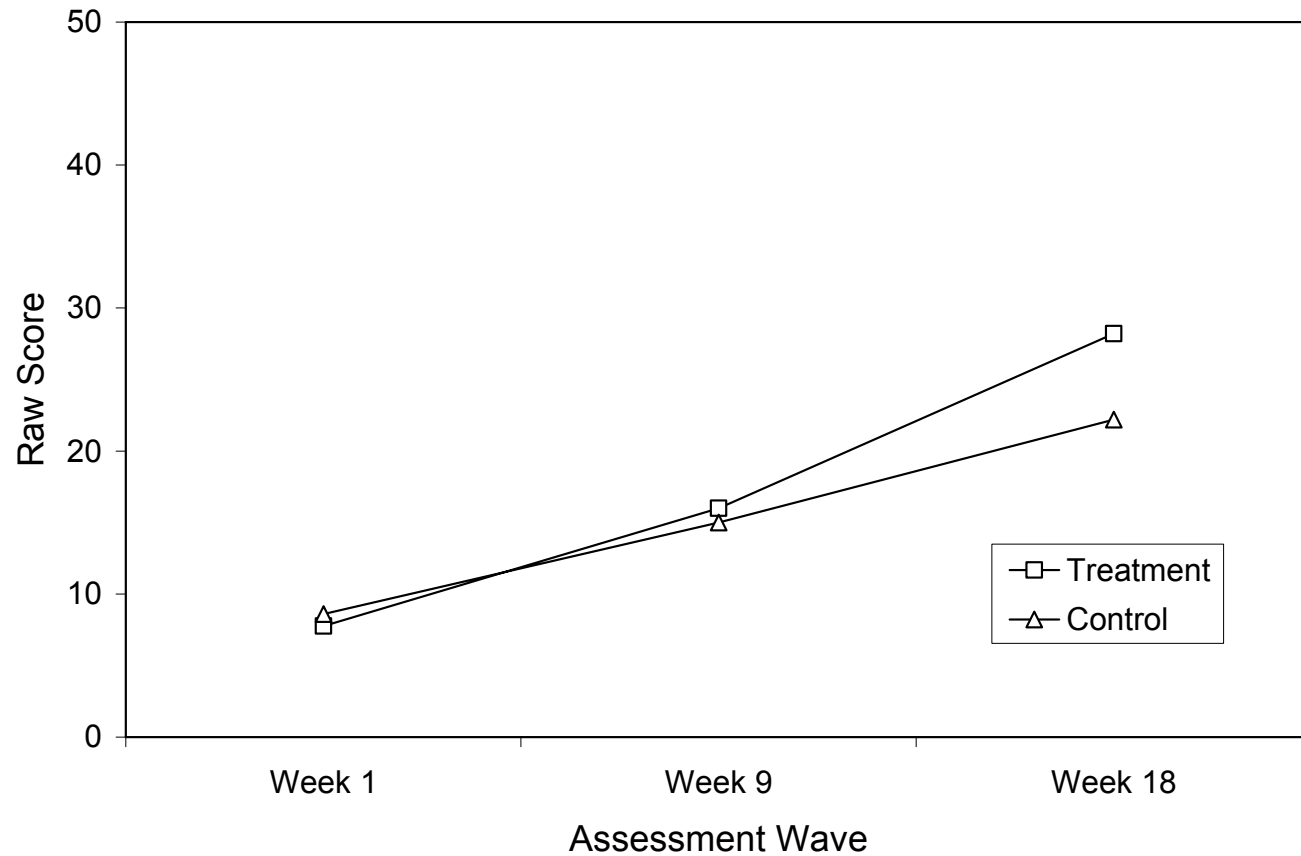
Group Comparability

- Tier 1 unresponsive students (in Spring tutoring and control groups) comparable on:
 - IQ
 - Vocabulary
 - CTOPP Rapid Digit Naming, Elision, Memory for Digits
 - WRMT WID and WA
 - TOWRE Sight Word and Phonemic Decoding
 - Teacher Ratings of Effort and Distractibility
- They were:
 - ~ 2/3 *SD* < mean on WIF local norms
 - ~ 2/3 *SD* < national norms on IQ, Vocabulary, Phonological Processing
 - 1/3 to 2/3 *SD* < national norms on reading measures
 - Teachers' mean effort rating ~ 60%
 - Teachers' mean distractibility rating between “sometimes” and “very often”

Effects of Tutoring: Progress Monitoring Data

- Multi-level modeling with HLM (time was nested within the child; child was nested within tutoring condition)
- 2-piece model: an intercept (at mid-year) and two slope terms (fall and spring)
- WIF was adequately explained with a 2-piece model
- Spring Tutored and Control groups showed similar growth from Fall to mid-year, prior to tutoring (slope 1).
- Spring Tutored group had greater growth from mid-year to end-year, during tutoring (slope 2).

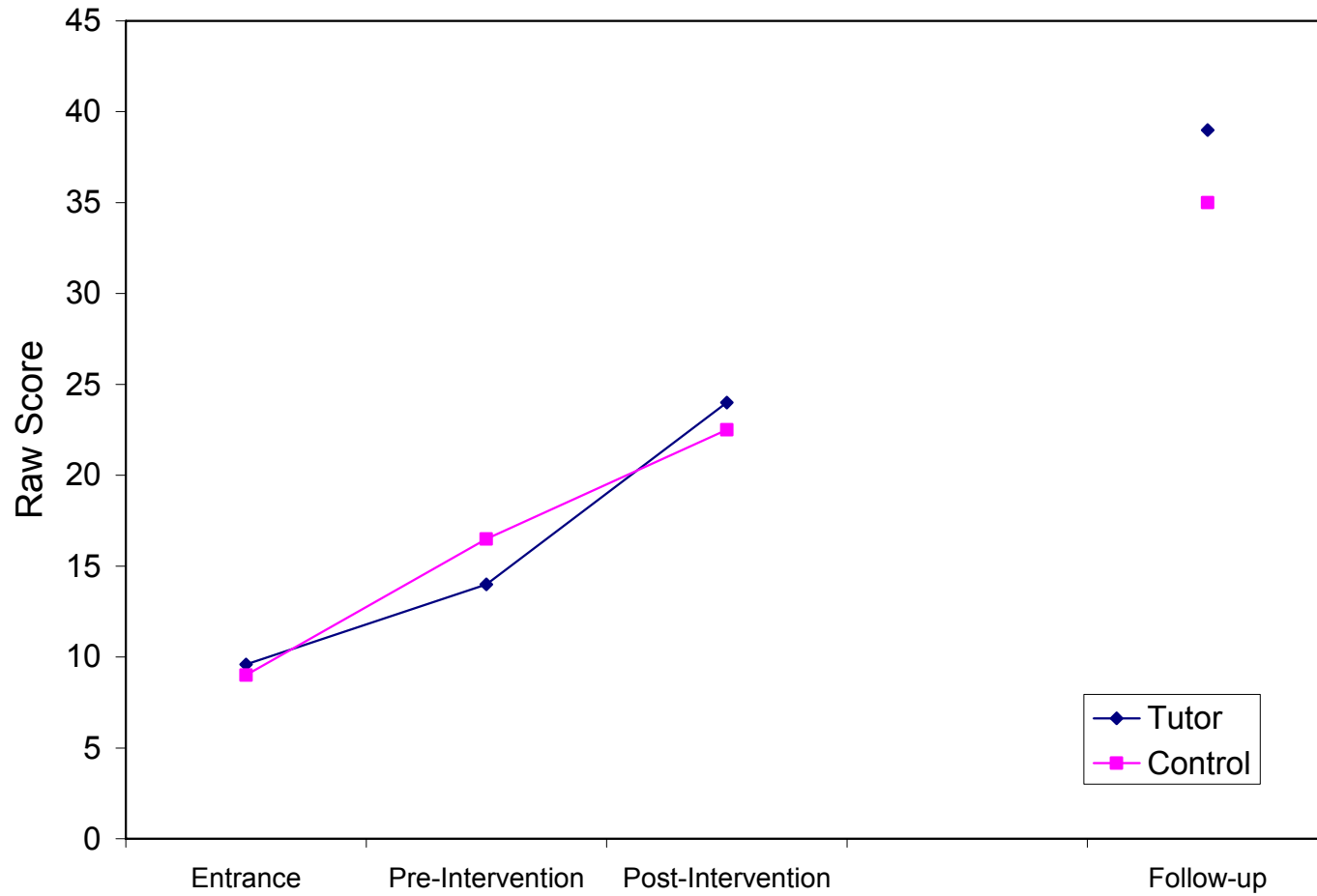
Word Identification Fluency



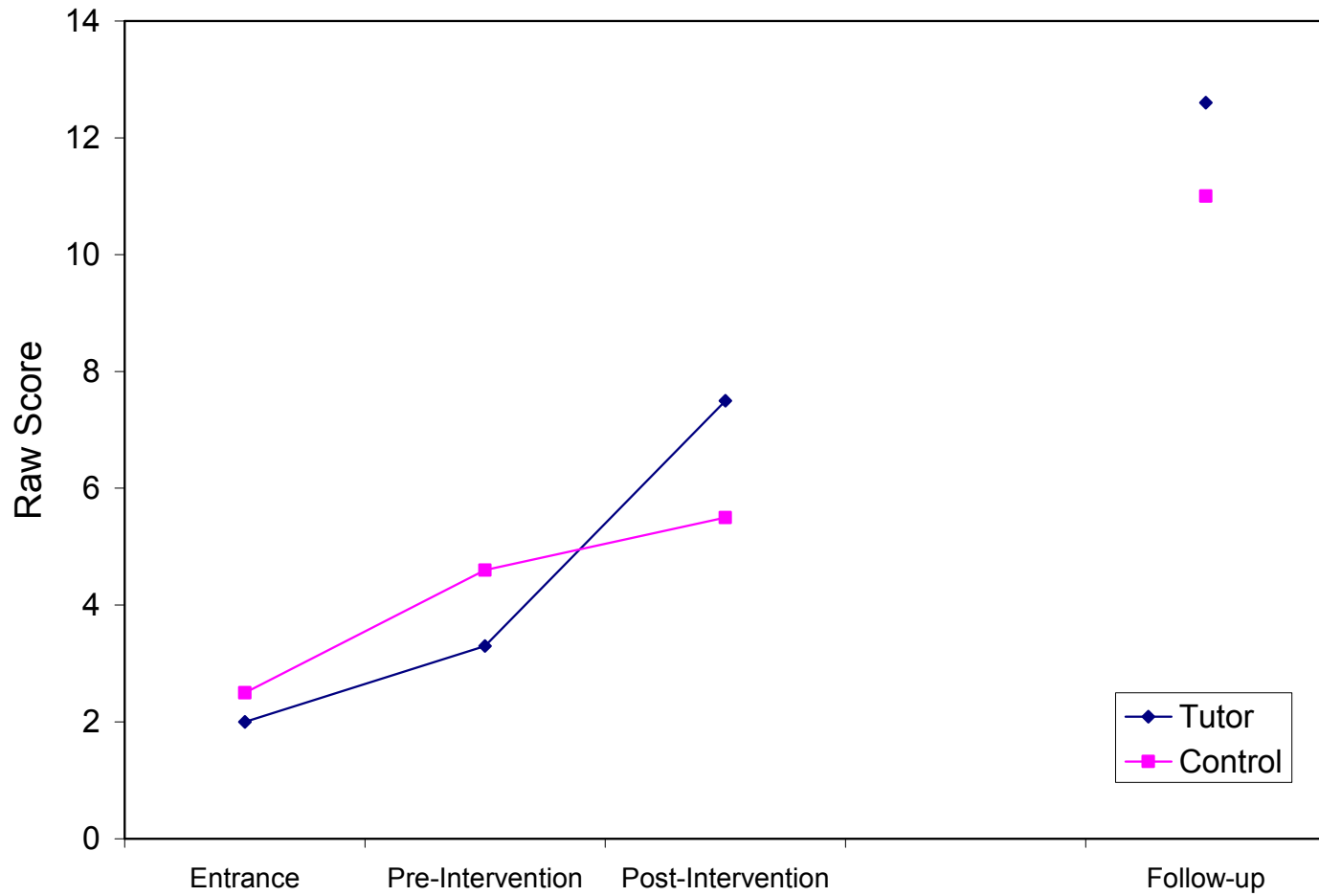
Effects of Tutoring: Standardized Reading Measures

- 2-way ANOVAs
 - Condition (Spring Tutored/Control) as between-subjects factor
 - Time (pretest vs. mid-year; mid-year vs. posttest) as within-subjects factor
 - Outcome measures: WID, WA, Sight Word Efficiency, Phonemic Decoding Efficiency
- For 3 of 4 measures (all but Sight Word Efficiency): interaction between condition and time, whereby
 - Contrast from pretest to mid-year was comparable for Spring Tutored and Control
 - Contrast from mid-year to posttest was significant, with Spring Tutored outperforming Control.

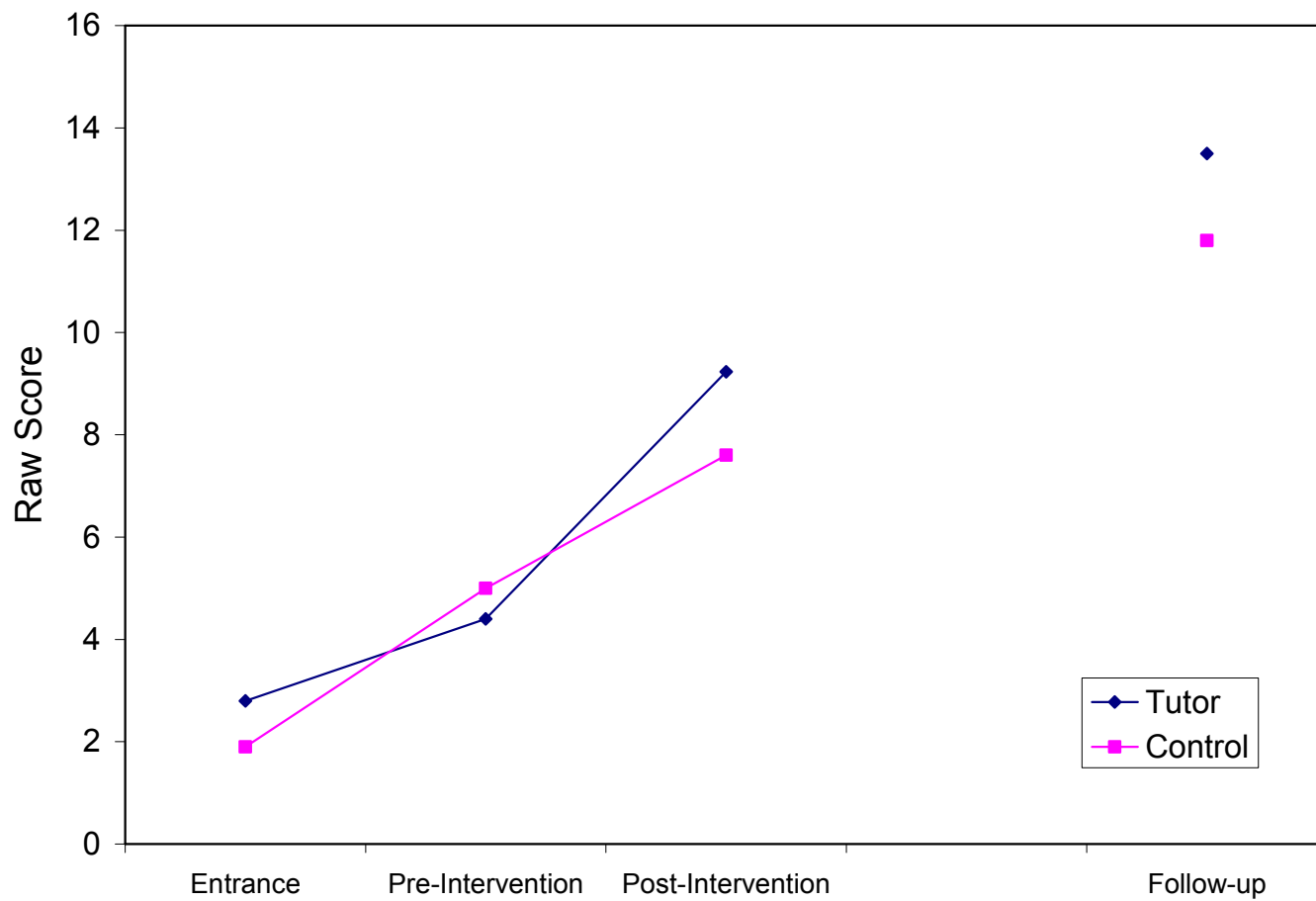
Sight Word Efficiency



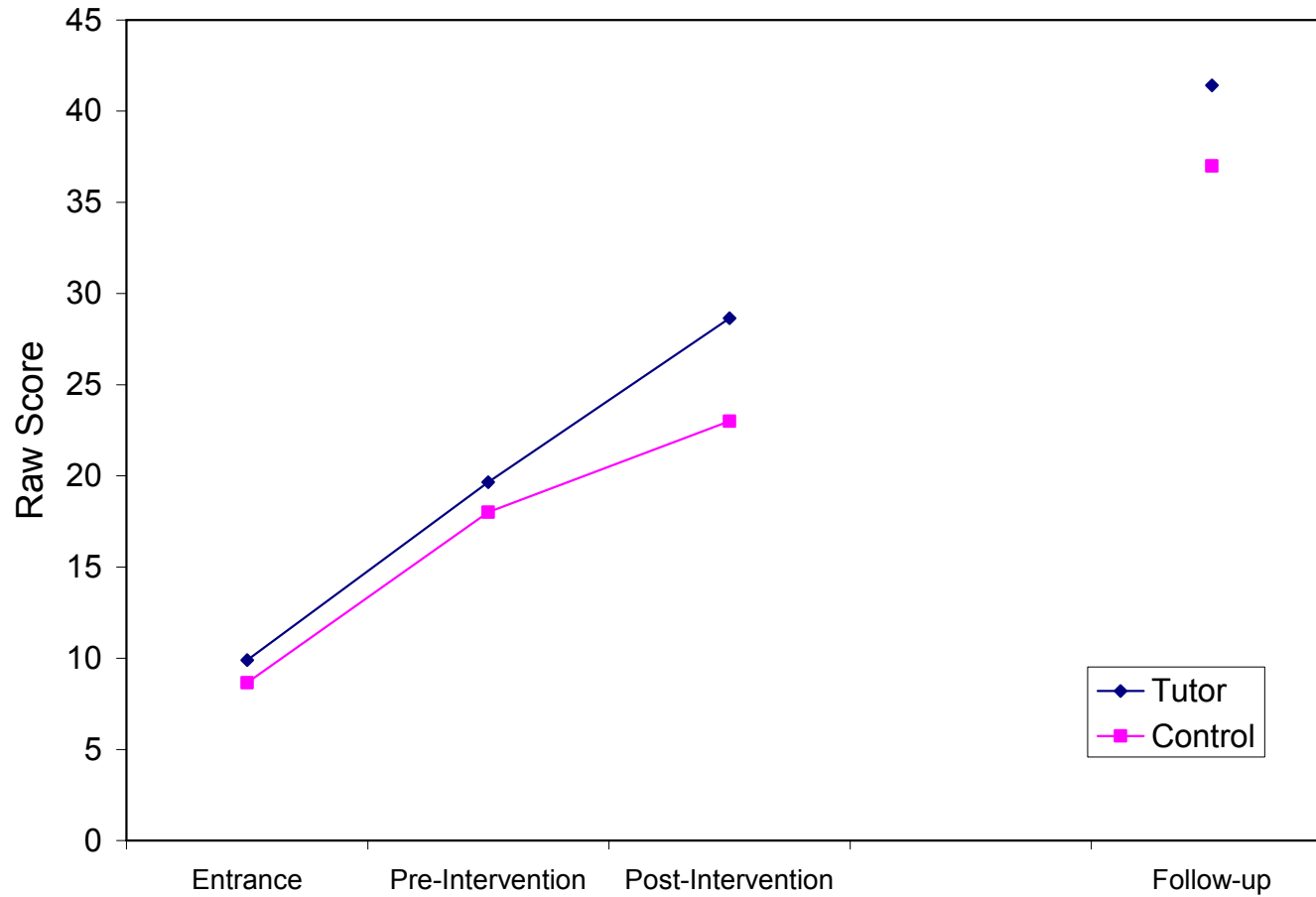
Decoding Efficiency



Word Attack



Word ID



Maintenance of Effects to End of Grade 2

- ANOVAs on posttest at Grade 1 to end of grade 2 (condition; time)
- Main effect for time, but not for condition or interaction; so, effects maintained to end of grade 2

Table 10

Effect Sizes for Tutoring and Control Groups during Grade 1 and One-Year Follow-Up

Measure	Pretest to Posttest	Pretest to Follow-up
Word ID	.34	.28
Word Attack	.28	.40
Sight Word	.45	.41
Phonemic Decode	.55	.42

Note. Word ID = Woodcock Reading Mastery Test – Revised/NU: Word Identification subtest; Word Attack = Woodcock Reading Mastery Test – Revised/NU: Word Attack subtest; Sight Word = Test of Word Reading Efficiency: Sight Word Efficiency subtest; Phonemic Decoding = Test of Word Reading Efficiency: Phonemic Decoding Efficiency subtest; d = Cohen’s d corrected for the correlation between pretest and posttest.

Did tutoring decrease RD prevalence at end of 1st grade?

Defining RD = 1st-grade WID slope $< .75 SD$
below normative mean slope

Yes: RD rates significantly lower in Spring
Tutored (43.5%) than Control (81.8%)

Classification Methods

Questions

- Using longitudinal data (grade 1 to grade 2), what is the sensitivity and specificity of various classification methods and measures and their associated RD prevalence rates?
- What degrees of RD severity are associated with the methods/measures?

Measures, Methods, RD Criterion

- Definitional Methods/Measures in 1st Grade
 - WRMT-WID, TOWRE-SWE, CBM-WIF, CBM-PRT
 - Initial Low Achievement, Discrepancy, Normalization, Benchmark, Slope, Dual Discrepancy
- RD Criterion at End of 2nd Grade
 - Composite Score
 - WRMT-WID and WA (one-third weight)
 - TOWRE-SWE and DE (one-third weight)
 - WRMT-PC (one-third weight)

Criteria for Judging Definitional Methods

1. Sensitivity with respect to end of grade 2 RD ~ .80
2. Specificity with respect to end of grade 2 RD ~ .80
3. Severity ESs (RD vs. non-RD) at end of grade 1 and at end of grade 2 ~ 1.00, across various reading measures (but excluding measure used in definition)
 - Which 1st-grade definitional methods/measures identify RD children one year later, while yielding severe reading deficits and expected prevalence?

Table 3. RD Prevalence (Percent RD), Severity (ES), and Stability (from End of Grade 1 to End of Grade 2) for Alternative Methods of Classifying RD

First-Grade Definition	Percent LD	End Grade 1		End Grade 2		Hit Rate	Sensitivity	Specificity
		Avg ES	ES Range	Avg ES	ES Range			
Initial Low Achievement								
WID (-1SD)	3.29	1.47	0.96 –1.74	1.37	0.98 –1.65	.91	.57	.88
SWE (-1SD)	8.48	0.80	0.60 –1.16	0.70	0.55 –0.87	.78	.70	.80
WIF (-1SD)	6.87	1.24	0.97 –1.43	0.99	0.78 –1.47	.85	.78	.85
Final IQ-Ach Discrepancy								
WID (1SD)	0.79	1.78	1.25 –2.61	1.57	1.14 –1.92	.89	.13	.98
SWE (1SD)	1.12	1.15	0.46 –1.99	0.94	0.65 –1.25	.90	.22	.99
WIF (1SD)	5.08	0.42	0.20 –0.72	0.21	0.01 –0.42	.78	.26	.86
Final Normalization								
WID (90SS)	2.36	1.98	1.45 –2.58	1.95	1.43 –2.57	.94	.56	.99
SWE (90SS)	8.28	1.24	0.99 –1.43	1.17	0.88 –1.31	.83	.87	.82
WIF (90SS)	17.67	1.12	0.88 –1.36	1.04	0.69 –1.54	.52	1.00	.49
Benchmark								
WIF (-1SD)	15.67	1.15	0.90 –1.41	1.04	0.64 –1.60	.63	1.00	.58
PRT (<40)	18.12	1.40	0.98 1.90	1.20	0.91 1.57	.53	1.00	.49
Slope								
WIF Median	16.09	1.10	0.91 –1.36	1.07	0.92 –1.36	.61	1.00	.56
WIF (-1SD)	7.65	1.34	0.88 –1.80	1.18	0.85 –1.54	.83	.83	.84
Slope Split	15.47	1.27	0.95 1.60	1.27	1.01 1.69	.62	1.00	.58
Dual								
WIF (-1SD) ^a	9.31	1.63	1.08 –2.12	1.42	0.95 –2.06	.81	.83	.81
PRT (<40 with WIF Slope - 1SD)	7.34	1.47	0.92 1.95	1.42	0.94 1.99	.82	.75	.83

Notes: ESs were based on WRMT-NU WID, WRMT-NU WA, TOWRE SWE, TOWRE DE, & WRMT-NU Passage Comp (excluding measure used for identification). Hit rate, sensitivity, and specificity were based on children scoring 1 SD below national norm on a composite measure of WRMT-NU WID, TOWRE SWE, & WRMT-NU Passage Comp administered at the end of grade 2. ^aDual model based on an N of 159, whereas Slope models were based on an N of 206.

Methods that Meet Criteria

- Initial Low Achievement using WIF (< -1 SD)
- Normalization using SWE (< 90 SS)
- Slope using WIF (-1 SD)
- Dual Discrepancy using PRT (< 40) and WIF Slope (< -1 SD)

Observations

- Poor sensitivity associated with low prevalence
- Poor specificity associated with high prevalence
- None of the discrepancy options performed well
- Sensitivity, specificity, severity, and prevalence change as a function of which RTI definitional option is used, suggesting the need to become clear on which options “work”
- Could the answer be as simple as poor initial performance on a measure that provides fine discriminations, like WIF?

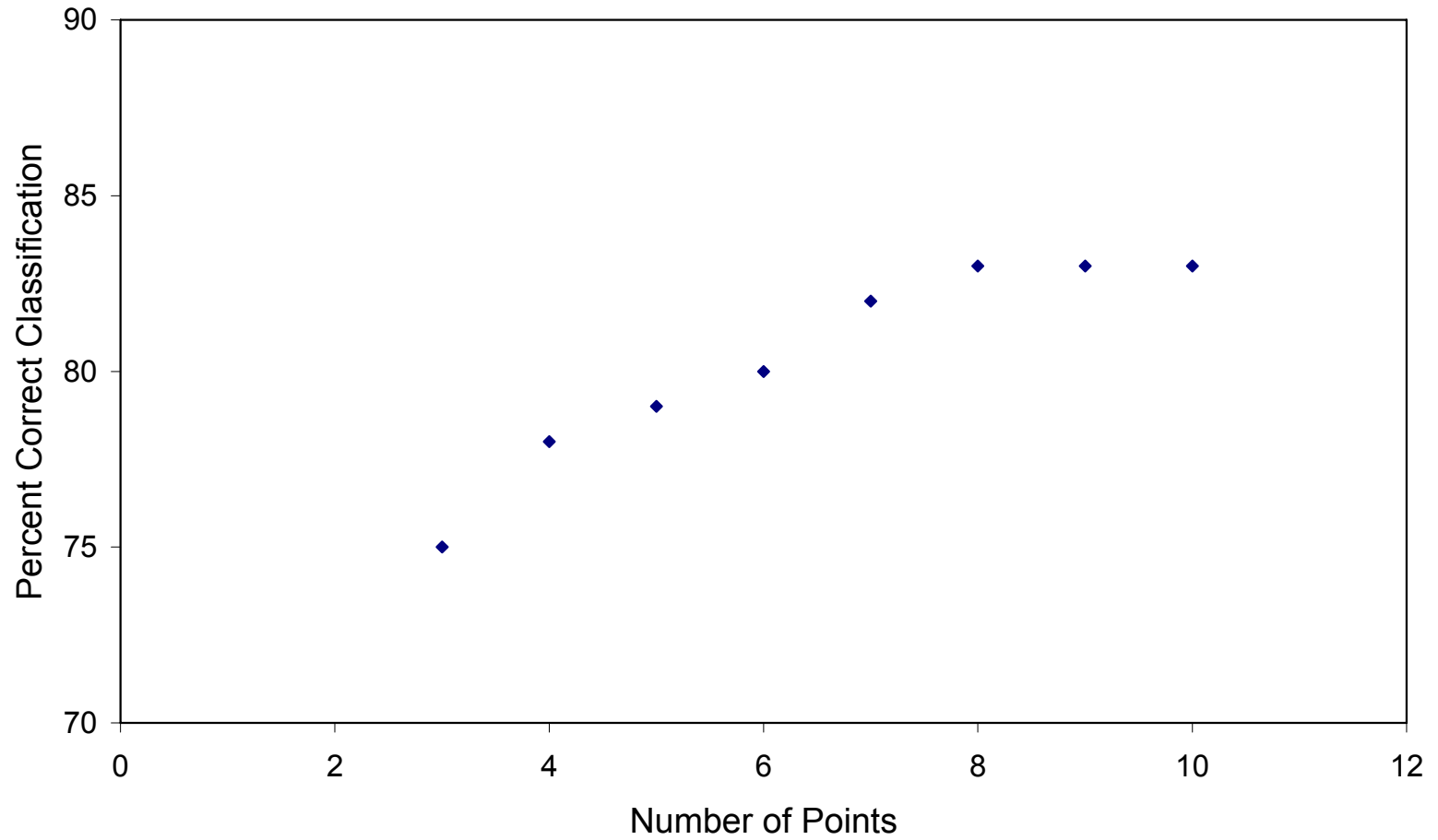
How many data points are
necessary to achieve a reliable
slope?

Table

Correlations Among Slope Terms Based on 3-18 Data Points

Slope Estimates	Slope 9 points	Slope 18 points
Slope 3 points	.31	.19
Slope 4 points	.72	.53
Slope 5 points	.82	.63
Slope 6 points	.89	.68
Slope 7 points	.95	.73
Slope 8 points	.98	.78
Slope 9 points		.82
Slope 10 points		.86
Slope 11 points		.90
Slope 12 points		.93
Slope 13 points		.95
Slope 14 points		.97
Slope 15 points		.98
Slope 16 points		.99
Slope 17 points		.99

Classification Accuracy as a Function of Number of Points



Conclusions

- Prevention: For students who fail to respond to Tier 1 reading instruction in the fall of 1st grade, 9 weeks of tutoring in the spring semester can improve reading outcomes, which are still evident at end of 2nd grade.
- Identification: None of the discrepancy options performed well; Sensitivity, specificity, severity, and prevalence change as a function of RTI definitional option, suggesting the need to become clear on which options “work.”
- Slope reliability: Using WIF, 8-10 data points provides a reliable estimate of longer-term slope.