



# Cognitive Labs to Explore Potential Causes of DIF in Students With and Without Reading-Based Learning Disabilities

Teresa King  
Educational Testing Service



# Overview

- Background
- Purpose
- Procedures
- Findings
  - Item trends
  - Text-based
  - Test-based
  - RLD differences
- Implications and Next Steps



# Background

- Build off the findings of two of last year's research study findings
  - Large scale differential boost study
  - DIF analysis of the GMRT data



# Cognitive Labs

- Cognitive labs using the think aloud method on reading comprehension questions
  - Means of measuring mental processes
- Involves 2 steps
  - Academic task (reading assessment)
  - Verbal reporting task (thinking aloud)



# Cognitive Labs Advantages

- Unique and optimal way to capture otherwise unattainable information
- Beneficial to learn about components of mental processes of reading (Afflerbach & Johnston, 1984, Alavi, 2005; Pressley & Afflerbach, 1995)
- Beneficial in the development of assessments (Caspar, Lessler, & Willis, 1999; Desimone & LeFloch, 2004; Willis, 2005)
- Open flexible procedure can be catered to the specific situation and activity (Davison, Vogel & Coffman, 1997)



# Research Questions

1. What are the differences in the errors on reading comprehension questions for students with and without disabilities?
2. Do students with reading-based disabilities differ in their *explanations* to reading comprehension questions compared to students without disabilities?



# Research Questions

3. What *kind* of information does a cognitive lab provide for future test development for a reading assessment for students with and without reading-based learning disabilities?
4. Is this *type of information gathering and data quality* worthwhile to conduct in future large scale studies?



# Instruments

- Comprehension section of the Gates MacGinitie Reading test
  - Reading passages and multiple choice questions
  - Think aloud protocol for administrator to administer the GMRT
- Three short oral reading fluency measures
- Student survey



# Methodology

- Individual administration
- 1 hour testing session
- Audio recorded sessions
- Test full GMRT passage sets with DIF for non-LD and RLD students



# Think aloud protocol

- Directions were detailed and thorough
  - Test administration began with an example of a think aloud response given to each student
  - Two strategies to answer the same question were given to demonstrate the process of thinking aloud in an effort to deter the potential bias of student responses
- Sufficient practice was encouraged
  - Novel testing method for all students
  - There was concern that students, particularly those with LD, would be less verbal during the reporting task



# Think aloud protocol

Concurrent and retrospective verbal reports were included

- Concurrent reports
  - + More likely to remember mental processes
  - Working memory may be overloaded when too many tasks are performed at once (Afflerbach & Johnston, 1984)
- Retrospective reports
  - + Frees up working memory to perform experimental task (Afflerbach & Johnston, 1984)
  - Retrieval of thoughts allow more room for error (Leighton, 2004)



# Sample

4<sup>th</sup> grade

- RLD = 16
  - M = 10
  - F = 6
- Non-LD = 18
  - M = 9
  - F = 9

8<sup>th</sup> grade

- RLD = 7
  - M = 2
  - F = 5
- Non-LD = 8
  - M = 4
  - F = 4



# 4<sup>th</sup> Grade Basic Findings

## RLD Students

- Percent correct = 55%
- Mean # of items completed = 9
- Mean ORF score = 81
- Mean NJASK score = 219

## NLD Students

- Percent correct = 84%
- Mean # of items completed = 12
- Mean ORF score = 136
- Mean NJASK score = 227



# 8<sup>th</sup> Grade Basic Findings

## RLD Students

- Percent correct = 35%
- Mean # of items completed = 13
- Mean ORF score = 96
- Mean NJASK score = 172

## NLD Students

- Percent correct = 71%
- Mean # of items completed = 14
- Mean ORF score = 161
- Mean NJASK score = 226



# Grade 4 Test Forms

Item	Passage	Form A	Form B	RLD Percent Answered	NLD Percent Answered
1	1	17	2	50%	61%
2		17	2	50%	61%
3		17	2	50%	61%
4		17	2	50%	61%
5	2	16	6	50%	78%
6		15	6	44%	78%
7		15	6	44%	78%
8		15	6	44%	78%
9	3	8	16	63%	78%
10		7	16	63%	72%
11		6	16	56%	72%
12		6	16	50%	78%
13	4	3	17	50%	67%
14		3	17	50%	67%
15		3	17	50%	67%
16		2	17	50%	67%
17		2	17	50%	61%
18		2	17	50%	61%

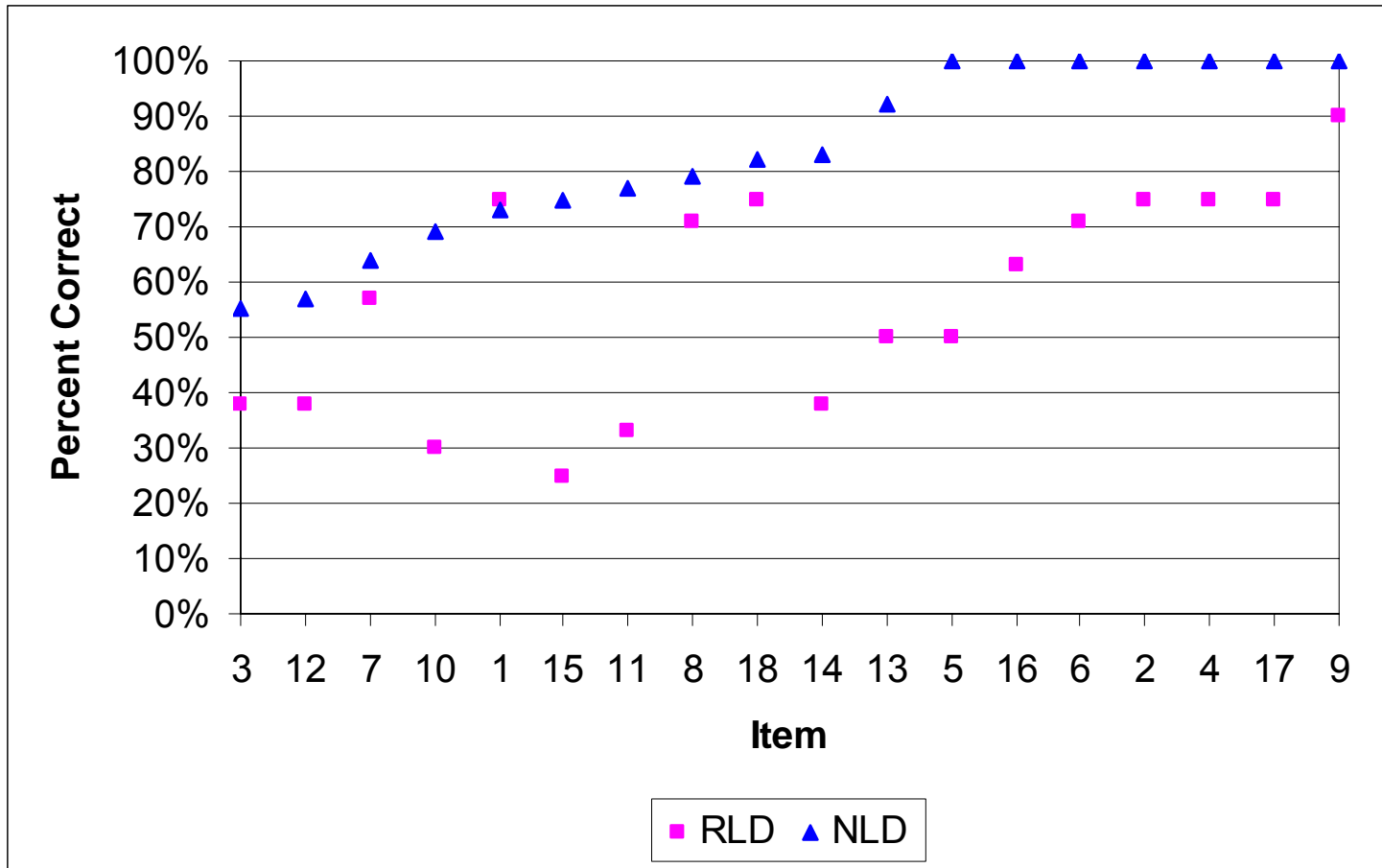


# Grade 8 Test Forms

Item	Passage	Form A	Form B	RLD Percent Answered	NLD Percent Answered
1	1	7	4	57%	88%
2		7	4	57%	88%
3		7	4	57%	88%
4	2	7	7	100%	88%
5		7	7	100%	88%
6		7	7	100%	88%
7		7	7	100%	88%
8		7	7	100%	88%
9	3	7	8	100%	100%
10		7	8	100%	100%
11		7	8	100%	100%
12		7	8	100%	100%
13	4	5	8	71%	100%
14		5	8	71%	100%
15		5	8	71%	100%



# 4<sup>th</sup> Grade Percent Correct by Item

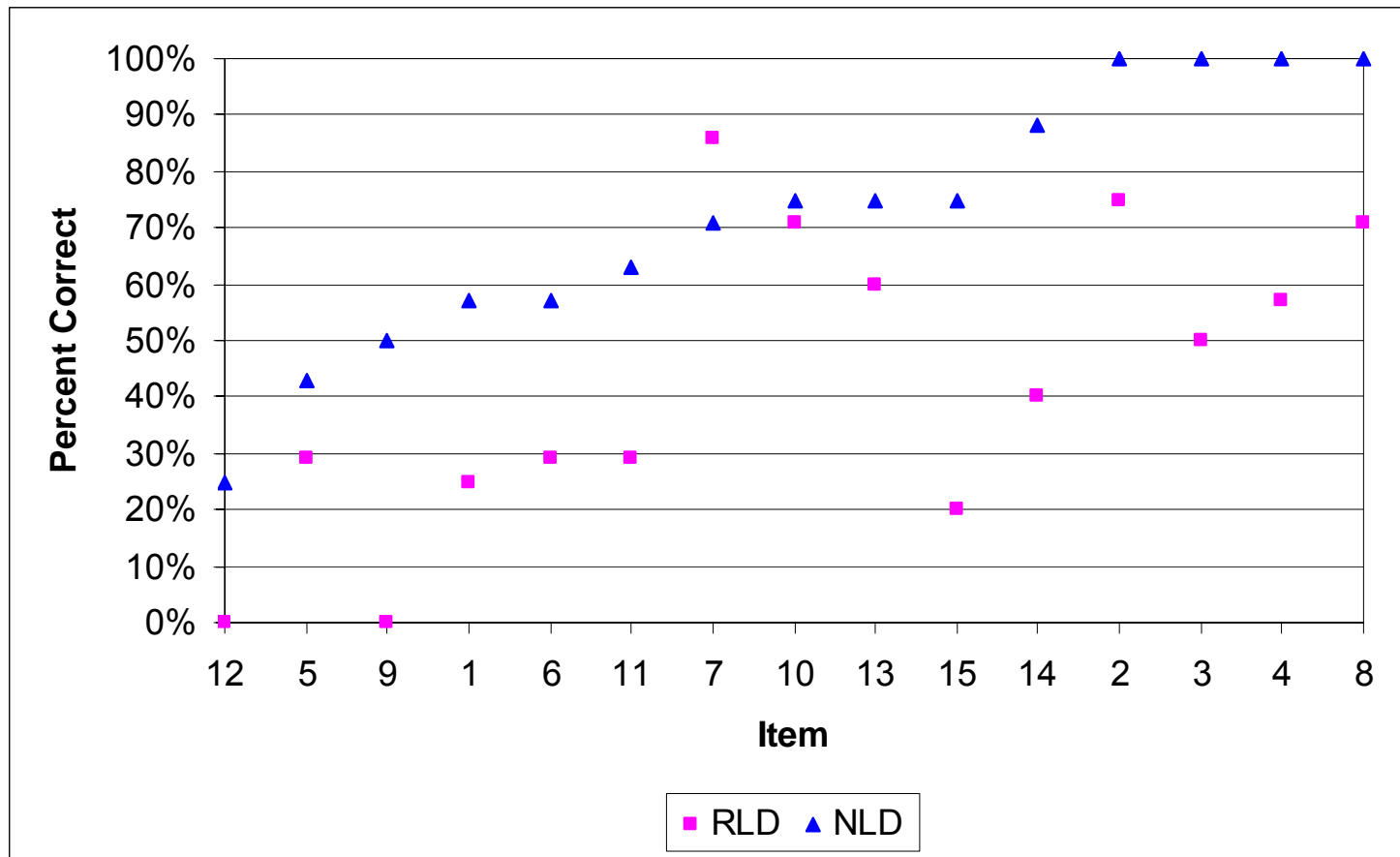




# 4<sup>th</sup> Grade DIF Items

Item	RLD Percent Answered Correctly	NLD Percent Answered Correctly	Difference
3	38%	55%	17%
12	38%	57%	19%
7	57%	64%	7%
10	30%	69%	39%
1	75%	73%	-2%
15	25%	75%	50%
11	33%	77%	44%
8	71%	79%	8%
18	75%	82%	7%
14	38%	83%	45%
13	50%	92%	42%
2	75%	100%	25%
4	75%	100%	25%
5	50%	100%	50%
6	71%	100%	29%
9	90%	100%	10%
16	63%	100%	37%
17	75%	100%	25%

# 8<sup>th</sup> Grade Percent Correct by Item





# 8<sup>th</sup> Grade DIF Items

Item	RLD Percent Answered Correctly	NLD Percent Answered Correctly	Difference
12	0%	25%	25%
5	29%	43%	14%
9	0%	50%	50%
1	25%	57%	32%
6	29%	57%	28%
11	29%	63%	34%
7	86%	71%	-15%
15	20%	75%	55%
13	60%	75%	15%
10	71%	75%	4%
14	40%	88%	48%
3	50%	100%	50%
4	57%	100%	43%
8	71%	100%	29%
2	75%	100%	25%



# Coding Schema

Errors on reading comprehension test can occur for many reasons:

- Problems with issues related to an aspect of the test
- Inability or inaccurate understanding of a question and/or the answer options
- Presence of a learning disability



# Text-Based Findings

- Vocabulary issues
- Subject issues
- Understanding/comprehension
  - Accuracy
- Text layout
  - Images and pictures
  - Item placement on the page



# Test-Based Findings

- Test taking strategies
  - Crossing off
  - Reading passage aloud
  - Rereading passage to answer questions
  - Guessing
  - Monitoring meaning



# RLD Explanation Trends

- There was a pattern of students getting test questions correct using the incorrect rationale
- LD students were much more likely to use outside knowledge to help support test responses (that didn't necessarily apply)
- LD students were less likely to refer back to the passage to select an answer



# Implications

- Complements the findings of the previous DARA research
- This methodology is helpful to gain insight on the specific errors students make given a *component approach* to measuring reading
  - When there was an error, it was easy to identify the problem.



# Implications

- This procedure will be useful to test new item types and specific subject areas on target populations (LD, VI, blind, hearing impaired, etc.)
- This is a good means to examine non-content related aspects of text and testing such as motivation, interest level, and prior knowledge



# Next Steps

- Examine variables not yet analyzed including:
  - Follow up questions examining confidence
  - Students' perception of item and passage difficulty
- Work with test developers to create guidelines of factors to avoid that were found to be problematic such as:
  - Effects of prior knowledge
- Revise protocol for use in future research