

Examining Differential Item Functioning for a Reading Test Administered With a Read Aloud Modification

Linda Cook, Mary Pitoniak, Cara Cahalan-Laitusis
and Fred Cline

Educational Testing Service

Data Collection

- Data collected for large experimentally designed study
- All NJ public school 4th and 8th grades were invited to participate
- Samples
 - Fourth grade
 - 1170 students
 - 522 (RLD)
 - 648 (NLD)
 - Eighth Grade
 - 855 students
 - 394 (RLD)
 - 461 (NLD)

Data Collection

- Two forms of a reading comprehension test
 - Form S
 - Form T
- Two grades
 - Grade 4
 - Grade 8
- All forms and grade levels contained 48 passage based items
- All items were multiple-choice

Data Collection Design

Group	Session 1		Session 2	
	Form	Modification	Form	Modification
1	S	Standard	T	Audio (read aloud)
2	S	Audio (read aloud)	T	Standard
3	T	Standard	S	Audio (read aloud)
4	T	Audio (read aloud)	S	Standard

Differential Item Functioning (DIF) Analyses

- Comparisons of item responses for 4 groups
 - Reading based learning disability (RLD), no accommodation (Standard)
 - Reading based learning disability (RLD), audio modification (Audio)
 - No disability (NLD), no accommodation (Standard)
 - No disability (NLD), audio modification (Audio)

Summary Statistics for Grade 4 Form S DIF Study Samples

Group	N	Mean	SD
RLD Standard	258	19.18	9.051
RLD Audio	269	24.36	8.808
NLD Standard	326	30.08	9.680
NLD Audio	328	32.44	8.806

Comparison Groups for DIF Analyses

Reference Group	Focal Group		
	RLD Standard	RLD Audio	NLD Audio
RLD Standard	--	1	--
NLD Standard	3	4	2

Comparison Groups for DIF Analyses

	Focal Group		
Reference Group	RLD Standard	RLD Audio	NLD Audio
RLD Standard	--	1	--
NLD Standard	3	4	2

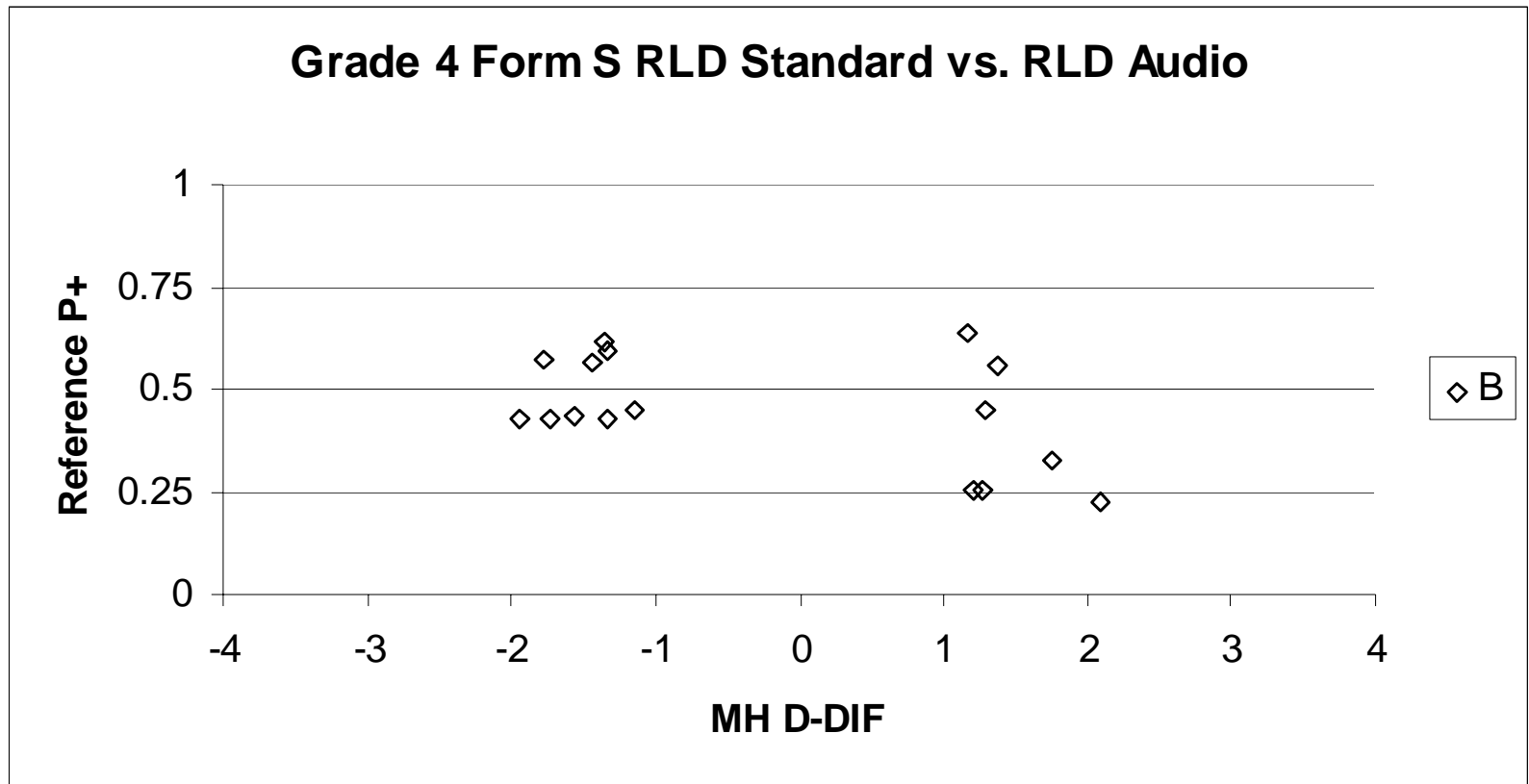
Comparison Groups for DIF Analyses

	Focal Group		
Reference Group	RLD Standard	RLD Audio	NLD Audio
RLD Standard	--	1	--
NLD Standard	3	4	2

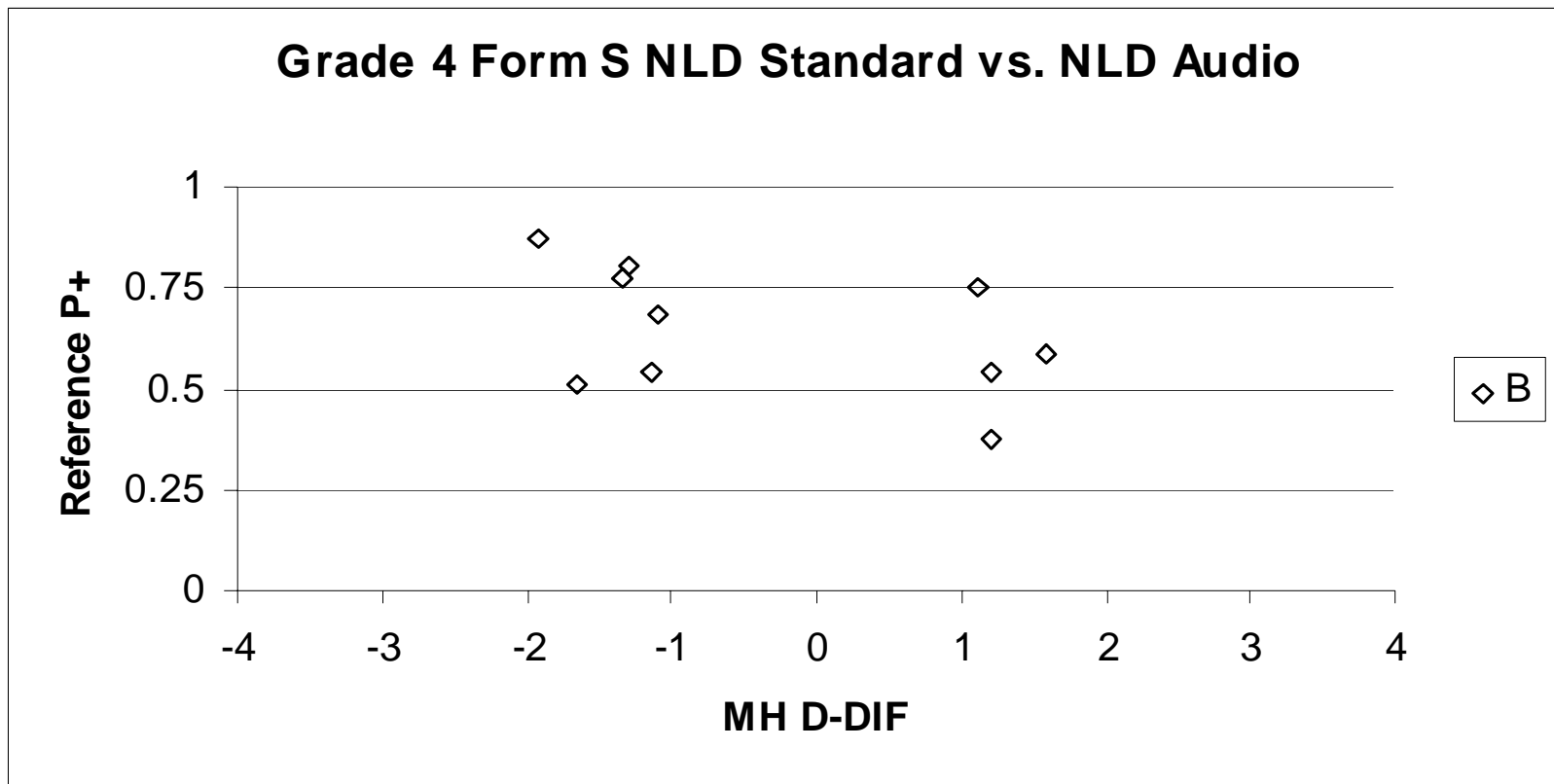
Differential Item Functioning (DIF) Analyses

- Used Mantel-Haenszel procedure with total score as criterion
- Mantel-Haenszel categorization
 - A—negligible DIF
 - B—slight to moderate DIF
 - C—moderate to large DIF
- Direction of DIF Flags
 - Negative favors reference group
 - Positive favors focal group

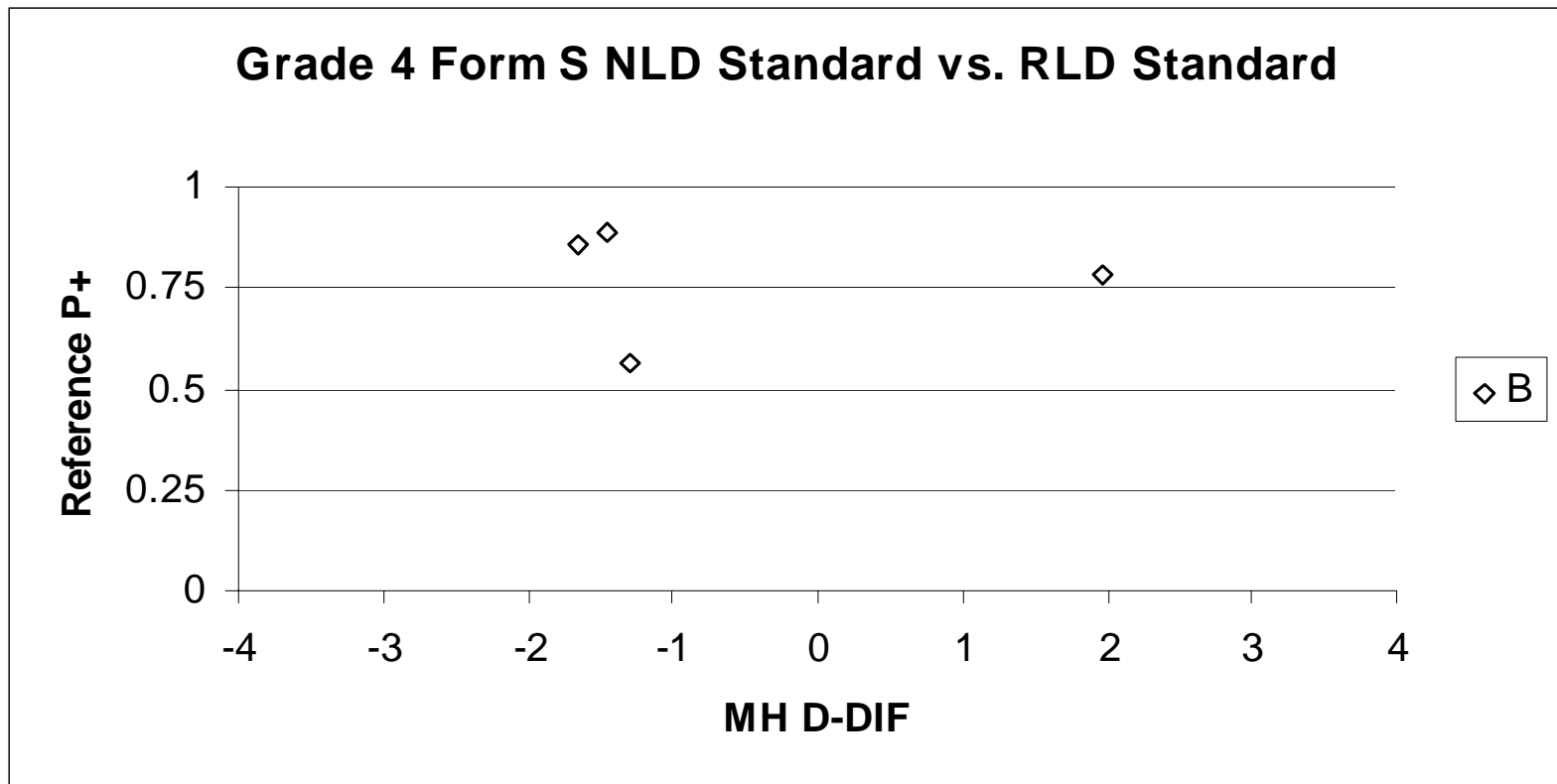
Comparison 1: RLD Standard vs. RLD Audio



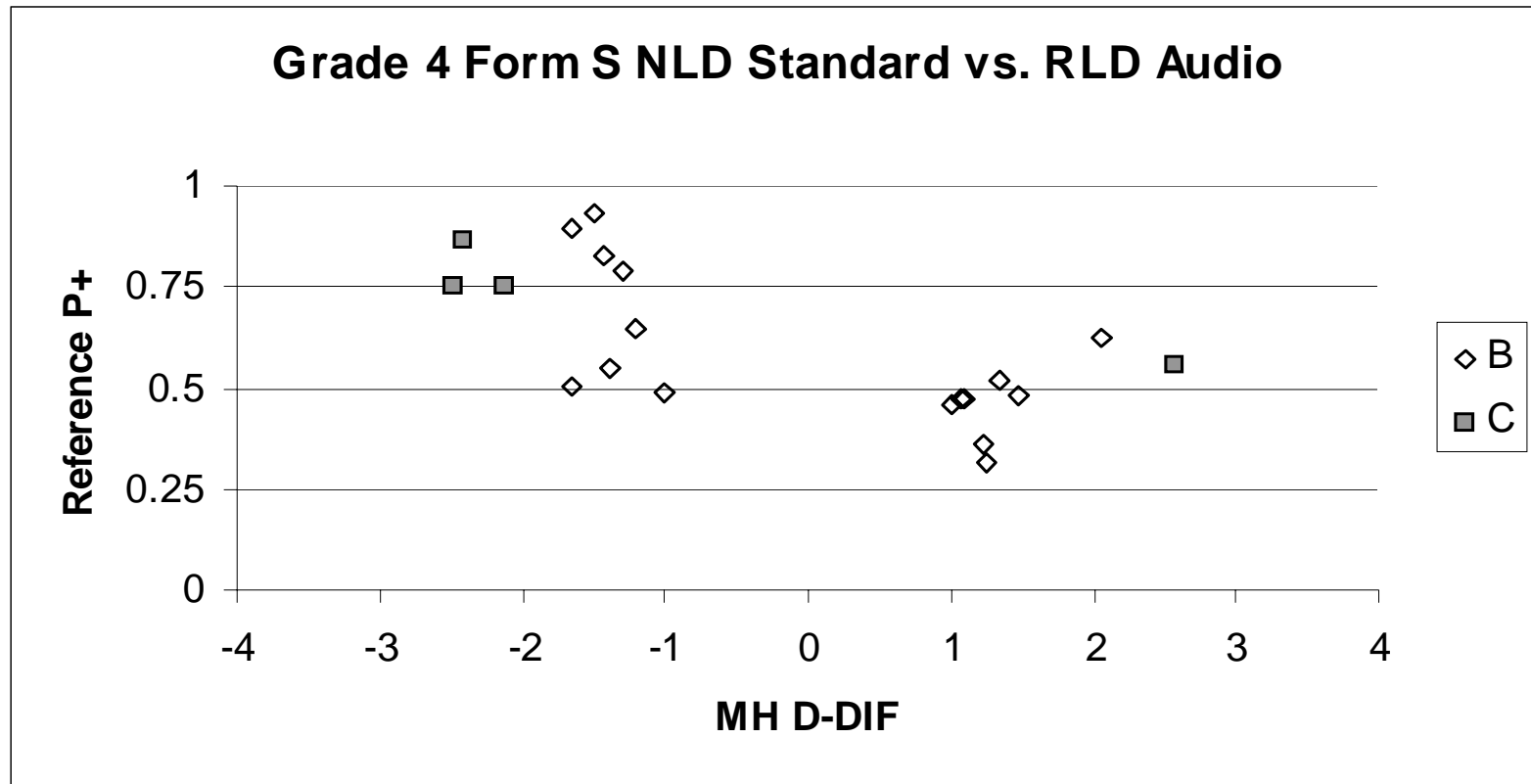
Comparison 2: NLD Standard vs. NLD Audio



Comparison 3: NLD Standard vs. RLD Standard



Comparison 4: NLD Standard vs. RLD Audio



Summary of DIF Results

- Comparisons focusing on mode of administration showed little DIF
- Comparisons focusing on disability status showed little DIF
- Comparisons that crossed mode of administration and disability status showed substantial DIF
 - Positive DIF—unusual proper names
 - Negative DIF
 - **Difficult inferences**
 - **Unusual prose style**
 - **Distracting art work**
 - **Words with multiple meanings**