



Designing Accessible Reading Assessments (DARA) A Summary of Research Cara Cahalan Laitusis Educational Testing Service



Designing Accessible Reading Assessments (DARA)

- Educational Testing Service (ETS)
- Focus on students with learning disabilities
- Focus on component approach to assessing reading skills. Primary areas are:
 - Word Recognition
 - Reading Fluency
 - Vocabulary Knowledge
 - Comprehension



Primary Questions for Year 2

- Can comprehension be assessed in audio format if word recognition and fluency are assessed separately?
 - Are listening comprehension and reading comprehension similar constructs (highly correlated) in proficient readers?
 - Do students with reading-based learning disabilities receive differential performance gains from read aloud?
 - Do tests and test items taken with and without read aloud perform the same psychometrically (same factor structure, no evidence of differential item performance)?



Year 2 Research

- Differential Boost from Read Aloud on Reading Test
- Psychometric Studies of ELA test
 - Differential Item Function
 - Differential Distractor Analysis
 - Factor Analysis



Differential Boost Study

- Does a read-aloud accommodation provide a differential boost for students with learning disabilities?
- How well do test scores predict teacher's ratings of a student's reading comprehension level?
- Are teachers able to predict which students will benefit from a read-aloud accommodation?



Differential Boost Description

- Framework for studying impact of changes in testing conditions
- Change in testing condition may be viewed as an accommodation if students with disabilities receive a significant increase in scores compared to students without disabilities
- Most researchers refer to differential boost as an interaction between group and change in testing condition



Differential Boost Study

Data Collected

Primary Measure

- 2 Reading Comprehension Tests (Form S and Form T)
 - Extra time
 - Extra time with Read Aloud via CD

Additional Measures

- 2 Fluency Measures
- 2 Decoding Measures (4th grade only)
- Student Survey
- Teacher Survey



Sample

- 1181 4th grade students
 - 527 students with reading based learning disabilities (RLD)
 - 654 students without a disability (NLD)
- 855 8th Graders
 - 394 RLD
 - 461 NLD



Differential Boost Data Collection Design

Group	Session 1		Session 2	
	Form	Accommodation/ Modification	Form	Accommodation/ Modification
1	S	Standard	T	Audio
2	S	Audio	T	Standard
3	T	Standard	S	Audio
4	T	Audio	S	Standard



Grade 4 Differential Boost Analysis

Differential Boost Study: Grade 4 Summary Statistics for GMRT						
	Non-LD			RLD		
Condition	N	Mean	SD	N	Mean	SD
Audio	654	502	33	527	477	30
Standard	654	497	38	527	457	32
Boost	654	5	24	527	20	29



Results of the Differential Boost Study

- Students with reading-based learning disabilities had a significantly larger boost from audio (read aloud) accommodations than students without disabilities
 - Controlling for other factors (e.g., reading fluency, decoding, or ceiling effects) does not change these findings



Additional Findings on how to predict who might benefit from audio

- Teachers were not good predictors of which students would benefit from read aloud accommodation
- Boost from audio is more strongly correlated with the standard (non-audio) score than other measures:
 - Standard-Boost correlations range from $-.42$ to $-.51$
 - Fluency-Boost correlations range from $-.03$ to $-.25$
 - Decoding-Boost correlations range from $-.22$ to $-.28$
 - Audio-Boost correlations range from $+.14$ to $+.41$



Additional Findings on Validity of Scores

- Teachers ratings of “reading comprehension” are more highly correlated with the standard (non-audio) score than the audio score for grade 4 (RLD and NLD) and grade 8 RLD, correlations were similar for grade 8 NLD
- Standard (non-audio) and Audio scores are more highly correlated in students without disabilities
 - Standard-Audio correlations were .78 and .79 for NLD
 - Standard-Audio correlations were .56 and .65 for RLD
 - Standard-Standard correlations reported by test publisher are .85



Additional Studies

- Factor Analysis of State ELA assessment (grade 4 and 8)
- Differential Item Functioning
 - state ELA assessments (grades 4 and 8)
 - GMRT (grades 4 and 7/9)
- Differential Distractor Functioning
 - state ELA assessment (grade 4)



Results of the Factor Analysis

Question: Is the internal structure of grade 4 and 8 ELA assessments similar for these 4 groups:

- Students without disabilities
- Students with disabilities (no accommodations or modifications)
- Students with disabilities (504/IEP accommodations)
- Students with disabilities (read-aloud modification)

Answer: Results suggest test measures same single dimension for all groups



Differential Item Functioning (DIF)

- Comparison of item difficulty on two samples of students with and without learning disabilities matched on total test score
- Test we examined were:
 - GMRT (grade 4 and 7/9)
 - State ELA Assessment (grade 4 and 8)



Differential Item Functioning (DIF)

The purpose of the DIF studies was to

- examine whether or not the assessments measured the same construct(s) for the groups in each study
- Examine if some standards (e.g., reading or writing) performed differently for different subgroups
- Identify characteristics of test questions that performed differently for different subgroups (e.g., are longer reading passages differentially more difficult for students who use read aloud)



Summary of DIF Results

- Our a priori hypotheses about which test questions would exhibit DIF (and why) were not successful
- 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 in the experimental study
 - very little DIF in the state assessment study



Answers to Questions for Year 2

Question 1: Are listening comprehension and reading comprehension similar constructs (highly correlated) in proficient readers?

Answer: Yes

Question 2: Do students with reading-based learning disabilities receive differential performance gains from read aloud

Answer: Yes

Question 3: Do tests and test items taken with and without read aloud perform the same psychometrically (same factor structure, no evidence of differential item performance)?

Answer: Sometimes



Primary Questions for Year 3

- Can we explain why some items function differently
 - for students with learning disabilities?
 - when a read aloud accommodation is used?
- Can we design a multi-stage reading assessment that assesses different components of reading (decoding, fluency, comprehension of text) in isolation (and simultaneously) and report scores on the same scale?
 - Test 1: Audio comprehension, decoding, oral reading fluency
 - Test 2: Comprehension without read aloud



Year 3 Research

- Think Aloud studies with LD and non-LD students to examine how students approach
 - items shown to have DIF
 - new item types designed to assess fluency and word recognition in a large scale assessment
 - Families of items with slight variations (e.g., operational item and universally designed items)
- Simulation studies using real data to examine test reliability and comparability on multi-stage reading assessment.