

**Technical Report # 28**

**Analysis of Reading Fluency and Comprehension Measures for Fifth  
Grade Students**

Ron Beghetto

Gerald Tindal

University of Oregon



behavioral research & teaching

Published by

Behavioral Research and Teaching  
University of Oregon • 175 Education  
5262 University of Oregon • Eugene, OR 97403-5262  
Phone: 541-346-3535 • Fax: 541-346-5689  
<http://brt.uoregon.edu>

Copyright © 2004. Behavioral Research and Teaching. All rights reserved. This publication, or parts thereof, may not be used or reproduced in any manner without written permission.

The University of Oregon is committed to the policy that all persons shall have equal access to its programs, facilities, and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation. This document is available in alternative formats upon request.

## *Introduction*

In response to the No Child Left Behind Act of 2001, school districts are working to develop assessment systems to monitor student progress. In the area of reading, three measures can provide useful information about students' developing proficiency: a test of oral reading fluency (ORF), a vocabulary test, and a reading comprehension test comprised of selection response and constructed response items. To be most useful at the district level, it is helpful to have a variety of comparable forms available for each of these measures to allow multiple testing during the school year.

This report details an analysis conducted on the fifth-grade version of District Reading Tests (Oral Reading Fluency, Vocabulary, and Reading Comprehension). The analysis had the purpose of: (a) examining the comparability between forms; (b) examining group performance, item difficulty, and item discrimination power; and (c) abbreviating two forms of selection and constructed response items. First, we describe the methods of the analysis and then we present results before closing with a brief conclusion.

## *Methods*

### *Setting and Subjects*

This report summarizes the spring 2003 reading achievement data for fifth-grade students ( $N=1443$ ) from 29 schools representing a school district in a mid-sized city in the Pacific Northwest.

### *Design and Operational Procedures*

Three databases of student reading achievement scores were used in this analysis: (a) District Reading Comprehension, (b) District Vocabulary, and (c) state reading achievement. Prior to merging the three databases, formulas and scoring for Reading Comprehension and

Vocabulary were verified by randomly selecting cases and testing the formulas used to compute total scores. A copy of both forms of the Reading Comprehension Test (with scores) was examined and the scoring formula used in the database was verified. No copy of the constructed response scoring rubric or the Vocabulary Test was available for this analysis. Therefore, the accuracy of the scoring formula for Constructed Responses and Vocabulary was not verified other than checking the accuracy of the calculation of total and percent correct.

Three data files were merged on common student identification numbers (ID). Other relevant data (e.g., student names) were used to match the few records missing IDs. The combined database included students who had multiple records. In cases such as this, multiple records were combined by deleting empty cells or creating average scores for students who had multiple scores on the state reading test.

Dependent variables analyzed in this report include scores from the following measures: Oral Reading Fluency (ORF) ( $n = 1376$ ), District Vocabulary Test ( $n = 1369$ ), and District Reading Comprehension Test ( $n = 1392$ ). Prior to analysis, schools in the district were coded into two regions that roughly corresponded with income level. Independent variables analyzed in this report include income level, gender, ethnicity, and Special Education (SPED) and English Language Learner (ELL) designation.

#### *Measurement/Instrument Development*

##### *Oral Reading Fluency (ORF)*

The test of Oral Reading Fluency was administered individually to each student by trained assessors. Students read aloud for exactly one minute one of two comparable passages deemed grade-level appropriate on the Flesch-Kincaid reading scale. At the end of one minute,

assessors marked the last word read then counted the total words read as well as any words read incorrectly to arrive at a final ORF score.

### *Vocabulary*

Students were administered a 25-word, multiple-choice Vocabulary Test. Each item on the test consisted of one correct answer and two distracters. Students bubbled in their answers on the form itself and all tests were machine scored.

### *Reading Comprehension*

Students were administered one of two Reading Comprehension Tests (Form A or Form B). Form A was based on the fiction passage entitled “The Stag.” Form B was based on the non-fiction passage entitled “Bessie Coleman.” Each form of the Reading Comprehension Test consisted of a reading passage followed by multiple-choice as well as constructed response questions. Multiple-choice selection response (SR) questions were machine scored while constructed response (CR) questions were all scored by the same scorer using scoring guides provided by the district. Scorers were trained by two district administrators who also checked every fifth paper to ensure that their scores were consistent with district expectations. Responses for which the scorer was unable to decide on an appropriate score were discussed with both trainers before having a final score assigned.

### *Data Preparation and Analysis*

*Comparable forms.* Comparability of the two forms of District ORF and Reading Comprehension were examined using Analysis of Variance (ANOVA).

*Group performance.* Group performance on District Reading Comprehension and District Vocabulary measures was examined using multiple ANOVAs. Type I Error in post hoc comparisons was controlled using Tukey’s Honestly Significant Difference adjustment.

*Item analysis.* Item analyses were conducted for the District Reading Vocabulary Test and both forms of the District Reading Comprehension Test. First, the top and bottom 27% of examinees were classified based on total score. Responses for top and bottom 27% of examinees were tabulated for each item on each test (i.e., percent correct for each item, percent of students selecting distracters). Tabulated item data were used to determine: (a) performance of each distracter, (b) item difficulty (total correct divided by total number of top and bottom 27% of examinees), and (c) item discriminating power (using item-total score correlations of the entire dataset). Item difficulty is reported as a percentage. The percentage represents how many students, in the upper and lower 27%, correctly answered the item. The higher the percentage, the easier the item.

Item discrimination is reported as a correlation coefficient, ranging from .00 to 1.0. The closer to 1, the greater the item discriminates. Higher item discrimination coefficients indicate that the item is closely aligned with the total score (based on data from all examinees). Poorer functioning items will have lower item discrimination coefficients (approaching zero).

### *Results*

#### *Oral Reading Fluency (ORF)*

*Comparability of forms.* Oral Reading Fluency (ORF) did not significantly differ between Form A ( $M=130$ ,  $SD=36$ ) and Form B ( $M=130$ ,  $SD=33$ ) of the fifth-grade District Reading Comprehension Test,  $F(1, 1374)=0.149$ ,  $p =.700$ .

*Group performance.* Analysis of variance revealed significant differences between groups on ORF scores (across both forms). Females outperformed males, Asian students outperformed African American and Hispanic students, regular education students outperformed SPED students, and students from high income regions outperformed students from low income

regions. Non-parametric analysis (Mann-Whitney U) indicated non-ELL students significantly outperformed ELL students ( $Z=2.645$ ,  $p = .008$ ). Descriptive statistics and results of ANOVAs are depicted in tables 1 and 2.

Table 1

*Descriptive Statistics for Grade 5 District ORF Test*

	Group	<i>n</i>	<i>M</i>	<i>SD</i>
Gender	Male	693	128	33
	Female	683	132	36
Ethnicity	White	959	132	33
	Hispanic	51	119	36
	African American	30	120	35
	Asian	70	142	36
	Native American	33	122	40
	Other	58	134	29
SPED	SPED	168	99	36
	Non-SPED	1208	134	32
ELL	ELL	6	99	10
	Non-ELL	1370	130	34
Income	Low	807	126	33
	High	569	134	35
Total		1376	130	34

Table 2

*Analysis of Variance Summary Table for Grade 5 District ORF Test*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>P</i>
Gender	1 1374	5.54*	.004	.019
Ethnicity	5 1124	4.01**	.016	.001
SPED	1 1374	172.96**	.112	0.00
Income	1 1374	20.09**	.014	.000

\* $p < .05$ , \*\* $p < .01$ .

*District Vocabulary Test*

*Group performance.* Analysis of variance revealed significant differences between groups on the District Vocabulary scores. White and “other” students outperformed Hispanic students, regular education students outperformed SPED students, and students from high income regions outperformed students from low income regions. Non-parametric analysis (Mann-Whitney U) indicated non-ELL students significantly outperformed ELL students ( $Z=3.773$ ,  $p < .000$ ). No significant gender differences in Vocabulary scores were found.

Descriptive statistics and results of ANOVAs are depicted in tables 3 and 4.



Table 3

*Descriptive Statistics for Grade 5 District Vocabulary Test*

	Group	<i>N</i>	<i>M</i>	<i>SD</i>
Gender	Male	686	22	4
	Female	683	22	4
Ethnicity	White	956	23	3
	Hispanic	52	20	4
	African American	31	21	4
	Asian	69	22	4
	Native American	32	21	4
	Other	57	22	4
	SPED	SPED	171	19
	Non-Sped	1198	22	3
ELL	ELL	7	15	6
	Non-ELL	1362	22	4
Income	Low	799	22	4
	High	570	23	4
Total		1369	22	4

Table 4

*Analysis of Variance Summary Table for Grade 5 District Vocabulary Test*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>P</i>
Gender	1	1.39	.001	.239
	137	(45197.17)		
Ethnicity	5	5.73**	.023	.000
	14728.34			
SPED	1	139**	.092	.000
	1367	(19217.92)		
Income	1	15.27**	.011	.000
	1367			

\* $p < .05$ , \*\* $p < .01$ .

*Item analysis.* The upper 27% of examinees ( $n=370$ ) scored between 25 and 26 ( $M=25$ ,  $SD=.50$ ) and the lower 27% of examinees ( $n=370$ ) scored between 3 and 21 ( $M=17$ ,  $SD=4$ ) on the fifth-grade District Vocabulary Test. Responses to each item by upper and lower 27% of examinees, item difficulties, and item discriminating power are summarized in Appendix A.

*District Reading Comprehension Test*

*Comparability of Forms.* Students taking Form A ( $M=16$ ,  $SD=4$ ) of the *SR* portion of the fifth-grade District Reading Comprehension Test scored significantly higher than students taking Form B ( $M=14$ ,  $SD=3$ ),  $F(1, 1390)=68.31$ ,  $p < .001$ . However, students taking Form A ( $M=5.2$ ,  $SD=2.2$ ) of the *CR* portion of the fifth-grade District Reading Comprehension Test did not score significantly higher than students taking Form B ( $M=5.3$ ,  $SD=2.2$ ),  $F(1, 1383)=1.48$ ,  $p = .224$ .

*Group performance (Form A).* Analysis of variance revealed significant differences between groups on Form A of the District Reading Comprehension Test. Females outperformed

males, regular education students outperformed SPED students, and students from high income regions outperformed students from low income regions. Numbers were too small to analyze potential differences between ELL ( $n=1$ ) and non-ELL students. No significant differences in total scores across ethnic groups were found. Descriptive statistics and results of ANOVAs are depicted in tables 5 and 6.

*Group performance (Form B).* Analysis of variance revealed significant differences between groups on Form B of the District Reading Comprehension Test. Females outperformed males; Asian, White, and “other” students outperformed African American students; and regular education students outperformed SPED students. Numbers were too small to analyze potential differences between ELL ( $N=4$ ) and non-ELL students. No significant differences in total scores of students from high and low income regions were found. Descriptive statistics and results of ANOVAs are depicted in tables 7 and 8.

Table 5

*Descriptive Statistics for Grade 5 District Reading Test: SR (Form A)*

	Group	<i>n</i>	<i>M</i>	<i>SD</i>
Gender	Male	319	15.4	4.2
	Female	317	16.1	3.8
Ethnicity	White	377	16.3	3.7
	Hispanic	17	14.9	4.4
	African American	13	17.2	3.1
	Asian	30	16.6	3.8
	Native American	19	14.8	4.4
	Other	17	15.7	4.1
SPED	SPED	62	13.8	4.2
	Non-Sped	574	16.0	3.9
Income	Low	390	15.0	4.1
	High	246	16.9	3.6
Total		636	15.7	4.0

Table 6 Analysis of Variance Summary Table  
*Grade 5 District Reading Test: SR (Form A)*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>P</i>
Gender	1	4.02*	.006	.045
Error	634			
Ethnicity	5	1.20	.013	.309
Error	467			
SPED	1	16.11**	.025	.000
Error	634			
Income	1	34.5**	.052	.000
Error	634			

*Note.* Values enclosed in parentheses represent mean square errors.

\* $p < .05$ , \*\* $p < .01$ .

Table 7

*Descriptive Statistics for Grade 5 District Reading Test: SR (Form B)*

	Group	<i>n</i>	<i>M</i>	<i>SD</i>
Gender	Male	385	14.0	3.0
	Female	371	14.4	2.8
Ethnicity	White	595	14.3	2.8
	Hispanic	35	13.2	3.8
	African American	17	11.8	3.4
	Asian	38	14.3	3.1
	Native American	15	13.9	3.7
	Other	41	14.8	2.5
	SPED	SPED	111	12.4
	Non-Sped	645	14.5	2.6
Income	Low	420	14.3	2.9
	High	336	14.1	2.9
Total		756	14.2	2.9

Table 8 Analysis of Variance Summary Table  
 Grade 5 District Reading Test: SR (Form B)

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>P</i>
Gender	1	3.94*	.005	.047
	754			
Ethnicity	5	3.82**	.025	.002
	735			
SPED	1	56.51**	.070	.000
	754			
Income	1	1.15	.002	.285
	754			

*Note.* Values enclosed in parentheses represent mean square errors.

\* $p < .05$ , \*\* $p < .01$ .

*Item analysis: SR.* Separate item analyses were conducted on Form A, “Bessie Coleman,” ( $n=636$ ) and Form B, “The Stag” ( $n=756$ ). The upper 27% of examinees ( $n=172$ ) scored between 19 and 21 ( $M=20$ ,  $SD=.74$ ) and the lower 27% of examinees ( $n=172$ ) scored between 1 and 14 ( $M=10$ ,  $SD=3$ ) on the SR portion of the Reading Comprehension Test of Form A. Responses to each item by the upper and lower 27% of examinees, item difficulties, and item discriminating power for Form A “Bessie Coleman” are summarized in Appendix B.

The upper 27% of examinees ( $n=204$ ) scored between 16 and 18 ( $M=17$ ,  $SD=.57$ ) and the lower 27% of examinees ( $n=204$ ) scored between 2 and 13 ( $M=10$ ,  $SD=3$ ) on the SR portion of the Reading Comprehension Test of Form B. Responses to each item by the upper and lower 27% of examinees, item difficulties, and item discriminating power for Form B “The Stag” are summarized in appendix C.

*Item analysis: CR.* Separate item analyses were conducted on Form A “Bessie Coleman” ( $n=630$ ) and Form B “The Stag” ( $n=746$ ). The upper 27% of examinees ( $n=170$ ) scored between 7 and 8 ( $M=7.5$ ,  $SD=.50$ ) and the lower 27% of examinees ( $n=170$ ) scored between 0 and 4 ( $M=2.2$ ,  $SD=1.3$ ) on the CR portion of the Reading Comprehension Test of Form A. Responses to each item by the upper and lower 27% of examinees, item difficulties, and item discriminating power for Form A “Bessie Coleman” are summarized in table 9.

The upper 27% of examinees ( $N=201$ ) scored between 7 and 8 ( $M=7.8$ ,  $SD=.40$ ) and the lower 27% of examinees ( $N=201$ ) scored between 0 and 4 ( $M=2.5$ ,  $SD=1.2$ ) on the CR portion of the Reading Comprehension Test of Form B. Responses to each item by the upper and lower 27% of examinees, item difficulties, and item discriminating power for Form B “The Stag” are summarized in table 9.

Table 9  
*Item Analysis: CRs*

Form	Item	Difficulty	Differentiation (Item-Total Correlation)
A	1	66%	.65
A	2	63%	.71
A	3	60%	.76
A	4	55%	.76
B	1	69%	.65
B	2	59%	.75
B	3	74%	.58
B	4	56%	.73



*Item reduction.* Items were removed, based on a combination of difficulty score and differentiation power, from Forms A and B for both SRs and CRs. Discrimination power was given the most weight followed by item difficulty (i.e., easier items that did not effectively discriminate). Table 10 summarizes eliminated items and the resulting impact on mean difficulty and mean discrimination.

Table 10

*Items for Removal from Grade 5 Reading Test and Impact of Removal*

Item Type and Form	Item #s Removed	Mean Difficulty		Mean Discrimination (Item-Total Correlation)	
		Before Removal	After Removal	Before Removal	After Removal
SR: A	1,2,7,8,13,14	72%	70%	.46	.49
SR: B	4, 8, 18	74%	76%	.43	.46
CR: A	1, 2	61%	58%	.72	.76
CR: B	1, 3	65%	58%	.68	.74

SR: Selection Response CR: Constructed Response

*Discussion**Oral Reading Fluency*

The ORF scores did not differ between forms A and B of the District Reading Comprehension Assessment. This finding suggests that ORF is stable across both forms. However, the present analysis suggests that there are differences in performance across gender, ethnicity, SPED classification, income region, and ELL classification.

*District Vocabulary Test*

The present analysis found significant differences in performance on the District Vocabulary Test for Grade 5 across ethnicity, special education classification, income region, and English language learner classification. The average difficulty of the District Vocabulary Test was 82%. The average discrimination power was .45. Eliminating easy items and items that do not sufficiently discriminate (e.g., item 17\_Adulthood) could help improve the functioning of the Vocabulary assessment.

*Fifth-Grade District Reading Comprehension Test: SR*

Student performance on Form A, “Bessie Coleman,” and Form B, “The Stag,” of the SR portion of the District Reading Assessment differed significantly – suggesting lack of comparability of forms. Significant differences on Form A in performance by group (i.e., gender, SPED classification, income region) were found. On Form B, significant group differences (gender, ethnicity, special education classification) also were found. SR items, on Forms A and B, were reduced to 15 questions. CR items, of Forms A and B, were reduced to two items. The removal of items improved the discrimination power of selection and CRs for both Forms A and B.

The forms remain non-comparable. This may be due to the content of the forms (fiction vs. non-fiction). After having removed the items (bringing both forms to 15 items), students on Form B ( $M=12.2$ ,  $SD=2.6$ ), on average, significantly outperform students on Form A ( $M=11.6$ ,  $SD=2.8$ ). Making Form B more difficult would help align the performance between forms.

The district’s current Reading Assessment kit is a commendable model. It can offer insights into strengths of particular programs, schools, and teachers and provides school

personnel with information that can help them measure their progress toward promoting reading proficiency for all students.

## Appendix A

## Item Analysis -- District Vocabulary Test

Item	Group	%	%	%	Difficulty	Discrimination (Item-Total Correlation)
		Selecting Alternative 1	Selecting Alternative 2	Selecting Alternative 3		
1_Ambition	Lower	<b>42.2%</b>	23.6%	34.2%	70%	.46
	Upper	<b>97.8%</b>	1.1%	1.1%		
2_Blunder	Lower	<b>28.9%</b>	54.0%	17.2%	60%	.38
	Upper	<b>90.0%</b>	8.6%	1.4%		
3_Powerless	Lower	4.6%	7.3%	<b>88.0%</b>	94%	.34
	Upper			<b>100.0%</b>		
4_Postpone	Lower	<b>58.1%</b>	31.2%	10.7%	79%	.52
	Upper	<b>99.5%</b>	.3%	.3%		
5_Boldness	Lower	32.9%	19.6%	<b>47.6%</b>	72%	.44
	Upper	1.6%	1.1%	<b>97.3%</b>		
6_Distress	Lower	<b>30.2%</b>	17.2%	52.6%	56%	.31
	Upper	<b>81.9%</b>	1.1%	17.0%		
7_Visual	Lower	9.5%	10.9%	<b>79.6%</b>	90%	.51
	Upper			<b>100.0%</b>		
8_Captivity	Lower	14.2%	<b>63.1%</b>	22.7%	81%	.46
	Upper	.3%	<b>99.2%</b>	.5%		
9_Navigate	Lower	11.7%	<b>78.5%</b>	9.8%	89%	.49
	Upper	.3%	<b>99.7%</b>			
10_Bleach	Lower	4.6%	10.9%	<b>84.5%</b>	92%	.43
	Upper			<b>100.0%</b>		
11_Transfer	Lower	<b>83.7%</b>	7.1%	9.2%	92%	.48
	Upper	<b>99.7%</b>	.3%			
12_Vanish	Lower				95%	.47
	Upper		<b>100.0%</b>			
13_Penalty	Lower	11.8%	23.1%	<b>65.1%</b>	82%	.47
	Upper		.8%	<b>99.2%</b>		
14_Difficult	Lower	4.6%	3.0%	<b>92.4%</b>	96%	.40
	Upper			<b>100.0%</b>		
15_Effort	Lower	7.7%	<b>83.8%</b>	8.5%	92%	.52
	Upper		<b>100.0%</b>			
16_Notice	Lower	4.6%	20.2%	<b>75.1%</b>	87%	.33
	Upper		1.9%	<b>98.1%</b>		

17_Adulthood	Lower Upper	13.2% .8%	<b>84.1%</b> <b>99.2%</b>	2.7%	92%	.38
18_Modern	Lower Upper	<b>34.7%</b> <b>95.1%</b>	15.7% .3%	49.6% 4.6%	65%	.45
19_Rude	Lower Upper	<b>85.8%</b> <b>100.0%</b>	7.9%	6.3%	93%	.47
20_Defend	Lower Upper	10.8% .3%	8.6% .3%	<b>80.7%</b> <b>99.5%</b>	90%	.48
21_Anxious	Lower Upper	21.5% .8%	16.5%	<b>62.0%</b> <b>99.2%</b>	81%	.53
22_Gossip	Lower Upper	<b>63.5%</b> <b>97.3%</b>	17.9% .8%	18.7% 1.9%	81%	.46
23_Paralyze	Lower Upper	36.3% 3.5%	<b>53.0%</b> <b>96.5%</b>	10.7%	75%	.45
24_Pledge	Lower Upper	<b>54.6%</b> <b>97.6%</b>	22.4% 1.9%	23.0% .5%	76%	.44
25_Usatisfactory	Lower Upper	<b>66.9%</b> <b>100.0%</b>	13.3%	19.9%	84%	.56
26_Clatter	Lower Upper	24.3% 1.1%	24.0% .3%	<b>51.7%</b> <b>98.6%</b>	75%	.49

## Appendix B

## Item Analysis: SR (Form A)

Item	Group	% Selecting A	% Selecting B	% Selecting C	% Selecting D	Difficulty	Discrimination (Item-Total Correlation)
1	Lower	13.5%	5.8%	<b>78.4%</b>	2.3%	89%	.37
	Upper	.6%	.6%	<b>98.8%</b>			
2	Lower	5.8%	9.9%	7.0%	<b>77.2%</b>	89%	.40
	Upper				<b>100.0%</b>		
3	Lower	<b>42.1%</b>	33.9%	11.7%	12.3%	71%	.53
	Upper	<b>99.4%</b>			.6%		
4	Lower	26.9%	32.7%	<b>22.8%</b>	17.5%	60%	.52
	Upper	2.9%	.6%	<b>96.0%</b>	.6%		
5	Lower	<b>37.1%</b>	31.8%	18.2%	12.9%	65%	.45
	Upper	<b>92.5%</b>	1.2%	6.4%			
6	Lower	15.7%	<b>68.6%</b>	3.5%	12.2%	84%	.57
	Upper		<b>100.0%</b>				
7	Lower	29.4%	7.6%	14.7%	<b>48.2%</b>	71%	.38
	Upper	6.4%	.6%		<b>93.1%</b>		
8	Lower	<b>63.7%</b>	7.6%	23.4%	5.3%	81%	.41
	Upper	<b>98.3%</b>	1.2%	.6%			
9	Lower	20.7%	10.7%	6.5%	<b>62.1%</b>	81%	.51
	Upper				<b>100.0%</b>		
10	Lower	9.4%	<b>51.2%</b>	20.6%	18.8%	75%	.48
	Upper		<b>98.8%</b>	.6%	.6%		
11	Lower	<b>52.9%</b>	5.3%	25.3%	16.5%	76%	.55
	Upper	<b>98.8%</b>		.6%	.6%		
12	Lower	17.8%	<b>46.2%</b>	10.7%	25.4%	69%	.46
	Upper	6.4%	<b>91.3%</b>		2.3%		
13	Lower	13.5%	46.5%	12.4%	<b>27.6%</b>	50%	.31
	Upper	2.3%	22.7%	2.9%	<b>72.1%</b>		
14	Lower	20.0%	6.5%	<b>66.5%</b>	7.1%	83%	.41
	Upper			<b>98.8%</b>	1.2%		
15	Lower	22.5%	17.2%	<b>55.0%</b>	5.3%	76%	.49
	Upper	1.7%	1.2%	<b>96.5%</b>	.6%		

16	Lower Upper	<b>24.9%</b> <b>80.3%</b>	16.6% 4.0%	37.3% 15.6%	21.3%	53%	.42
17	Lower Upper	<b>35.5%</b> <b>95.4%</b>	27.8% 4.6%	24.3%	12.4%	66%	.48
18	Lower Upper	10.1% <b>33.1%</b> <b>94.8%</b>		22.5% .6%	34.3% 4.6%	64%	.47
19	Lower Upper	9.2% <b>34.8%</b> <b>87.9%</b>		34.8% 6.4%	21.3% 5.8%	64%	.45
20	Lower Upper	13.5% 17.0%		22.0% 7.6%	<b>47.5%</b> <b>92.4%</b>	72%	.47
21	Lower Upper	14.0%	5.9%	<b>51.5%</b> <b>100.0%</b>	28.7%	78%	.56

## Appendix C

## Item Analysis: SR (Form B)

Item	Group	% Selecting A	% Selecting B	% Selecting C	% Selecting D	Difficulty	Discrimination (Item-Total Correlation)
1	Lower	9.90%	7.40%	<b>77.30%</b>	5.40%	88%	.40
	Upper	0.50%	0.50%	<b>98.50%</b>	0.50%		
2	Lower	3.90%	6.90%	4.90%	<b>84.30%</b>	92%	.43
	Upper				<b>100.00%</b>		
3	Lower	6.40%	<b>69.50%</b>	19.70%	4.40%	83%	.41
	Upper	2.00%	<b>97.10%</b>	0.50%	0.50%		
4	Lower	<b>80.90%</b>	6.90%	4.40%	7.80%	89%	.34
	Upper	<b>98.00%</b>		2.00%			
5	Lower	23.20%	<b>38.90%</b>	30.50%	7.40%	67%	.52
	Upper	3.90%	<b>94.60%</b>	1.00%	0.50%		
6	Lower	<b>48.80%</b>	21.70%	9.90%	19.70%	70%	.44
	Upper	<b>91.20%</b>	3.90%		4.90%		
7	Lower	17.70%	8.40%	12.80%	<b>61.10%</b>	80%	.48
	Upper	1.50%			<b>98.50%</b>		
8	Lower	<b>36.80%</b>	26.90%	28.40%	8.00%	42%	.13
	Upper	<b>47.10%</b>	1.50%	51.50%			
9	Lower	<b>75.90%</b>	3.90%	5.40%	14.80%	87%	.46
	Upper	<b>98.00%</b>			2.00%		
10	Lower	15.80%	<b>41.90%</b>	22.70%	19.70%	67%	.42
	Upper	5.90%	<b>91.70%</b>	1.50%	1.00%		
11	Lower	19.70%	14.80%	<b>29.60%</b>	36.00%	56%	.38
	Upper	2.90%	2.50%	<b>81.90%</b>	12.70%		
12	Lower	7.40%	13.80%	22.20%	<b>56.70%</b>	77%	.39
	Upper		2.00%	1.50%	<b>96.60%</b>		
13	Lower	16.90%	15.40%	7.50%	<b>60.20%</b>	79%	.48
	Upper	1.50%	1.50%	0.50%	<b>96.60%</b>		
14	Lower	16.70%	16.70%	<b>59.10%</b>	7.40%	79%	.49
	Upper		0.50%	<b>99.50%</b>			
15	Lower	<b>41.60%</b>	20.30%	30.20%	7.90%	69%	.53
	Upper	<b>96.60%</b>	1.00%	2.00%	0.50%		



16	Lower	10.90%	16.80%	24.80%	<b>47.50%</b>	71%	.48
	Upper	2.00%	0.50%	2.90%	<b>94.60%</b>		
17	Lower	9.40%	33.70%	<b>52.50%</b>	4.50%	75%	.52
	Upper	2.50%	1.00%	<b>96.60%</b>			
18	Lower	14.10%	<b>38.70%</b>	30.20%	17.10%	64%	.38
	Upper	4.40%	<b>89.20%</b>	2.50%	3.90%		