

Differential Item Functioning in a State-wide Visual Arts Assessment Using Item-focused Rasch Tree Methodology

Ning Jiang
Kelvin Pompey
Yin Burgess
Ashlee Lewis Ph. D.



SCEPUR– March 1st, 2019

1

South Carolina Arts Assessment Program (SCAAP)

- Administered by REM Center - Funded by SCDE
- Statewide, low stakes standards-based assessment
- Schools who receive Distinguished Arts Program grants
- Opt-in model
- “Report cards” provided to schools

2

South Carolina Arts Assessment Program (SCAAP)

- 4th grade
- Multiple-choice
 - Administered online
 - 45-items
- Performance Assessment
 - Two tasks
 - One-on-one administration

3

South Carolina Arts Assessment Program (SCAAP)

- Test Development
 - Multiple-choice
 - Each item targets a standard and an indicator
 - Each item has a designated Blooms' taxonomy level
 - Test form developed to match Test Specifications
 - Many items have multi-media stimuli (sound files or visual stimuli)
 - Performance Tasks
 - Target standards related to performance (music), creation, and critique

4

South Carolina Arts Assessment Program (SCAAP)

The Study

- In educational assessment, many decisions are made utilizing student assessment results.
- The validity of interpretations is vital.
- Items should not provide a clear advantage for one subgroup over another. Otherwise, items are biased or have differential item functioning (DIF), which is a validity issue (Crocker & Algina, 1986; Hidalgo & López-Pina, 2004).
- An important step in the validation process, especially for all large-scale educational assessments, involves evaluating items for DIF (Walker, 2011).

5

South Carolina Arts Assessment Program (SCAAP)

The Study (Cont'd)

- Studies have suggested gender difference in academic achievements in various school subjects (Agak, 2002), but few have focused on visual arts.
- Reports have shown ethnicity differences in school performance and standardized test scores (Duncan & Magnuson; 2005, NAEP, 2008; as in Singh, Chang, & Dika, 2010), but few on visual arts.
- Research on relationship between visual arts and gender is limited.
- Research on relationship between visual arts and ethnicity is still lacking.

6

South Carolina Arts Assessment Program (SCAAP)

Method 1: Rasch Trees

- Measurement tool used to detect the variables/combinations of variables that induce DIF
- Method: Model-based Recursive Partitioning
 1. Estimate Rasch model over entire sample
 2. Select variable with greatest potential for inducing DIF based on Parameter Instability Statistical Test
 3. Determine optimal way to divide sample using covariate
 4. Repeat steps 1-3 on subsamples until no significant DIF is found or until a minimum sample size has been reached.
- End Result: Tree structure displaying how data was partitioned, and which variables were used to divide data, and the estimated item difficulty parameters of each subsample.
- Software: psychotree package in R

7

South Carolina Arts Assessment Program (SCAAP)

Method 2: Item-focused Trees

- Measurement tool used to detect the variables/combinations of variables that induce DIF on the item level
- Method: Model-based Recursive Partitioning
 1. Select the item, variable, and optimal way to split the sample with greatest potential for inducing DIF.
 2. Statistically test evidence of DIF.
 3. Repeat until no item, variables, split point combination is significant.
- End Result: Tree structure for each item displaying the variables used to split the sample and the estimated item difficulty parameters for each subsample.
- Software: DIFtree package in R

8

South Carolina Arts Assessment Program (SCAAP)

Results-Dimensionality

- Raw variance explained by items: 13.8%
Unexplained variance in the first contrast: 3.0%
- Secondary dimension: eigenvalue of 1.7
- Loadings from the first contrast: 0.00 to 0.40;
-0.38 to -0.02

9

South Carolina Arts Assessment Program (SCAAP)

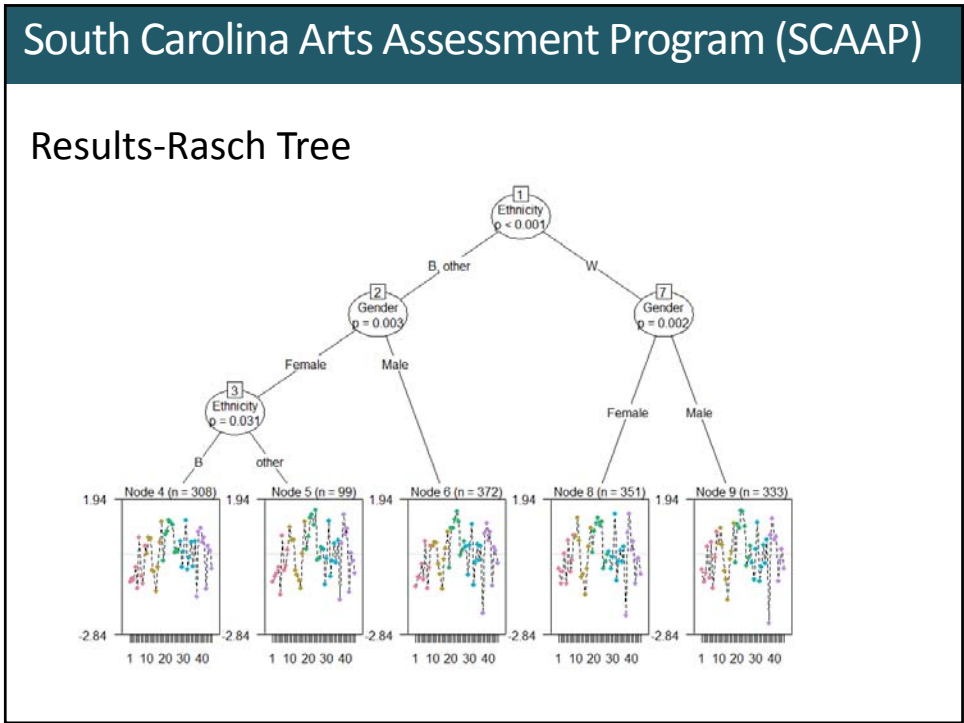
Results-Rasch Tree DIF

Table 1. *Rasch trees parameter stability test results*

Node	Covariates			
	Ethnicity		Gender	
	S	P value	S	P Value
1	234.2	< .001 ⁺ *	125.6	< .001 ⁺
2	74.4	.006 ⁺	76.4	.003*
3	63.0	.031 ⁺ *	-	-
6	56.2	.103	-	-
7	-	-	75.4	.002 ⁺ *

⁺ indicates a statistically significant .05 Bonferroni adjusted result and * indicates the *p* value associated with the covariate that was partitioned.

10



11

South Carolina Arts Assessment Program (SCAAP)

Results-Item-focused Tree

Table 2. DIF items in the visual arts assessment

Item	Type	Favorable Groups	Number of splits	Difference between favorable and non-favorable group
2	Uniform	White	1	0.58
6	Uniform	White	1	0.49
7	Uniform	Female	1	0.64
10	Uniform	Other	1	0.53
11	Uniform	White	1	0.73
13	Uniform	Female	1	0.39
16	Uniform	Female, Other	2	0.52/1.25
20	Uniform	Female	1	0.49
22	Uniform	White	1	0.38
23	Uniform	Female	1	0.56
27	Uniform	Female	1	0.43
30	Uniform	Female	1	0.40

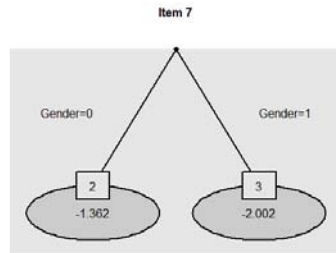
12

South Carolina Arts Assessment Program (SCAAP)

Results-Item-focused Tree (Cont'd)

7. What type of line may **BEST** represent lightning?

- A. Curvy
- B. Straight
- C. Wavy
- D. Zigzag



13

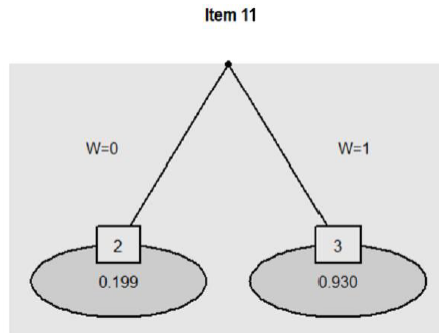
South Carolina Arts Assessment Program (SCAAP)

Results-Item-focused Tree (Cont'd)



11. The South Carolina artist, Philip Simmons, is famous for his wrought iron gates. His design uses symmetry, shape, and

- A. color
- B. line
- C. texture
- D. value




14

South Carolina Arts Assessment Program (SCAAP)

Results-Item-focused Tree (Cont'd)

Use the painting and read the critiques to answer the question.



Susie's critique

I like this painting a lot. The dresses are very beautiful. I like the girl's long black hair. I hope I will be able to grow my hair that long. I also like it because there is a lot of blue in this picture. I like the yellow sash on the girl's dress too.

Shawna's critique

I like this painting because the artist made the proper look like they are sitting. Standing, they also puts the arms in front of the woman to create depth. He also used bright colors to make the painting look happy.

16. Who wrote a better critique and why?

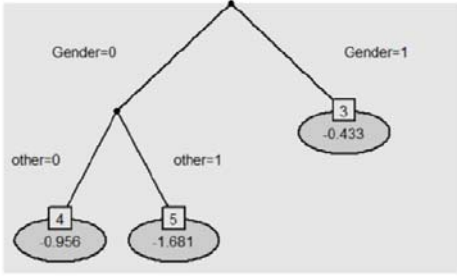
A. Susie because it is the longer one.

B. Shawna because she uses bigger words.

C. Susie because it is based on her opinion.

D. Shawna because she talks about the elements of art.

Item 16



15

South Carolina Arts Assessment Program (SCAAP)

Limitations

- Rasch Trees method is on the global test level and should only be an indicator of differential test functioning.
- Due to the computational intensity of the Item-focused trees method, DIF detection could not be conducted on a large set of items.
- Due to testing DIF in a subset of items at a time, item purification tests may be required to remove the effect of DIF.
- Results of the Item-focused Trees method should be used exploratorily. Careful interpretation is necessary because results may be more of an indication of measurement invariance and not DIF.

16

South Carolina Arts Assessment Program (SCAAP)

Thank you!

Ginger Jiang

njiang@email.sc.edu

