The Reading Excellence Act Grant Application Process: A Comprehensive Example of Applied Rasch Measurement

Ramazan Basturk
Willaim E. Loadman

The Ohio State University

The Paper Presented at the annual Meeting of the Mid-western Educational Research Association
Chicago, IL
October 25, 2001

Introduction

School districts in Ohio were given the opportunity to apply for funding to support district based reading programs through a request for proposal procedures. An effort was made to reliably and equitably score the proposed applications. The grant selection process for reading excellence program included two phases. In the first phase, applications were selected that meet established criteria and were thought to have sufficient merit to be considered for phase two. The criteria that were used to selection in the first phase were: a) basic requirements, b) required information, c) significance, d) the quality of proposal, e) adequacy of resources, and f) quality of project evaluation.

Phase two, on the other hand, was an on-site evaluation conducted by a three-member review team in order to assess the strengths of the proposed Reading Excellence Act (REA) grant application. The purpose of the phase two was: a) verify application information; b) assess the capacity and commitment of the districts; and c) ensure likely success of the proposal (ODE., 1999). The criteria that were used to evaluate the strengths of the proposed REA grant application in the phase two were to:

a) Coordination of REA activities that impact classroom instruction and focus on improving student academic performance;
b) Connections among professional development for teachers, family literacy activities, and extended learning opportunities for students;
c) Evidence of significally-based reading research in literacy plan;
d) Commitment of staff members at targeted elementary schools;
e) Active involvement of districts leadership in REA plan;

f) Partnerships between targeted schools and early childhood educators in community;

g) Involvement of parents and other community stakeholders;

h) Alignment of districts resources and REA proposed budget to meet literacy goals identified in grant proposal;

i) Capacity to implement proposed literacy plan.

In order to analyze the quality of applications, ordinary analysis procedures were not considered appropriate for a fair grant funding procedure. The Rasch Item Response Theory was considered a more effective data analysis procedure to identify the quality of applications because multiple raters were going to be reading and evaluating each grant application. According to Linacre (1993), the many-facet Rasch model has distinct advantages over classical data analysis. These advantages include the use of person measures rather than raw scores and the adjustment of person measures for facets included in the model (Weigle, 1994). Another advantage of many-facet Rasch model is the facet "connectedness" that is required if linear "rules" are to be created for each facet (Schumacker, 1996). The facet analysis "can provide a framework for obtaining objective and fair measurement" (Engelhard, 1992).

Method of Inquiry and Instrumentation

The data set include 106 applications from over 3000 buildings in 612 school districts. Applications have been assigned randomly to 114 readers to evaluate the quality of applications and at least 4 raters were assigned per application.

The survey that was used to evaluate Grant Applications for Reading Excellence Act contains three subscales and 26 questions. The first six items related to the ***district's***

reading tutor program as discussed in the application. The next eight questions outline school-***based tutor service***. Finally, the last eleven questions relate only to the ***contracted tutorial services***. An overall evaluation of grant application quality is found in the last question, Question (#26). Six-point, likert type, scale was used to evaluate the quality of application using the following criteria:

1 = No evidence; very weak; lowest rating

2 = Minimal evidence; little support; hardly observable; vague; weak concepts

3 = Some evidence; some potential for effectiveness; partially developed concepts; needs more work

4 = Enough evidence to indicate a fairly good chance of success; good concepts; on the right track

5 = Strong evidence, easily seen; several success seen; well developed concepts; well underway

6 = Exceptionally strong evidence; outstanding potential; high quality; exceptional quality; highest rating.

Results

An extension of the Rasch model to include multiple facets (FACETS model) was used in analyzing Reading Excellence Act Tutorial Assistance survey. Basically, FACETS analysis provides estimates of examine ability, rater severity and item difficulty on a common log-linear metric or logit scale (Linacre, 1993). The mathematical definition of the three-faceted model with facets of application, rater and item can be expressed as follows:

$$\text{Log}\ (P_{njk} / P_{njk-1}) = B_n - D_i - C_j - F_k$$

Where,

P_{njk} = Probability that person n on item i is rated by judge j with score of k

B_n = ability n

D_i = difficulty I

C_j = severity j

F_k = Challenge k.

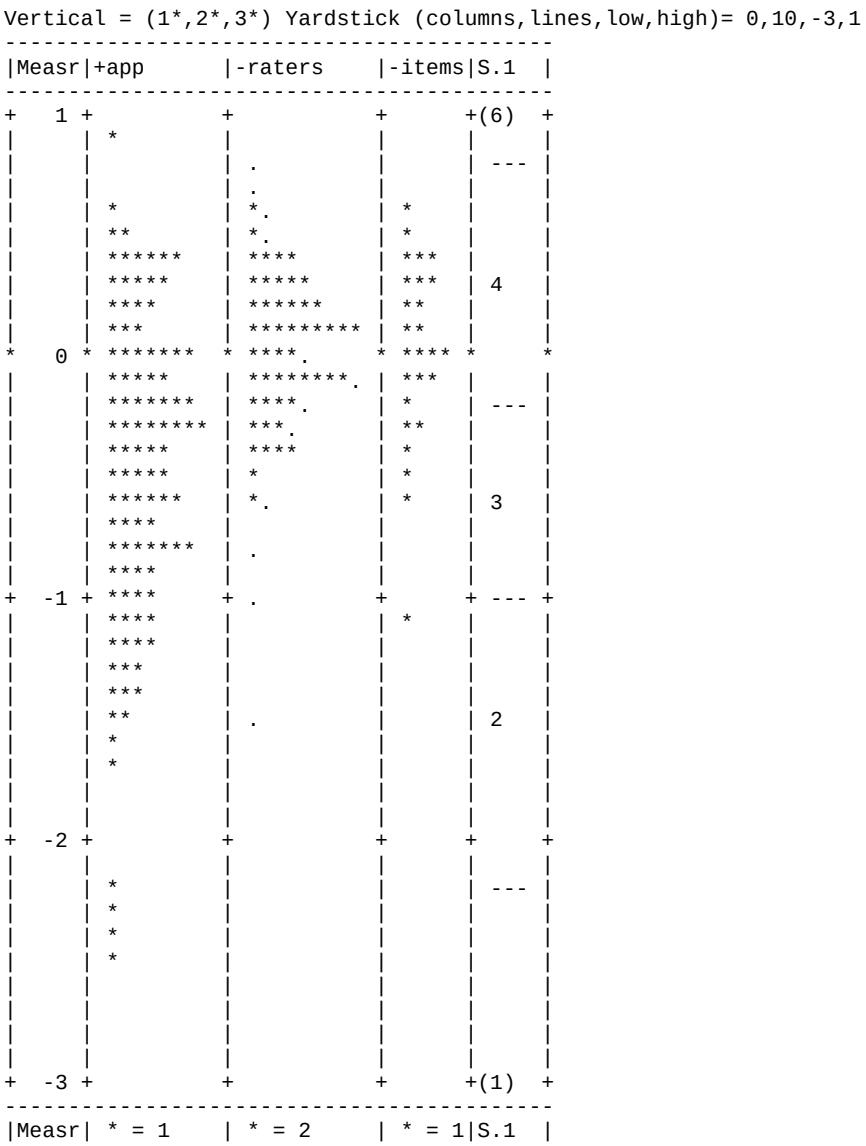
In addition to providing logit estimates of the ability, severity or difficulty of each element of each facet, FACETS also provides statistics indicating the relative spread of these estimates within each facet. In other words, the analysis provides information about the significance of any differences that may exist among elements of a facet; for example differences in severity among raters or quality among applications.

Another important feature of the FACETS analysis is that it provides fit statistics for each element, which provide an indication of degree to which each element is behaving in a manner that is predicted by the model. In the case of raters, the fit statistics are indicators of rater consistency. Thus a detailed picture of the behavior of each rater in terms of both severity and consistency can be formed.

As an overall introduction to the Reading Excellence Act Tutorial Application Rating analysis results, Figure 1 shows graphically the measures for applications, raters and items for the data. The figure is to be interpreted as follows. The scale along the left of the figure represents the logit scale, which is the same for all three facets. Each application is represented by the star (*). Applications are ordered so that the most quality application at the top, and the application with lowest quality is at the bottom. The other facets are ordered so that the most difficult element of each facet is towards the

top, and the least difficult elements towards the bottom. In terms of Raters, for example, the most severe rater is the uppermost rater in the figure 1. Similarly, the most acceptable item is uppermost in the figure 1. The figure 1 thus shows pictorially the differences across the different elements of each facet.

Figure 1. Reading Excellence Act Tutorial Assistance All Facet Summary.



In figure 1, applications are ordered with the highest quality applications at the top and the lowest quality applications at the bottom. As the figure 1 indicates,

application estimates range from a high of about +1 logits to a low of close to –3 logits. Looking at the column for applications on the figure 1, we can see that applications mainly are low quality than high quality. Of the application clustered around the mean (0), 22 out of 104 applications (21%, respectively) above the 0.

Looking at the column for raters on the figure 1, we can see that raters mainly are neither severe nor lenient. Of the raters clustered around the mean (0), 54 out of 114 readers (47%, resectively) of the raters tend to be more severe than the leinent

Looking at the column for items on the figure 1, we can see that items mainly are neither high nor low quality. Of the raters clustered around the mean (0), 12 out of 26 items (46%, resectively) have more quality than less quality.

Application Analysis

Application's Quality

A more detailed analysis about applications is found in table 7.1.1, with the title of the Application Measurement Report for the Reading Excellence Act Tutorial Assistance. Applications are presented in descending order of quality; in other words, Application **#3029** and **#3026** are the highest quality applications and application **#3060** and **#3061** are the lowest quality applications, as was seen in Table 7.1.1.

Reading Excellence Act Tutorial Assistance 04-30-2000 01:27:25
Table 7.1.1 app Measurement Report (arranged by mN).

Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Model Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Num app
691	156	4.4	4.57	.92	.09	1.1	0	1.0	0	**3029 3029**
547	130	4.2	4.25	.56	.09	1.2	1	1.2	1	**3026 3026**
531	130	4.1	4.20	.51	.09	1.0	0	1.0	0	3000 3000
529	130	4.1	4.20	.51	.09	1.1	0	1.1	0	3062 3062
494	130	3.8	4.10	.40	.08	1.0	0	1.0	0	3050 3050
426	104	4.1	4.08	.38	.10	0.8	-1	0.8	-1	3066 3066
520	130	4.0	4.07	.37	.09	0.8	-2	0.8	-1	3014 3014
495	130	3.8	4.07	.37	.09	1.0	0	1.0	0	3078 3078

526	130	4.0	4.05	.35	.09	0.8	-1	0.8	-1	3079	3079	
554	130	4.3	4.05	.35	.09	1.1	0	1.1	0	3093	3093	
393	104	3.8	4.04	.35	.10	0.8	-1	0.8	-2	3107	3107	
525	130	4.0	4.04	.34	.09	**0.7**	**-3**	**0.7**	**-3**	**3069**	**3069***	
524	130	4.0	4.03	.33	.09	**0.6**	**-3**	**0.6**	**-3**	**3031**	**3031***	
521	130	4.0	4.02	.33	.09	**0.7**	**-2**	**0.8**	**-2**	**3047**	**3047***	
547	130	4.2	3.98	.29	.09	0.9	-1	0.8	-2	3090	3090	
479	130	3.7	3.92	.23	.08	1.1	0	1.1	0	3019	3019	
513	130	3.9	3.91	.22	.09	1.1	0	1.1	0	3038	3038	
533	130	4.1	3.89	.20	.09	**0.7**	**-2**	**0.7**	**-2**	**3010**	**3010***	
534	130	4.1	3.88	.19	.09	1.1	1	1.2	1	3070	3070	
479	130	3.7	3.83	.14	.08	**1.6**	**4**	**1.6**	**4**	**3043**	**3043****	
503	130	3.9	3.75	.07	.09	0.9	0	0.9	0	3028	3028	
488	130	3.8	3.75	.07	.08	1.1	0	1.1	0	3094	3094	
458	130	3.5	3.70	.03	.08	**0.5**	**-5**	**0.5**	**-5**	**3018**	**3018***	
473	130	3.6	3.70	.03	.08	0.8	-2	0.8	-2	3075	3075	
466	130	3.6	3.69	.01	.08	0.8	-2	0.8	-2	3095	3095	
10576	2912	3.6	3.67	.00	.02	1.1	3	1.1	3	3	3	
486	130	3.7	3.68	.00	.08	1.0	0	1.1	0	3045	3045	
576	156	3.7	3.63	-.04	.08	1.0	0	1.0	0	3025	3025	
447	130	3.4	3.62	-.05	.08	1.0	0	1.1	0	3013	3013	
467	130	3.6	3.59	-.07	.08	**0.6**	**-3**	**0.6**	**-3**	**3081**	**3081***	
472	130	3.6	3.57	-.09	.08	0.9	-1	0.9	-1	3020	3020	
452	130	3.5	3.57	-.09	.08	0.9	0	0.9	0	3042	3042	
462	130	3.6	3.52	-.13	.08	0.8	-1	0.8	-1	3006	3006	
433	130	3.3	3.53	-.13	.08	**0.6**	**-4**	**0.6**	**-4**	**3056**	**3056***	
461	130	3.5	3.50	-.15	.08	**0.5**	**-5**	**0.5**	**-5**	**3080**	**3080***	
433	130	3.3	3.48	-.17	.08	**0.6**	**-4**	**0.6**	**-4**	**3086**	**3086***	
541	156	3.5	3.45	-.19	.08	0.9	0	0.9	0	3073	3073	
401	130	3.1	3.44	-.20	.08	**0.6**	**-4**	**0.6**	**-4**	**3057**	**3057***	
444	130	3.4	3.43	-.21	.08	**0.7**	**-2**	**0.8**	**-2**	**3030**	**3030***	
450	130	3.5	3.40	-.23	.08	**0.6**	**-3**	**0.6**	**-3**	**3085**	**3085***	
508	156	3.3	3.38	-.25	.07	0.9	0	0.9	0	3083	3083	
420	130	3.2	3.37	-.26	.08	1.1	0	1.1	0	3027	3027	
426	130	3.3	3.37	-.26	.08	**1.5**	**3**	**1.5**	**3**	**3099**	**3099****	
387	130	3.0	3.35	-.28	.08	**0.7**	**-2**	**0.7**	**-2**	**3040**	**3040***	
387	130	3.0	3.34	-.29	.08	**0.4**	**-6**	**0.4**	**-6**	**3041**	**3041***	
443	130	3.4	3.34	-.29	.08	**0.7**	**-3**	**0.7**	**-3**	**3046**	**3046***	
400	130	3.1	3.32	-.31	.08	**0.6**	**-3**	**0.6**	**-3**	**3044**	**3044***	
415	130	3.2	3.31	-.31	.08	1.3	2	1.3	2	3058	3058	
447	130	3.4	3.27	-.35	.08	**0.5**	**-6**	**0.5**	**-5**	**3005**	**3005***	
437	130	3.4	3.21	-.39	.08	**0.6**	**-3**	**0.6**	**-3**	**3072**	**3072***	
442	130	3.4	3.21	-.40	.08	0.8	-1	0.9	-1	3064	3064	
415	130	3.2	3.20	-.41	.08	**0.7**	**-2**	**0.7**	**-2**	**3001**	**3001***	
446	130	3.4	3.20	-.41	.09	0.8	-1	0.8	-2	3102	3102	
425	130	3.3	3.15	-.44	.08	0.9	0	0.9	0	3049	3049	
417	130	3.2	3.13	-.46	.08	1.1	0	1.1	0	3015	3015	
298	104	2.9	3.11	-.48	.09	**1.4**	**3**	**1.4**	**3**	**3074**	**3074****	
383	130	2.9	3.06	-.52	.08	0.8	-1	0.8	-1	3016	3016	
399	130	3.1	3.06	-.52	.08	1.3	2	1.3	2	3021	3021	
477	156	3.1	3.04	-.54	.07	0.9	0	0.9	0	3012	3012	
386	130	3.0	3.02	-.55	.08	**1.9**	**6**	**1.9**	**6**	**3052**	**3052****	
373	130	2.9	2.99	-.58	.08	0.8	-2	0.8	-1	3035	3035	
403	130	3.1	2.99	-.58	.08	1.1	0	1.1	0	3097	3097	
356	130	2.7	2.97	-.60	.08	**0.7**	**-2**	**0.7**	**-2**	**3065**	**3065***	
522	182	2.9	2.97	-.60	.07	0.9	-1	0.9	-1	3067	3067	
369	130	2.8	2.92	-.63	.08	0.9	0	0.9	-1	3053	3053	
404	130	3.1	2.88	-.67	.08	**1.4**	**3**	**1.4**	**3**	**3011**	**3011****	
379	130	2.9	2.84	-.70	.08	**0.6**	**-3**	**0.6**	**-3**	**3089**	**3089***	
390	130	3.0	2.81	-.73	.08	0.8	-2	0.8	-2	3008	3008	
350	130	2.7	2.79	-.74	.08	0.9	-1	0.9	-1	3068	3068	
372	130	2.9	2.77	-.76	.08	**1.6**	**4**	**1.6**	**4**	**3022**	**3022****	
360	130	2.8	2.73	-.79	.08	0.9	-1	0.9	-1	3103	3103	
346	130	2.7	2.74	-.79	.08	**1.7**	**5**	**1.7**	**5**	**3106**	**3106****	
379	130	2.9	2.71	-.81	.08	**1.6**	**4**	**1.6**	**4**	**3007**	**3007****	
331	130	2.5	2.70	-.82	.09	**0.6**	**-4**	**0.6**	**-4**	**3076**	**3076***	
434	156	2.8	2.67	-.84	.08	0.9	-1	0.9	-1	3084	3084	
366	130	2.8	2.67	-.85	.08	**0.5**	**-5**	**0.6**	**-4**	**3096**	**3096***	
318	130	2.4	2.63	-.88	.09	**1.5**	**4**	**1.6**	**4**	**3071**	**3071****	
325	130	2.5	2.61	-.90	.09	**1.6**	**4**	**1.6**	**4**	**3082**	**3082****	
315	130	2.4	2.57	-.93	.09	1.3	2	1.3	2	3063	3063	

409	156	2.6	2.55	-.95	.08	**0.7**	**-3**	**0.7**	**-3**	**3039**	**3039***
309	130	2.4	2.52	-.98	.09	1.3	2	1.3	2	3023	3023
333	130	2.6	2.49	-1.01	.08	**1.9**	**6**	**2.0**	**6**	**3004**	**3004****
345	130	2.7	2.46	-1.03	.08	**1.4**	**3**	**1.4**	**3**	**3002**	**3002****
302	130	2.3	2.45	-1.04	.09	**2.1**	**7**	**2.1**	**7**	**3017**	**3017****
305	130	2.3	2.35	-1.13	.09	1.1	1	1.1	0	3100	3100
284	130	2.2	2.35	-1.13	.09	**1.4**	**3**	**1.5**	**3**	**3104**	**3104****
328	130	2.5	2.34	-1.14	.09	**1.4**	**3**	**1.4**	**3**	**3088**	**3088****
292	130	2.2	2.34	-1.14	.09	0.9	-1	0.9	0	3105	3105
277	130	2.1	2.32	-1.16	.09	0.8	-1	0.8	-1	3098	3098
284	130	2.2	2.26	-1.21	.09	1.3	2	1.3	2	3036	3036
294	130	2.3	2.27	-1.21	.09	1.1	0	1.1	0	3048	3048
303	130	2.3	2.24	-1.24	.09	1.0	0	1.0	0	3055	3055
280	130	2.2	2.20	-1.28	.09	0.8	-1	0.8	-2	3033	3033
287	130	2.2	2.17	-1.31	.09	0.9	0	0.9	0	3059	3059
255	130	2.0	2.13	-1.35	.10	0.9	0	1.0	0	3003	3003
231	104	2.2	2.12	-1.35	.10	**0.6**	**-3**	**0.6**	**-3**	**3024**	**3024***
292	130	2.2	2.06	-1.42	.09	0.9	0	1.0	0	3051	3051
249	130	1.9	2.04	-1.44	.10	1.0	0	0.9	0	3092	3092
248	130	1.9	1.97	-1.52	.10	1.1	0	1.1	0	3037	3037
246	130	1.9	1.95	-1.54	.10	1.2	1	1.1	0	3077	3077
237	130	1.8	1.90	-1.61	.10	1.1	1	1.4	2	3009	3009
238	130	1.8	1.84	-1.68	.10	**0.7**	**-2**	**0.7**	**-2**	**3101**	**3101***
244	156	1.6	1.52	-2.18	.11	0.9	-1	0.9	-1	3034	3034
191	130	1.5	1.47	-2.29	.13	1.1	0	1.3	1	3087	3087
205	130	1.6	1.43	-2.36	.12	1.1	0	1.2	1	3061	3061
180	130	1.4	1.37	-2.52	.14	0.8	-1	0.7	-1	3060	3060

Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Model Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Num	app
-497.6	157.7	3.1	3.09	-.51	.09	1.0	-0.4	1.0	-0.4	Mean (Count: 106)	
988.5	269.0	0.7	0.73	.68	.01	0.3	2.9	0.3	2.9	S.D.	

RMSE (Model) .09 Adj S.D. .67 Separation 7.70 Reliability .98
Fixed (all same) chi-square: 5760.1 d.f.: 105 significance: .00
Random (normal) chi-square: 104.4 d.f.: 104 significance: .47

Note: *** Muting**
**** Noise**

The FACETS analysis provides a number of indications of the magnitude of the differences among elements of a facet: in this case, the quality among applications. These are the ***RMSE, Reliability***, ***Separation Index*** and ***Fixed*** (all same) & ***Random*** (normal) ***Chi Square***., ***Infit*** and ***Outfit*** statistics.

Root Mean Square Standard Error, ***RMSE***, is produced for all non-extreme measures over application. RMSE score, .09, illustrate that application error is very low. After application variance has been adjusted for measurement error, Adjusted standard deviation found below the 1.0 (.67, respectively).

The ***Reliability*** statistics provided by the FACETS analysis indicates the degree to which the analysis reliability distinguishes between different levels of quality among the

elements of the facets (in this case, the different applications). For applications, the reliability is .98, indicating that the analysis is fairly reliably separating applications into different levels of quality.

The ***Separation index*** is the ratio of the corrected standard deviation of elements measures (in these cases, applications) to the root mean-square estimation error. If the applications were equally quality, the standard deviation of the application's quality estimates should be equal to or smaller than the mean estimation error of the entire data set. However, the Application Separation Index is 7.70, indicating that the variance among applications is about eight times the error of estimates.

Finally, the ***Fixed (***all same***) Chi-square*** tests the null hypothesis that all of the elements of the facet are equal. The Chi-square of 5760.1 with 105 df. is significant at p = .00, indicated that the null hypothesis must be rejected; in other words, the applications' quality are not equal.

The FACETS analysis also provides two measures of fit, or consistency: the infit and the outfit. The infit is the weighted mean-squared residual that is sensitive to unexpected responses near the point where decisions are being made. Less then .8 indicates muting: too little variation, lack of independence. More than 1.3 indicates noise: unmodelled excess variation. On the other hand, the outfit is the the unweighted mean-squared residual and is sensitive to extreme scores. This fit statistics has the same form as infit, but is the conventional mean-square which is more sensitive to outliers.

In addition to the mean squares, FACETS provides standardized infit and outfit statistics, which have an expected mean of 0 and standard deviation of 1. These statistics are useful for comparing the elements of a facet with each other, as they show the degree

of variability in addition raters' ratings relative to the amount of variability in the entire set. Standardized fit statistics greater than 6 or less than – 6 are generally signs of misfit.

Applying these information on the Table 7.1.1, we can see that 41 out of 106 application had very high (>1.4) or low (< 0.6) infit and/or high (>1.4) or low (< 0.6) outfit statistics: These Applications are: **# 3069, #3031, #3043, #3018, #3000, #3081, #3056, #3080, #3086, #3057, #3085, #3099, #3041, #3046, #3044, #3005, #3072, #3074, #3052, #3011, #3089, #3022, #3106, #3007, #3076, #3096, #3071, #3082, #3039, #3004, #3002, #3017, #3104, #3088, #3024, #3101, #3010, #3047, #3001, #3065 & #3030**. (Please look at the table 7.1.1 for more detail). These infit/outfit statistics indicate that these applications were not consistent with the estimated quality measures with 114 the score showing "noise" and 5 outfit scores "muting". In most cases the noice or muting is not severe. In a couple of instance, the noise is substantial (#3052, #3004 and #3017) and replacing an aberrant rater may be a valuable solution.

Rater Analysis

Rater Quality

A more detailed analysis of rater's behavior is found in table 7.2.1, under the title of the Raters Measurement Report for the Reading Excellence Act Tutorial Assistance survey results. Raters are presented in descending order of severity; in other words, Rater **#24** and **#61** are the most severe and Rater **#144** and **#39** are the least severe, as was seen in Table 7.2.1.

Reading Excellence Act Tutorial Assistance 04-30-2000 01:27:25
Table 7.2.1 raters Measurement Report (arranged by mN).

Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Num raters
287	130	2.2	2.21	.75	.09	0.9	0	0.8	-1	**24 24**
346	130	2.7	2.22	.74	.09	1.1	1	1.2	1	**61 61**
361	130	2.8	2.33	.63	.08	0.8	-1	0.8	-1	51 51
340	156	2.2	2.37	.60	.08	1.3	2	1.3	2	10 10
402	130	3.1	2.41	.56	.09	1.3	2	1.4	3	115 115
650	234	2.8	2.50	.48	.06	**1.6**	**5**	**1.7**	**6**	**27 27****
332	130	2.6	2.52	.47	.09	**0.7**	**-3**	**0.7**	**-2**	**1 1***
337	130	2.6	2.53	.45	.09	**0.5**	**-5**	**0.5**	**-5**	**59 59***
493	182	2.7	2.55	.44	.07	0.8	-1	0.9	0	63 63
288	104	2.8	2.56	.43	.10	**0.6**	**-3**	**0.7**	**-2**	**62 62***
391	130	3.0	2.60	.39	.08	0.9	-1	0.9	-1	11 11
373	156	2.4	2.61	.39	.08	1.1	0	1.3	2	105 105
391	130	3.0	2.62	.38	.08	0.9	0	0.9	-1	40 40
324	130	2.5	2.61	.38	.09	**0.5**	**-4**	**0.6**	**-3**	**118 118***
467	156	3.0	2.63	.37	.08	0.8	-1	0.8	-2	17 17
442	156	2.8	2.64	.36	.08	**0.7**	**-3**	**0.7**	**-3**	**26 26***
547	182	3.0	2.67	.33	.07	1.2	2	1.2	2	87 87
429	156	2.8	2.69	.32	.08	**2.1**	**7**	**1.9**	**6**	**50 50****
394	130	3.0	2.69	.31	.09	1.0	0	1.0	0	46 46
448	156	2.9	2.70	.31	.08	**0.6**	**-4**	**0.6**	**-4**	**99 99***
397	156	2.5	2.70	.31	.08	**0.7**	**-3**	**0.7**	**-2**	**175 175***
549	182	3.0	2.71	.30	.07	**0.6**	**-5**	**0.6**	**-5**	**73 73***
256	78	3.3	2.71	.30	.11	1.0	0	1.0	0	122 122
345	130	2.7	2.72	.29	.09	0.8	-1	0.9	-1	82 82
372	130	2.9	2.73	.28	.08	0.8	-1	0.8	-1	8 8
420	156	2.7	2.74	.27	.08	0.9	0	0.9	0	85 85
494	156	3.2	2.78	.24	.08	0.9	-1	0.9	-1	6 6
401	156	2.6	2.78	.24	.08	**0.7**	**-2**	**0.7**	**-2**	**176 176***
328	130	2.5	2.82	.21	.09	1.0	0	1.1	0	98 98
363	130	2.8	2.84	.19	.08	**1.5**	**3**	**1.5**	**3**	**12 12****
378	130	2.9	2.85	.18	.08	0.9	-1	0.9	-1	126 126
963	312	3.1	2.85	.18	.05	**1.6**	**6**	**1.6**	**6**	**146 146****
738	286	2.6	2.87	.17	.06	1.0	0	1.0	0	28 28
353	104	3.4	2.86	.17	.09	0.8	-1	0.8	-1	53 53
534	156	3.4	2.86	.17	.08	**0.5**	**-6**	**0.5**	**-6**	**57 57***
344	130	2.6	2.86	.17	.09	**0.7**	**-3**	**0.7**	**-3**	**102 102***
964	312	3.1	2.87	.16	.06	1.1	0	1.0	0	16 16
884	286	3.1	2.87	.16	.06	1.3	3	1.4	3	45 45
327	130	2.5	2.89	.15	.09	1.2	1	1.2	1	19 19
423	130	3.3	2.89	.15	.09	0.8	-1	0.8	-1	65 65
407	130	3.1	2.90	.14	.09	1.0	0	0.9	0	58 58
363	130	2.8	2.89	.14	.08	0.9	-1	0.9	-1	116 116
390	156	2.5	2.91	.13	.08	**1.4**	**3**	**1.4**	**3**	**71 71***
499	156	3.2	2.91	.13	.08	1.1	0	1.1	0	96 96
418	156	2.7	2.91	.13	.08	0.8	-1	0.9	0	100 100
373	130	2.9	2.93	.12	.09	**1.7**	**5**	**1.7**	**5**	**88 88***
422	130	3.2	2.92	.12	.09	0.9	0	0.9	0	129 129
560	182	3.1	2.93	.11	.07	1.0	0	1.0	0	30 30
365	130	2.8	2.94	.10	.08	**0.5**	**-5**	**0.5**	**-5**	**125 125***
440	130	3.4	2.97	.08	.08	1.2	1	1.2	1	52 52
356	130	2.7	2.97	.08	.09	**1.7**	**5**	**1.7**	**4**	**54 54****
436	156	2.8	2.97	.08	.08	1.1	0	1.1	0	79 79
338	104	3.3	2.98	.08	.09	**0.6**	**-3**	**0.6**	**-3**	**109 109***
378	130	2.9	2.98	.07	.09	0.9	-1	0.9	0	80 80
418	156	2.7	2.98	.07	.08	0.9	-1	0.9	-1	170 170
458	156	2.9	2.99	.07	.08	1.3	2	1.3	2	179 179
548	182	3.0	3.02	.04	.07	**0.7**	**-3**	**0.7**	**-2**	**86 86***
526	156	3.4	3.02	.04	.08	**0.7**	**-3**	**0.7**	**-3**	**171 171***
510	156	3.3	3.05	.02	.08	1.0	0	1.0	0	2 2
482	156	3.1	3.07	.00	.08	0.8	-2	0.8	-2	4 4
322	104	3.1	3.07	.00	.09	**1.8**	**5**	**1.8**	**5**	**21 21****
392	130	3.0	3.07	.00	.09	**0.7**	**-3**	**0.7**	**-2**	**23 23***
429	130	3.3	3.06	.00	.08	**0.6**	**-4**	**0.6**	**-4**	**180 180***
376	104	3.6	3.11	-.04	.09	0.8	-2	0.8	-2	108 108

487	130	3.7	3.12	-.04	.08	0.8	-1	0.8	-2	168	168	
518	156	3.3	3.16	-.07	.08	1.0	0	1.0	0	18	18	
574	182	3.2	3.16	-.07	.07	**1.4**	**3**	**1.3**	**2**	**174**	**174**	**
346	104	3.3	3.16	-.08	.09	1.1	0	1.2	1	140	140	
430	130	3.3	3.17	-.08	.08	1.0	0	1.0	0	177	177	
478	156	3.1	3.18	-.09	.08	**0.7**	**-3**	**0.7**	**-3**	**112**	**112**	*
594	156	3.8	3.17	-.09	.08	1.2	1	1.2	1	130	130	
447	130	3.4	3.18	-.09	.08	**0.6**	**-3**	**0.6**	**-3**	**185**	**185**	*
193	52	3.7	3.19	-.10	.13	**0.6**	**-2**	**0.6**	**-2**	**38**	**38**	*
477	130	3.7	3.18	-.10	.09	1.0	0	1.1	0	93	93	
441	130	3.4	3.20	-.10	.08	1.0	0	1.0	0	132	132	
473	130	3.6	3.21	-.11	.08	0.8	-1	0.8	-1	13	13	
225	78	2.9	3.21	-.12	.11	1.0	0	1.0	0	110	110	
497	130	3.8	3.22	-.12	.09	1.3	2	1.3	2	166	166	
532	156	3.4	3.21	-.12	.08	0.9	-1	0.9	-1	181	181	
471	156	3.0	3.23	-.13	.08	1.3	2	1.3	2	101	101	
508	156	3.3	3.25	-.15	.08	**1.6**	**5**	**1.6**	**5**	**5**	**5**	**
476	130	3.7	3.25	-.15	.08	0.8	-1	0.9	-1	14	14	
362	130	2.8	3.24	-.15	.09	1.2	1	1.1	1	182	182	
205	52	3.9	3.28	-.17	.14	**0.6**	**-2**	**0.6**	**-2**	**42**	**42**	**
526	156	3.4	3.27	-.17	.08	0.8	-2	0.8	-1	123	123	
2045	598	3.4	3.27	-.17	.04	**0.5**	**-9**	**0.5**	**-9**	**136**	**136**	*
583	156	3.7	3.30	-.19	.08	**2.2**	**8**	**2.2**	**8**	**133**	**133**	**
654	182	3.6	3.32	-.21	.07	0.8	-2	0.8	-2	44	44	
526	156	3.4	3.32	-.21	.08	0.9	0	0.9	0	49	49	
554	182	3.0	3.35	-.24	.07	0.9	0	0.9	0	15	15	
490	156	3.1	3.36	-.24	.08	0.9	-1	0.9	-1	104	104	
564	156	3.6	3.39	-.27	.08	**0.7**	**-3**	**0.7**	**-2**	**114**	**114**	*
499	156	3.2	3.41	-.28	.08	**0.7**	**-3**	**0.7**	**-2**	**56**	**56**	*
491	130	3.8	3.42	-.29	.09	0.8	-1	0.8	-1	9	9	
475	130	3.7	3.42	-.29	.08	1.0	0	1.0	0	90	90	
385	130	3.0	3.42	-.29	.08	**1.6**	**4**	**1.6**	**4**	**172**	**172**	**
547	156	3.5	3.44	-.31	.08	**0.5**	**-5**	**0.5**	**-5**	**66**	**66**	*
589	156	3.8	3.44	-.31	.08	0.9	0	0.9	0	95	95	
546	182	3.0	3.52	-.38	.08	**1.5**	**3**	**1.3**	**2**	**25**	**25**	**
291	78	3.7	3.53	-.39	.11	**1.5**	**2**	**1.5**	**2**	**89**	**89**	**
587	156	3.8	3.53	-.39	.08	**1.6**	**4**	**1.5**	**4**	**178**	**178**	**
472	130	3.6	3.56	-.41	.09	**1.5**	**3**	**1.5**	**3**	**167**	**167**	**
499	130	3.8	3.56	-.42	.09	**1.5**	**3**	**1.5**	**3**	**81**	**81**	**
491	130	3.8	3.57	-.43	.09	1.0	0	1.0	0	22	22	
293	78	3.8	3.59	-.44	.11	**1.5**	**2**	**1.5**	**2**	**7**	**7**	**
397	104	3.8	3.59	-.44	.10	**0.7**	**-2**	**0.7**	**-2**	**92**	**92**	*
547	130	4.2	3.68	-.52	.09	**0.7**	**-3**	**0.6**	**-3**	**141**	**141**	*
482	130	3.7	3.69	-.53	.09	**0.7**	**-2**	**0.7**	**-2**	**143**	**143**	*
298	104	2.9	3.77	-.60	.10	1.3	1	1.3	1	67	67	
724	182	4.0	3.81	-.64	.07	1.1	0	1.1	1	20	20	
186	52	3.6	3.82	-.65	.14	**0.5**	**-3**	**0.5**	**-3**	**169**	**169**	*
566	130	4.4	4.00	-.82	.09	1.1	0	1.1	0	75	75	
243	52	4.7	4.17	-.98	.16	0.9	0	0.9	0	**39**	**39**	
383	78	4.9	4.59	-1.46	.14	**0.3**	**-5**	**0.3**	**-5**	**144**	**144**	*
Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Num raters		
462.7	146.6	3.2	3.07	.00	.08	1.0	-0.4	1.0	-0.3	Mean (Count: 114)		
198.5	59.8	0.5	0.40	.35	.02	0.4	3.1	0.3	3.1	S.D.		

RMSE (Model) .09 Adj S.D. .34 Separation 3.89 Reliability .94
Fixed (all same) chi-square: 1615.2 d.f.: 113 significance: .00
Random (normal) chi-square: 110.2 d.f.: 112 significance: .53

Note : *** Muted**
**** Noise**

The FACETS analysis provides a number of indications of the magnitude of the differences among elements of a facet: in this case, in severity among raters. These are

the ***RMSE, Reliability***, ***Separation Index*** and ***Fixed*** (all same) and ***Random*** (normal) ***Chi Square, Infit*** and O***utfit*** statistics.

Roat Mean Square Standard Error, ***RMSE***, is produced for all non-extreme measures over the raters. RMSE score, .09, shows that rater's error is very low. After raters error variance has been adjusted for measurement error, adjusted standard deviation found below the 1.0 (.33), thus anr rater score is likely to be with .09 points for each item.

The ***Reliability*** statistics provided by the FACETS analysis indicates the degree to which the analysis reliability distinguishes between different levels of quality among the elements of the facets (in this case, the different Raters). Table 7.2.1 shows that the reliability for raters is .94. This indicates that the analysis is fairly reliably separating raters into approximately 4 different levels of leniency and severity.

The ***Separation index*** is the ratio of the corrected standard deviation of elements measures (in this case, Raters) to the root mean-square estimation error. If the Raters were equally severe, the standard deviation of the Raters difficulty estimates should be equal to or smaller than the mean estimation error of the entire data set. However, the Rater Seperation Index is 3.87, indicating that the variance among raters is about four times the error of estimates.

Finally, the ***Fixed Chi-square*** tests the null hypothesis that all of the elements of the facet are equal. The Chi-square of 1615.2 with 113 df. Is significant at p = .00, indication that the null hypothesis must be rejected; in other words, the reaters are not equally severe.

The FACETS analysis also provides two measures of fit, or consistency: the infit and the outfit scores. The infit is the weighted mean-squared residual that is sensitive to

unexpected responses near the point where decisions are being made. Less then 0.6 indicates muting: too little variation, lack of independence. More than 1.4 indicates noise: unmodelled excess variation. On the other hand, the outfit is the the unweighted mean-squared residual and is sensitive to extreme scores. This fit statistics has the same form as infit, but is the conventional mean-square which is more sensitive to outliers.

In addition to the mean squares, FACETS provides standardized infit and outfit statistics, which have an expected mean of 0 and standard deviation of 1. These statistics are useful for comparing the elements of a facet with each other, as they show the degree of variability in addition raters' ratings relative to the amount of variability in the entire set. Standardized fit statistics 2 or 3 or less than –2 or –3 are generally signs of misfit.

Applying these information on the Table 7.2.1, we can see that 23 out of 114 raters had either high (>1.4) or low (<0.6) infit statistics and/or either high (>1.4) or low (<.0.6) outfit statistics: These readers have high (>1.4) infit/outfit statistics: **# 27, #50, #12, #146, #71, #88, #54, #21, #5, #133, #172, #25, #89, #178, #81, #167**. Besides these readers, 7 readers have low infit/outfit statistics (< 0.6). **#118, #57, #125, #66, #92, #169, #144.** (Please check the table 7.2.1 for more detail). The low infit and outfit scores are less concern than the high infit and outfit scores. The low scores tend to reflect flat lining and lack of discrimination. The high infit and outfit scores are probably a concern for raters **#350, #321** and **#133.** Those raters might profit by addition training. It also might br prudent to remove these raters from the calibration of the application.

Item Analysis

Item Quality

A more detailed analysis of items are found in table 7.3.1, the Items Measurement Report for the Reading Excellence Act Tutorial Assistance survey results. Items are presented in descending order of quality to endorse; in other words, item #18 and #10 are the most difficult to endorse and items #3 and #4 are the least difficult items to endorse.

Reading Excellence Act Tutorial Assistance 04-30-2000 01:27:25
Table 7.3.1 items Measurement Report (arranged by mN).

Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Model Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Nu items
1643	643	2.6	2.39	.58	.04	1.0	0	1.0	0	**18 18**
1684	643	2.6	2.46	.51	.04	1.1	2	1.2	2	**10 10**
1755	643	2.7	2.59	.41	.04	0.9	-1	1.0	0	13 13
1769	643	2.8	2.61	.38	.04	1.0	0	1.0	0	8 8
1788	643	2.8	2.64	.36	.04	0.9	-1	0.9	-2	16 16
1794	643	2.8	2.65	.35	.04	1.0	0	1.0	0	21 21
1826	643	2.8	2.71	.30	.04	1.0	0	1.0	0	12 12
1823	643	2.8	2.70	.30	.04	**0.8**	**-4**	**0.8**	**-4**	**17 17***
1904	643	3.0	2.85	.18	.04	**1.7**	**9**	**1.8**	**9**	**14 14****
1926	643	3.0	2.88	.15	.04	**0.4**	**-9**	**0.4**	**-9**	**26 26****
1943	643	3.0	2.91	.13	.04	1.0	0	1.0	0	23 23
1961	643	3.0	2.95	.10	.04	1.0	0	1.0	0	9 9
2009	643	3.1	3.03	.03	.04	0.9	-1	1.0	0	11 11
2018	643	3.1	3.05	.02	.04	**1.3**	**6**	**1.3**	**5**	**22 22****
2034	643	3.2	3.07	-.01	.04	0.9	-1	0.9	-1	20 20
2055	643	3.2	3.11	-.04	.04	0.9	0	1.0	0	2 2
2090	643	3.3	3.17	-.09	.04	0.9	-1	0.9	-1	15 15
2102	643	3.3	3.19	-.10	.04	1.0	0	0.9	-1	19 19
2108	643	3.3	3.21	-.11	.04	1.0	0	1.0	0	24 24
2195	643	3.4	3.36	-.24	.04	1.1	1	1.0	0	25 25
2201	643	3.4	3.37	-.25	.04	1.0	0	1.0	0	7 7
2243	643	3.5	3.44	-.31	.04	0.9	-2	0.9	-2	6 6
2332	643	3.6	3.60	-.45	.04	1.0	0	1.0	0	5 5
2365	643	3.7	3.65	-.50	.04	**1.3**	**5**	**1.3**	**4**	**1 1****
2464	643	3.8	3.82	-.65	.04	0.8	-3	0.8	-3	**4 4**
2716	643	4.2	4.24	-1.06	.04	0.9	-1	0.9	-1	**3 3**
Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Model Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Nu items
2028.8	643.0	3.2	3.06	.00	.04	1.0	-0.3	1.0	-0.2	Mean (Count: 26)
250.5	0.0	0.4	0.44	.38	.00	0.2	3.4	0.2	3.3	S.D.

RMSE (Model) .04 Adj S.D. .37 Separation 9.66 Reliability .99
Fixed (all same) chi-square: 2324.2 d.f.: 25 significance: .00
Random (normal) chi-square: 25.0 d.f.: 24 significance: .41

*** : Muting**
**** : Noise**

As we pointed out before, The FACETS analysis provides a number of indications of the magnitude of the differences among elements of a facet: in this case, in quality of the items. These are the ***RMSE, Reliability***, ***Separation Index*** and ***Fixed*** (all same) and ***Random*** (normal) ***Chi Square***, ***Infit*** and ***Outfit*** statistics.

Roat Mean Square Standard Error, ***RMSE***, is produced for all non-extreme measures over the reaters. RMSE score, .04, shows that item quality error is very low. After item quality error variance has been adjusted for measurement error, adjusted standard deviation found below the 1.0 (.37, respectively).

The ***Reliability*** statistics is Rash equivalent to the KR-20 or Cronbach Alpha statistics, which is the ratio of "True variance" to "Observed variance". Reliability provided by the FACETS shows how different the measures are, which may or may not indicate how "good" the test is. High (near 1.0) item reliabilities are preferred. In this case, the reliability is .99, indicating that the analysis is very reliably separating items into different levels of difficulty.

The ***Separation index*** is the ratio of the corrected standard deviation of elements measures (in this case, items) to the root mean-square estimation error. If the Items were equal difficulty, the standard deviation of the item quality estimates should be equal to or smaller than the mean estimation error of the entire data set. However, the Item Seperation Index is 9.66, indicating that the variance among items is about ten times the error of estimates.

Finally, the ***Fixed Chi-square*** tests the null hypothesis that all of the elements of the facet are equal. The Chi-square of 2324.3 with 25 df. is significant at p = .00, indication

that the null hypothesis must be rejected; in other words, the items are not equal difficulty.

The FACETS analysis also provides two measures of fit, or consistency: the infit and the outfit. The infit is the weighted mean-squared residual that is sensitive to unexpected responses near the point where decisions are being made. Less then 0.6 indicates muting: too little variation, lack of independence. More than 1.4 indicates noise: unmodelled excess variation. On the other hand, the outfit is the the unweighted mean-squared residual and is sensitive to extreme scores. This fit statistics has the same form as infit, but is the conventional mean-square which is more sensitive to outliers.

In addition to the mean squares, FACETS provides standardized infit and outfit statistics, which have an expected mean of 0 and standard deviation of 1. These statistics are useful for comparing the elements of a facet with each other, as they show the degree of variability in addition items' ratings relative to the amount of variability in the entire set. Standardized fit statistics 2 or 3 or less than –2 or –3 are generally signs of misfit.

Applying these information on the Table 7.3.1, we can see that 1 out of 26 items had either high (>1.4) or low (< 0.6) infit/outfit statistics: Item **#14.** These statistics indicate that this item was not consistent with the estimated ability measures of the applications, and that the scores for this item may not be stable.

Application/Rater Interaction

A z score above 2.0 or below –2.0 would indicate an interaction effect. According to the Bias/Interaction report in FACETS analysis in Table 13.1.1, there were several raters who seemed to be too lenient or too severe on certain applications. Z scores in this

bias analysis ranged form – 8.14 to 6.0. many of this interactions effect came from the calibration (#3) application (used in the process of training the raters).

Reading Excellence Act Tutorial Assistance 04-30-2000 01:27:25
Table 13.1.1 Bias/Interaction Calibration Report (arranged by mN).

Bias/Interaction analysis specified by Model: ?B,?B,?,RATINGS

Obsvd Score	Exp. Score	Obsvd Count	Obs-Exp Average	Bias+ Measure	Model S.E.	Z-Score	Infit MnSq	Outfit MnSq	Sq	Num	app	measr	Num	rate	measr
31	46.2	26	-.59	1.43	.45	3.22	1.1	1.1	294	3101	3101	-1.68	71	71	.13
58	**92.7**	**26**	**-1.33**	**1.20**	**.20**	**6.00**	**1.0**	**0.9**	**222**	**3**	**3**	**.00**	**52**	**52**	**.08**
49	76.0	26	-1.04	1.04	.22	4.65	0.5	0.6	217	3	3	.00	51	51	.63
34	47.2	26	-.51	1.02	.35	2.88	0.7	0.6	48	3098	3098	-1.16	10	10	.60
59	88.3	26	-1.13	1.01	.20	5.11	0.6	0.6	357	3016	3016	-.52	90	90	-.29
34	45.7	26	-.45	.94	.35	2.65	0.6	0.7	86	3101	3101	-1.68	16	16	.16
59	84.5	26	-.98	.88	.20	4.46	0.9	0.9	566	3064	3064	-.40	168	168	-.04
36	48.8	26	-.49	.87	.32	2.74	0.7	0.8	80	3037	3037	-1.52	16	16	.16
34	44.5	26	-.40	.86	.35	2.44	1.4	1.1	137	3087	3087	-2.29	25	25	-.38
56	79.5	26	-.90	.84	.20	4.12	1.7	1.8	146	3025	3025	-.04	27	27	.48
47	66.7	26	-.76	.82	.23	3.56	1.7	1.8	163	3106	3106	-.79	28	28	.17
51	72.2	26	-.82	.81	.22	3.77	0.4	0.4	156	3067	3067	-.60	28	28	.17
78	101.6	26	-.91	.81	.18	4.43	0.2	0.2	511	3028	3028	.07	136	136	-.17
61	84.4	26	-.90	.80	.19	4.13	0.5	0.5	158	3073	3073	-.19	28	28	.17
51	71.9	26	-.80	.80	.22	3.72	1.2	1.1	547	3067	3067	-.60	146	146	.18
71	94.1	26	-.89	.77	.18	4.19	0.2	0.2	179	3107	3107	.35	40	40	.38
64	86.8	26	-.88	.77	.19	4.05	0.3	0.3	461	3018	3018	.03	122	122	.30
76	98.2	26	-.85	.75	.18	4.10	0.4	0.4	59	3	3	.00	13	13	-.11
52	71.8	26	-.76	.75	.21	3.53	1.3	1.3	300	3074	3074	-.48	73	73	.30
57	78.2	26	-.82	.75	.20	3.75	0.8	0.8	445	3	3	.00	115	115	.56
62	83.3	26	-.82	.73	.19	3.77	0.7	0.7	49	3	3	.00	11	11	.39
55	75.1	26	-.77	.73	.21	3.57	1.3	1.3	168	3052	3052	-.55	30	30	.11
69	90.7	26	-.84	.73	.19	3.90	0.3	0.3	450	3	3	.00	116	116	.14
74	95.5	26	-.83	.72	.18	3.93	0.5	0.5	117	3001	3001	-.41	22	22	-.43
60	80.9	26	-.80	.72	.20	3.70	0.6	0.5	151	3095	3095	.01	27	27	.48
53	72.1	26	-.73	.72	.21	3.41	1.3	1.3	631	3071	3071	-.88	181	181	-.12
43	57.4	26	-.55	.71	.25	2.81	1.5	1.2	212	3023	3023	-.98	50	50	.32
86	106.0	26	-.77	.71	.18	3.87	0.4	0.5	526	3047	3047	.33	140	140	-.08
93	111.1	26	-.70	.68	.19	3.65	1.1	1.1	321	3070	3070	.19	81	81	-.42
70	90.2	26	-.78	.67	.19	3.64	1.1	1.1	189	3	3	.00	45	45	.16
89	107.7	26	-.72	.67	.18	3.67	0.6	0.7	368	3078	3078	.37	93	93	-.10
52	69.3	26	-.66	.67	.21	3.13	0.8	0.8	431	3023	3023	-.98	110	110	-.12
51	67.9	26	-.65	.66	.22	3.08	0.4	0.4	157	3068	3068	-.74	28	28	.17
99	115.6	26	-.64	.66	.19	3.47	1.2	1.2	481	3029	3029	.92	129	129	.12
91	109.2	26	-.70	.66	.18	3.61	0.4	0.4	533	3	3	.00	143	143	-.53
86	104.5	26	-.71	.65	.18	3.57	1.0	1.0	196	3062	3062	.51	45	45	.16
78	97.2	26	-.74	.64	.18	3.54	0.5	0.5	524	3	3	.00	140	140	-.08
44	57.0	26	-.50	.63	.25	2.56	1.2	1.3	234	3055	3055	-1.24	54	54	.08
77	95.9	26	-.73	.63	.18	3.49	0.2	0.2	503	3006	3006	-.13	136	136	-.17
74	92.9	26	-.73	.63	.18	3.44	0.5	0.5	570	3	3	.00	170	170	.07
75	93.9	26	-.73	.63	.18	3.45	0.4	0.4	576	3	3	.00	171	171	.04
73	91.6	26	-.72	.62	.18	3.38	0.8	0.8	618	3013	3013	-.05	179	179	.07
93	109.5	26	-.63	.61	.19	3.28	0.3	0.3	505	3014	3014	.37	136	136	-.17
86	103.1	26	-.66	.60	.18	3.27	0.8	0.8	38	3	3	.00	9	9	-.29
54	70.0	26	-.61	.60	.21	2.90	0.4	0.4	220	3057	3057	-.20	51	51	.63
88	104.9	26	-.65	.60	.18	3.27	0.4	0.4	447	3029	3029	.92	115	115	.56
72	89.4	26	-.67	.58	.18	3.14	1.1	1.1	54	3	3	.00	12	12	.19
80	97.2	26	-.66	.58	.18	3.19	0.2	0.2	508	3020	3020	-.09	136	136	-.17
45	57.2	26	-.47	.58	.24	2.41	1.1	1.1	552	3104	3104	-1.13	146	146	.18
62	78.5	26	-.63	.57	.19	2.94	2.2	2.1	114	3052	3052	-.55	21	21	.00
47	59.5	26	-.48	.56	.23	2.41	0.8	0.8	235	3105	3105	-1.14	54	54	.08
100	113.9	26	-.53	.55	.19	2.87	0.2	0.2	510	3026	3026	.56	136	136	-.17
74	90.2	26	-.62	.54	.18	2.95	0.5	0.5	76	3	3	.00	16	16	.16
75	91.1	26	-.62	.54	.18	2.93	0.6	0.6	393	3	3	.00	100	100	.13

66	81.9	26	-.61	.54	.19	2.85	2.8	2.9	551	3099	3099	-.26	146	146	.18
82	97.7	26	-.61	.53	.18	2.94	0.8	0.8	373	3030	3030	-.21	95	95	-.31
52	65.3	26	-.51	.53	.21	2.47	1.4	1.3	404	3088	3088	-1.14	101	101	-.13
85	100.5	26	-.60	.53	.18	2.93	1.5	1.5	498	3045	3045	.00	133	133	-.19
74	89.8	26	-.61	.52	.18	2.86	0.2	0.3	502	3005	3005	-.35	136	136	-.17
71	86.3	26	-.59	.51	.18	2.75	0.9	0.9	323	3	3	.00	82	82	.29
79	94.4	26	-.59	.51	.18	2.83	0.4	0.4	529	3012	3012	-.54	141	141	-.52
62	76.8	26	-.57	.51	.19	2.66	2.2	2.2	586	3082	3082	-.90	172	172	-.29
59	72.9	26	-.53	.50	.20	2.52	0.4	0.4	494	3084	3084	-.84	132	132	-.10
44	53.8	26	-.38	.49	.25	2.01	2.3	2.3	216	3104	3104	-1.13	50	50	.32
97	109.6	26	-.49	.48	.19	2.53	0.3	0.3	425	3062	3062	.51	108	108	-.04
90	103.4	26	-.52	.48	.18	2.59	1.1	1.1	486	3038	3038	.22	130	130	-.09
47	57.3	26	-.40	.47	.23	2.04	0.7	0.7	16	3059	3059	-1.31	4	4	.00
49	59.9	26	-.42	.47	.22	2.11	1.5	1.6	317	3104	3104	-1.13	80	80	.07
51	62.4	26	-.44	.47	.22	2.16	1.0	1.0	386	3082	3082	-.90	98	98	.21
85	98.6	26	-.52	.46	.18	2.55	1.3	1.3	85	3090	3090	.29	16	16	.16
76	89.9	26	-.53	.46	.18	2.52	0.3	0.3	242	3	3	.00	57	57	.17
82	95.6	26	-.52	.46	.18	2.52	0.2	0.2	438	3081	3081	-.07	112	112	-.09
64	77.2	26	-.51	.45	.19	2.38	0.3	0.4	109	3055	3055	-1.24	20	20	-.64
53	64.3	26	-.44	.45	.21	2.13	0.4	0.4	128	3040	3040	-.28	24	24	.75
90	102.8	26	-.49	.45	.18	2.47	0.4	0.4	236	3	3	.00	56	56	-.28
60	72.8	26	-.49	.45	.20	2.32	0.7	0.6	258	3	3	.00	61	61	.74
101	112.4	26	-.44	.45	.19	2.33	1.3	1.2	374	3079	3079	.35	95	95	-.31
84	97.4	26	-.51	.45	.18	2.50	0.4	0.4	484	3	3	.00	130	130	-.09
72	85.6	26	-.52	.45	.18	2.46	0.4	0.4	609	3064	3064	-.40	177	177	-.08
69	82.1	26	-.50	.44	.19	2.36	1.3	1.3	23	3097	3097	-.58	5	5	-.15
95	107.0	26	-.46	.44	.19	2.36	1.4	1.4	30	3	3	.00	7	7	-.44
79	92.0	26	-.50	.43	.18	2.38	1.1	1.1	149	3066	3066	.38	27	27	.48
88	100.4	26	-.48	.43	.18	2.36	0.8	0.9	201	3000	3000	.51	46	46	.31
71	84.0	26	-.50	.43	.18	2.35	0.8	0.8	261	3078	3078	.37	61	61	.74
105	115.2	26	-.39	.42	.20	2.14	0.6	0.6	363	3069	3069	.34	92	92	-.44
100	110.9	26	-.42	.42	.19	2.21	0.5	0.6	442	3031	3031	.33	114	114	-.27
71	83.6	26	-.49	.42	.18	2.27	0.4	0.4	455	3	3	.00	118	118	.38
71	83.4	26	-.48	.41	.18	2.24	1.9	1.9	400	3021	3021	-.52	101	101	-.13
56	66.8	26	-.42	.41	.20	2.01	0.4	0.4	473	3084	3084	-.84	125	125	.10
78	90.1	26	-.47	.40	.18	2.21	0.9	0.9	21	3058	3058	-.31	5	5	-.15
96	106.9	26	-.42	.40	.19	2.15	0.8	0.8	184	3019	3019	.23	44	44	-.21
87	98.5	26	-.44	.39	.18	2.16	2.0	1.9	553	3	3	.00	166	166	-.12
86	74.8	26	.43	-.37	.18	-2.04	0.7	0.7	51	3040	3040	-.28	11	11	.39
87	75.9	26	.43	-.37	.18	-2.02	0.4	0.4	188	3096	3096	-.85	44	44	-.21
91	79.5	26	.44	-.38	.18	-2.09	0.8	0.8	183	3008	3008	-.73	44	44	-.21
86	74.6	26	.44	-.38	.18	-2.07	0.7	0.7	191	3016	3016	-.52	45	45	.16
96	84.1	26	.46	-.40	.19	-2.16	0.5	0.5	225	3046	3046	-.29	52	52	.08
59	49.6	26	.36	-.41	.20	-2.05	2.0	2.0	44	3017	3017	-1.04	10	10	.60
103	91.5	26	.44	-.41	.19	-2.12	0.6	0.6	479	3	3	.00	129	129	.12
94	81.8	26	.47	-.41	.19	-2.20	0.4	0.4	525	3016	3016	-.52	140	140	-.08
90	77.2	26	.49	-.42	.18	-2.31	0.5	0.5	67	3068	3068	-.74	14	14	-.15
108	96.6	26	.44	-.42	.20	-2.12	1.4	1.4	192	3038	3038	.22	45	45	.16
98	85.7	26	.47	-.42	.19	-2.23	0.6	0.6	200	3	3	.00	46	46	.31
65	53.9	26	.43	-.43	.19	-2.29	0.4	0.4	75	3101	3101	-1.68	15	15	-.24
58	48.3	26	.37	-.43	.20	-2.17	0.4	0.4	159	3077	3077	-1.54	28	28	.17
109	97.7	26	.44	-.43	.20	-2.12	0.3	0.3	365	3	3	.00	93	93	-.10
78	65.5	26	.48	-.43	.18	-2.34	0.2	0.2	523	3098	3098	-1.16	136	136	-.17
99	86.3	26	.49	-.44	.19	-2.31	0.6	0.6	154	3056	3056	-.13	28	28	.17
86	72.8	26	.51	-.44	.18	-2.39	0.4	0.4	453	3067	3067	-.60	116	116	.14
111	99.0	26	.46	-.46	.20	-2.26	0.3	0.3	60	3018	3018	.03	13	13	-.11
102	88.4	26	.52	-.47	.19	-2.47	0.1	0.2	520	3072	3072	-.39	136	136	-.17
94	79.6	26	.55	-.48	.19	-2.59	1.7	1.7	259	3019	3019	.23	61	61	.74
97	82.7	26	.55	-.48	.19	-2.57	0.4	0.4	385	3057	3057	-.20	98	98	.21
52	42.9	26	.35	-.48	.21	-2.26	0.5	0.5	464	3034	3034	-2.18	123	123	-.17
61	49.4	26	.45	-.49	.19	-2.54	1.2	1.2	270	3036	3036	-1.21	63	63	.44
104	90.1	26	.54	-.49	.19	-2.54	0.8	0.8	394	3025	3025	-.04	100	100	.13
111	98.3	26	.49	-.49	.20	-2.39	1.3	1.3	430	3	3	.00	110	110	-.12
108	94.4	26	.52	-.50	.20	-2.51	0.9	0.9	361	3015	3015	-.46	92	92	-.44
110	96.5	26	.52	-.51	.20	-2.52	0.5	0.5	413	3073	3073	-.19	104	104	-.24
96	80.5	26	.60	-.52	.19	-2.78	0.8	0.7	144	3	3	.00	27	27	.48
43	36.3	26	.26	-.52	.25	-2.06	1.0	1.1	195	3061	3061	-2.36	45	45	.16
115	102.1	26	.49	-.52	.21	-2.46	0.8	0.9	351	3078	3078	.37	88	88	.12
115	102.0	26	.50	-.52	.21	-2.49	0.5	0.5	580	3090	3090	.29	171	171	.04
106	90.7	26	.59	-.55	.20	-2.79	0.4	0.4	63	3099	3099	-.26	13	13	-.11
104	88.2	26	.61	-.56	.19	-2.87	0.1	0.1	516	3064	3064	-.40	136	136	-.17

74	58.1	26	.61	-.57	.18	-3.11	2.3	2.2	152	3106	3106	-.79	27	27	.48
117	103.0	26	.54	-.57	.21	-2.68	0.7	0.8	582	3	3	.00	172	172	-.29
118	104.1	26	.53	-.58	.22	-2.69	1.9	1.9	542	3000	3000	.51	146	146	.18
46	37.4	26	.33	-.60	.23	-2.54	0.5	0.5	398	3087	3087	-2.29	100	100	.13
124	110.9	26	.50	-.61	.23	-2.65	1.3	1.2	285	3	3	.00	67	67	-.60
113	97.1	26	.61	-.61	.21	-2.97	1.4	1.2	587	3	3	.00	174	174	-.07
96	77.6	26	.71	-.62	.19	-3.30	0.7	0.7	243	3001	3001	-.41	57	57	.17
66	50.6	26	.59	-.62	.19	-3.27	0.8	0.8	572	3037	3037	-1.52	170	170	.07
78	59.9	26	.70	-.63	.18	-3.48	0.7	0.7	415	3092	3092	-1.44	104	104	-.24
126	112.9	26	.51	-.63	.23	-2.70	2.0	2.1	555	3026	3026	.56	166	166	-.12
101	81.9	26	.73	-.65	.19	-3.42	1.0	1.1	50	3013	3013	-.05	11	11	.39
54	41.7	26	.47	-.65	.21	-3.13	1.6	1.6	145	3009	3009	-1.61	27	27	.48
85	65.6	26	.75	-.65	.18	-3.59	1.3	1.3	584	3033	3033	-1.28	172	172	-.29
104	85.0	26	.73	-.66	.19	-3.42	2.0	2.0	233	3027	3027	-.26	54	54	.08
61	45.7	26	.59	-.69	.19	-3.54	0.3	0.4	162	3101	3101	-1.68	28	28	.17
115	97.4	26	.68	-.69	.21	-3.28	1.9	2.0	214	3050	3050	.40	50	50	.32
118	101.2	26	.65	-.69	.22	-3.19	0.5	0.5	360	3081	3081	-.07	90	90	-.29
122	105.6	26	.63	-.71	.22	-3.19	1.7	1.6	353	3	3	.00	89	89	-.39
44	35.1	26	.34	-.71	.25	-2.91	0.5	0.5	396	3060	3060	-2.52	100	100	.13
118	100.5	26	.67	-.71	.22	-3.31	1.3	1.3	548	3078	3078	.37	146	146	.18
101	80.0	26	.81	-.72	.19	-3.75	0.4	0.4	241	3106	3106	-.79	56	56	-.28
88	66.2	26	.84	-.73	.18	-4.00	0.8	0.8	621	3082	3082	-.90	179	179	.07
104	82.3	26	.84	-.75	.19	-3.88	0.1	0.1	517	3067	3067	-.60	136	136	-.17
99	76.8	26	.85	-.75	.19	-3.96	0.2	0.2	536	3088	3088	-1.14	143	143	-.53
117	98.2	26	.72	-.75	.21	-3.50	0.5	0.5	628	3	3	.00	181	181	-.12
92	69.2	26	.88	-.76	.18	-4.12	0.6	0.6	10	3084	3084	-.84	2	2	.02
113	92.8	26	.78	-.76	.21	-3.69	0.6	0.6	224	3045	3045	.00	52	52	.08
114	93.5	26	.79	-.78	.21	-3.75	1.0	1.0	448	3062	3062	.51	115	115	.56
122	103.2	26	.72	-.81	.22	-3.61	0.5	0.5	493	3070	3070	.19	132	132	-.10
64	45.1	26	.73	-.83	.19	-4.36	0.5	0.5	262	3105	3105	-1.14	61	61	.74
109	86.0	26	.88	-.83	.20	-4.11	0.8	0.8	460	3	3	.00	122	122	.30
124	105.2	26	.72	-.83	.23	-3.64	0.3	0.3	565	3031	3031	.33	168	168	-.04
115	93.1	26	.84	-.84	.21	-3.98	1.3	1.3	20	3030	3030	-.21	5	5	-.15
122	102.3	26	.76	-.84	.22	-3.75	3.6	3.3	500	3094	3094	.07	133	133	-.19
109	84.3	26	.95	-.88	.20	-4.39	0.7	0.7	92	3095	3095	.01	17	17	.37
131	113.7	26	.66	-.89	.25	-3.55	0.7	0.7	136	3069	3069	.34	25	25	-.38
123	101.7	26	.82	-.91	.23	-4.02	1.0	0.9	110	3064	3064	-.40	20	20	-.64
127	107.4	26	.75	-.91	.24	-3.85	1.1	1.1	485	3014	3014	.37	130	130	-.09
88	61.0	26	1.04	-.92	.18	-5.04	0.7	0.8	527	3055	3055	-1.24	140	140	-.08
101	73.6	26	1.05	-.93	.19	-4.86	0.2	0.3	519	3071	3071	-.88	136	136	-.17
122	98.8	26	.89	-.97	.22	-4.32	1.3	1.2	399	3	3	.00	101	101	-.13
121	97.3	26	.91	-.97	.22	-4.38	0.7	0.7	606	3	3	.00	177	177	-.08
113	86.5	26	1.02	-.98	.21	-4.72	0.5	0.5	221	3107	3107	.35	51	51	.63
74	48.6	26	.98	-.99	.18	-5.39	0.8	0.8	449	3104	3104	-1.13	115	115	.56
102	72.4	26	1.14	-1.01	.19	-5.23	1.0	0.9	126	3	3	.00	24	24	.75
115	86.9	26	1.08	-1.05	.21	-4.98	0.4	0.4	328	3	3	.00	85	85	.27
102	70.8	26	1.20	-1.06	.19	-5.52	0.1	0.2	509	3023	3023	-.98	136	136	-.17
114	83.8	26	1.16	-1.11	.21	-5.31	1.2	1.1	175	3	3	.00	40	40	.38
121	90.1	26	1.19	-1.22	.22	-5.50	0.7	0.7	153	3	3	.00	28	28	.17
133	106.4	26	1.02	-1.32	.26	-5.12	1.4	1.5	318	3	3	.00	81	81	-.42
138	114.3	26	.91	-1.35	.28	-4.81	1.0	1.0	544	3029	3029	.92	146	146	.18
138	106.5	26	1.21	-1.67	.28	-5.95	1.1	1.1	116	3	3	.00	22	22	-.43
122	**73.6**	**26**	**1.86**	**-1.82**	**.22**	**-8.10**	**0.4**	**0.5**	**81**	**3052**	**3052**	**-.55**	**16**	**16**	**.16**

Obsvd Score	Exp. Score	Obsvd Count	Obs-Exp Average	Bias+ Measure	Model S.E.	Z-Score	Infit MnSq	Outfit MnSq	Sq	Num app		measr	Num rate		measr
82.0	82.0	26.0	.00	.00	.20	.00	0.9	0.9	Mean (Count: 643)						
22.2	19.6	0.0	.40	.41	.03	1.97	0.9	0.5	S.D.						

Fixed (all = 0) chi-square: 2484.3 d.f.: 643 significance: .00

For example, rater #16 with an expected score of 73.4 had an observed score of 122 on application #3052, translating into a z-score of – 8.14. Rater #52 with an expected

score 92.7 had an observed score 58 on application #0003, translating into a z-score of 6.00.

There was an overall statistically significant rater by application interaction effect χ^2 =2484.3, df. = 643, p <.01)

Summary

The FACETS analysis provided an assessment of quality in the applications, raters, and items. The specific indications of quality are *Separation Index*, *Reliability*, *RMSE,* and *Fixed* and *Random Chi-Square*. The FACETS analysis also provided two measures of fit, or consistency on each of the three facets: the *infit* and *outfit.* The infit is the weighted mean-squared residual that is sensitive to unexpected responses within expected response parameters. On the other hand, the outfit is the unweighted mean-squared residual and is sensitive to extreme scores.

In this analysis, the results showed that just 21% of the applications have acceptable quality. 11 out of 106 applications show very high infit statistics (3 > Infit > -3) and 5 out of 106 applications showed very high outfit statistics (.06> outfit > -3). These statistics indicated that most applications were consistent with the estimated quality measures, or their scores were highly predictable of the 11 applications with high infit/outfit, only 3 showed cause for concern.

For raters, results showed that 23 out of 114 raters have been found with high *infit* and *outfit* statistics. Sixteen of the raters had high infit/outfit statistics, but only 3 raters were problematic. None of the raters with low infit/outfit statistics were problematic. These statistics indicated that these raters' ordering of application was generally consistent with the estimated quality measures of the applications.

Finally, only 1 out of 26 items showed high *infit* and *outfit* statistics. This result indicated that this item may contribute minor noise to the overall calibration of the application. In general, the item functions are very well.

The Facet analysis also provides Root Mean Square Standard Error (RMSE) for all non-extreme measures over applications, raters and items (.09, .09 & .04 respectively). These RMSE scores illustrate that applications, rater and item measurement errors are very low. After application, raters and item variances have been adjusted for measurement error, three variances are below the 1.0 (Adj. SD = .67, Adj. SD = .33, Adj. SD = .37 for application, raters and items respectively). The ratio of Ad. SD to RMSE (7.70 for application, 3.89 for raters and 9.66 for items) for application raters and items separation are relatively high due to low RMSEs, indicating high calibration and low error.

The reliability statistics provided by the FACETS analysis indicates the degree to which the analysis reliability distinguishes level of quality among the elements of the application, raters and items. For applications, raters and items, FACETS analysis produce .98, .94 and .99, reliability scores respectively. These reliability scores indicate that the analysis is fairly reliably separating applications, raters and items into different levels of quality.

Conclusion

This study demonstrated the application of a sophisticated assessment procedure in addressing a significant educational problem, i.e., a fair and consistent way to assess

applications into a Reading Excellence program. This procedure has wide applicability, but is currently not well known.

References

Engelhard, J. G. (1992a). The measurement of writing ability with a many-faced Rasch model. *Applied Measurement in Education, 5,* 171-191.

Linacre, M. (1993, April). *Generalizability theory and many-facet Rasch measurement.* Paper presented at the annual meeting of the American Educational Research Association. Atlanta, GA.

Schumacker, R. E. (1996). *Many-faced Rasch model selection criteria: examining resudials and more.* Paper presented at the annual meeting of the American educational research Association, New York, NY.

U. S. Department of Education (2000). The Reading Excellence Program. *(http://www.ed.gov/offices/OESE/REA).*

Ohio Department of Education (1999). *Ohio's Reading Excellence Act Grant Program*. Local reading improvement subgrant tutorial assistance subgrant, Columbus, OH.