

OCTOBER 2003 GRADE 3 READING ACHIEVEMENT TEST

STATISTICAL SUMMARY

N-Count	131,701
Max. Raw Score	49
Raw Score Mean	31.65
Raw Score Standard Deviation	10.45
Raw Score SEM	3.47
Max. Scaled Score	498
Min. Scaled Score	261
Scaled Score Mean	401.87
Scaled Score Standard Deviation	28.86
Scaled Score SEM	9.57
Reliability	.89

These statistics were computed using the entire population of public school (including community school) students who were tested during the October 2003 administration.

CUT SCORE POINTS FOR BASIC, PROFICIENT, ACCELERATED, AND ADVANCED STANDARDS

Standard	Raw Score	Scaled Score
Limited	Below 25	Below 385
Basic	25	385
Proficient	33	400
Accelerated	38	415
Advanced	43	432

Number and Percentage of Students by Performance Levels from October 2003 Administration

Standard	Number of	
	Students	Percent
Limited	34,730	26.4
Basic	26,199	19.9
Proficient	23,183	17.6
Accelerated	26,984	20.5
Advanced	20,605	15.6

EQUATING AND SCALING: HOW RAW SCORES WERE COVERED INTO SCALED SCORES

Test Form Construction

The reading test included items that were field-tested in March 2003 and were used operationally for the first time. Form construction used the item locations from the Rasch/Item Response Theory (IRT) analysis to pre-equate the operational forms.

Common Item Equating

After the October 2003 operational test administration of Grade 3 Reading, the relative difficulty values of the items were computed on the basis of an early return sample (N=10,570), which was selected to be statistically representative of the whole Ohio grade 3 student population. The items had difficulty values in the item bank based on the March 2003 field test; these items served as anchor items in the equating process. Then, the current and the “bank” difficulty values of the anchor items were computed. The mean difference between the current and the bank difficulties of the anchor items is called the equating constant, and this value was added to each item difficulty value on the current test so that the mean difficulties were equal. The item (or items) with the largest absolute difference between the bank values and the equated current values was (or were) eliminated from the anchor set if this difference was greater than 0.3 in absolute value. This procedure was repeated until the largest difference was less than 0.3 to ensure that the anchor items were stable. When the equating was complete, item difficulties from the current administration were directly comparable with those from the bank.

Scaling

Because raw scores do not always have the same meaning across different test administrations, scaled scores are usually used in place of raw scores.

After the test was administered in October 2003, test items were calibrated and equated on the basis of the early return sample and Rasch ability estimates (θ) were computed for each possible raw score. For the test, the proficient standard was scaled to equal 400. After scaling, the basic standard was equal to a scaled score of 385, the proficient standard was equal to a scaled score of 400, the accelerated standard was equal to a scaled score of 415, and the advanced standard was equal to a scaled score of 432.

Raw Score to Scaled Score Conversion Table

Raw Score	Scaled Score
0	261
1	280
2	299
3	310
4	318
5	325
6	330
7	335
8	339
9	343
10	347
11	350
12	353
13	356
14	359
15	362
16	364
17	367
18	369
19	371
20	373
21	376
22	378
23	380
24	382
25	385

Raw Score	Scaled Score
26	386
27	388
28	390
29	392
30	394
31	397
32	399
33	401
34	403
35	406
36	408
37	411
38	415
39	417
40	420
41	424
42	428
43	432
44	438
45	444
46	451
47	462
48	480
49	498