

OCTOBER 2004 GRADE 3 READING TEST ADMINISTRATION

STATISTICAL SUMMARY

These statistics describe the entire population of Ohio 3rd-grade public school students (including community schools) tested during the October 2004 administration.

N-Count	127,737
Max. Raw Score	49
Max. Scaled Score	510
Raw Score Mean	31.24
Raw Score Standard Deviation	9.39
Raw Score SEM	4.28
Scaled Score Mean	402.24
Scaled Score Standard Deviation	28.41
Scaled Score SEM	12.95
Reliability	.89

Cut Score Points for Basic, Proficient, Accelerated, and Advanced Standards

Standard	Raw Score	Scaled Score
Limited	Below 26	Below 385
Basic	26	385
Proficient	32	400
Accelerated	37	415
Advanced	41	432

Percentage of Students by Performance Levels

Standard	Percent
Limited	26.58
Basic	18.88
Proficient	20.50
Accelerated	16.76
Advanced	17.27

Equating and Scaling: How Raw Scores Are Converted Into Scaled Scores

Test Form Construction

The October 2004 Grade 3 Reading Achievement Test is made up of items field-tested before, but not previously used in an operational test form. Item difficulty estimates from the field test administration are used to pre-equate operational forms during the form construction.

Common Item Equating

Following administration of the October 2004 Grade 3 Reading test, we re-estimated item difficulty values using an early return sample of 12,716 students. The early return sample is selected to be statistically representative of all Ohio 3rd-grade public school students. Because we already had item difficulty estimates from field test, all the operational items can potentially serve as anchor items in the equating process. AIR uses a stepwise deletion procedure to calibrate the early return data and calculate the linking constant needed to bring the set of operational items back to the reference scale established during the first operational administration. First, the current difficulty values (from the early return sample) are computed and compared with the “bank” or reference difficulty values. The mean difference between the current and the bank difficulties of the anchor items is called the equating constant. The equating constant is added to each difficulty value for items on the current test so that the mean item difficulties are equal. We then compare the “linked” current values with the original bank values to identify the item with the largest absolute difference between the two values. If the absolute value of the difference is greater than 0.3, the item is eliminated as an anchor item. This procedure is repeated until the largest difference between a linked current value and bank value is less than 0.3. This procedure ensures that the items used to anchor the operational test to the reference scale are stable. When the equating process is complete, item difficulties from the current administration are directly comparable with those from the bank.

Scaling

Because the meaning of raw scores changes across test forms and test administrations, scaled scores are usually used in place of raw scores.

As previously noted, after administering the October 2004 operational test, test items are calibrated and equated on the basis of the early return sample, and Rasch ability estimates (θ) are computed for each possible raw score. The Rasch ability estimates are then transformed to the Ohio 3rd-grade reading scale, which is scaled so that the proficient standard is equal to 400. After scaling, the basic standard on the 3rd-grade reading scale corresponds to a scaled score of 386, the proficient standard to a scaled score of 401, the accelerated standard to a scaled score of 416, and the advanced standard to a scaled score of 432.

Ohio Rounding Rule

When transforming raw scores to scaled scores, if the scaled score nearest to a proficiency level cut score is below the cut score, then the scaled score is rounded up to equal the proficiency level cut score. Otherwise, no special rounding is done. For example, if a raw score is associated with

an observed scaled score of 383.94, and 383.94 is the closest observed scaled score to the basic proficiency level cut score, this value is rounded up to 385, corresponding to the basic proficiency standard. Conversely, if the closest scaled score value to the proficient level cut score is 401.12, no special rounding rules are invoked, because the value is greater than the cut score.

Raw Score to Scaled Score Conversion Table

Raw Score	Scaled Score	Raw Score	Scaled Score
0	256	25	383
1	275	26	386
2	293	27	388
3	305	28	391
4	313	29	393
5	320	30	396
6	326	31	398
7	331	32	401
8	335	33	404
9	339	34	406
10	343	35	409
11	346	36	412
12	350	37	416
13	353	38	419
14	356	39	423
15	359	40	426
16	361	41	432
17	364	42	435
18	367	43	440
19	369	44	446
20	372	45	452
21	374	46	461
22	376	47	472
23	379	48	491
24	381	49	510