

OCTOBER 2011 ADMINISTRATION OF THE OHIO ACHIEVEMENT ASSESSMENT GRADE 3 READING TEST STATISTICAL SUMMARY

These statistics describe the entire population of Ohio grade 3 public school students (including community schools) tested during the October 2011 administration.

Table 1 – Summary of Parameters from the October 2011 Administration of the Ohio Achievement Assessment for Grade 3 Reading.

N-count	126,569
Max Raw Score	49
Raw Score Mean	31.16
Raw Score Standard Deviation	9.32
Raw Score SEM	3.27
Max Scaled Score	506
Scaled Score Mean	405.07
Scaled Score Standard Deviation	27.39
Scaled Score SEM	9.60
Reliability	0.88

Table 2 – Cut Score Points for Basic, Proficient, Accelerated, and Advanced Standards and Percentage of Students by Performance Levels

Standard	Raw Score	Scaled Score	Percent
Limited	Below 24	Below 385	22.4
Basic	24	385	18.6
Proficient	30	400	22.2
Accelerated	36	415	18.7
Advanced	41	432	18.2

Table 3 – Subscale (Raw Score) Bands for the Minimally Proficient Student in Grade 3, October 2011 Administration.

	Content Standard	Below	At or Near Proficient	Above	Possible Score
Grade 3 Reading	Vocabulary	0 – 7	8 – 9	10	10
	Reading process	0 – 9	10 – 12	13 – 18	18
	Informational text	0 – 4	5 – 7	8 – 11	11
	Literary text	0 – 6	7 – 9	10	10

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

Test Form Construction

All items in the October 2011 Ohio Grade 3 Reading Achievement Test were previously field-tested, and item difficulty estimates from the field test administrations were used to pre-equate operational test forms during form construction. Pre-equating based on field-test item parameters aids in the development of nearly equivalent Ohio Achievement Assessment (OAA) forms across test administrations.

Common Item Equating

Following administration of the October 2011 Grade 3 Reading Achievement Test, item difficulty values were re-estimated using an early return sample. The early return sample was selected to be statistically representative of all Ohio grade 3 public school students. Because bank item difficulty estimates for all test items were available from the previous field test administrations, all the operational items could potentially serve as anchor items in the equating process. When the equating process is complete, item difficulties from the current administration are directly comparable with those from the bank.

After calibrating items based on the early return sample data, all available linking items were used to calculate the linking constant needed to bring the set of operational items back to the appropriate OAA “bank,” or reference, scale values. First, the October 2011 difficulty values (from the early return sample) were computed and compared with the bank difficulty values. The mean difference between the current and the bank difficulties of the anchor items is called the equating constant. The equating constant was added to each difficulty value for the current test administration item parameters so that the mean item difficulties were equal, yielding linked parameter estimates for each item. When the equating process is complete, item difficulties from the current administration are directly comparable with those from the bank. This calibrates the test to the same scale as used in prior testing for the same content area and grade level.

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 2011 operational test, test items were calibrated and equated on the basis of the early return sample, and Rasch ability estimates were computed for each possible raw score. The Rasch ability estimates were then transformed to the Ohio Grade 3 Reading Achievement Test scale, which is scaled so that the Proficient standard is equal to 400.

Ohio Rounding Rule

When transforming raw scores to scaled scores, if the scaled score nearest to a performance standard cut score is below the cut score, then the scaled score is rounded up to equal the cut score. Otherwise, no special rounding is done. For example, if a raw score is associated with an observed scaled score of 398.94, and 398.94 is the closest observed scaled score to the Proficient level cut score, then this value is rounded up to 400, corresponding to the Proficient level performance 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.

Table 4 – Raw Score to Scaled Score Conversion Table
(cut scores indicated in **bold**)

Raw Score	Scaled Score	Raw Score	Scaled Score
0	251	25	388
1	270	26	390
2	290	27	393
3	302	28	395
4	311	29	398
5	319	30	400
6	325	31	402
7	331	32	405
8	336	33	407
9	340	34	410
10	344	35	413
11	348	36	415
12	352	37	418
13	355	38	421
14	359	39	424
15	362	40	428
16	365	41	432
17	367	42	435
18	370	43	440
19	373	44	445
20	376	45	451
21	378	46	459
22	381	47	470
23	383	48	488
24	386	49	506