MARCH 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 March 2004 administration.

N-Count	116,897
Max. Raw Score	49
Max. Scaled Score	503
Raw Score Mean	35.58
Raw Score Standard Deviation	8.76
Raw Score SEM	3.16
Scaled Score Mean	415.72
Scaled Score Standard Deviation	27.19
Scaled Score SEM	9.80
Reliability	.87

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

Standard	Raw Score	Scaled Score
Limited	Below 24	Below 385
Basic	24	385
Proficient	31	400
Accelerated	37	415
Advanced	42	432

Percentage of Students by Performance Levels

Standard	Percent
Limited	11.47
Basic	13.47
Proficient	19.47
Accelerated	26.02
Advanced	29.57

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

Test Form Construction

The March 2004 Grade 3 Reading Achievement Test is made up of items field-tested in March 2003, 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 form construction.

Common Item Equating

Following administration of the March 2004 Grade 3 Reading test, we re-estimated item difficulty values using an early return sample of 11,122 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 the March 2003 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 March 2004 operational test, test items are calibrated and equated on the basis of the early return sample, and Rasch ability estimates (theta) 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 385, the proficient standard to a scaled score of 400, the accelerated standard to a scaled score of 415, 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	260	25	386
1	278	26	388
2	297	27	391
3	308	28	393
4	317	29	395
5	324	30	397
6	329	**31**	**400**
7	334	32	402
8	339	33	404
9	343	34	406
10	347	35	409
11	350	36	412
12	354	**37**	**415**
13	357	38	417
14	360	39	420
15	363	40	424
16	365	41	428
17	368	**42**	**432**
18	370	43	436
19	373	44	442
20	375	45	448
21	377	46	456
22	380	47	467
23	382	48	485
24	**385**	49	503