Calibration, Equating, and Scaling

The Department of Education's psychometric team coordinates the work of a calibration team during dry-run and actual calibration, equating, and scaling (CES) activities each spring. The contractor is required to subcontract with at least one outside entity (currently HumRRO) to work on the calibration team during all CES activities. The outside entity reports results to the Department. The contractor's, the subcontractor's, and the Department's psychometric staff each independently perform CES to verify results.

Evaluation Criteria for Equating Solutions

1. Examination of anchor items with statistical shift different from other anchor items including the direction and degree of change in the item characteristic curves
2. Examination of flagged anchor items form by form, and year by year
3. Comparison of the resulting M1 and M2 coefficients among the solutions within year and across years
4. Comparison of TCC, TIF, and SEM curves by each equating solution (comparison of the curves before and after test administration)
5. Examination of scatterplots of old/new IRT A and B parameters and the magnitude of the correlations
6. Examination of the scale score frequency distributions on top of each other
7. Review of the changes in the percent proficienct rates by each equating solution
8. Content review of the anchors and content balance representation
9. Review average and individual item position shift between test administrations and previous item role.
10. Physical examination of anchor items, especially flagged items, in the test and answer books

2010 FCAT Grade 4 Reading
Calibration Result Summary

Recommended Solution across Parties

	HumRRO Solution A	**Pearson Solution A**	**FDOE**	**2009 Result**
N-count	196,909	196,909	196,909	195,826*
Mean Scale Score	323.013	323.03	323.013	324*
SD Scale Score	56.96	56.91	56.96	52*
% ACL 3 and above	71.67	72	71.67	74*
% ACL 1	15.46	15	15.46	13*
% ACL 2	12.88	13	12.88	13*
% ACL 3	32.21	32	32.21	35*
% ACL 4	28.68	29	28.68	30*
% ACL 5	10.78	11	10.78	9*
M1 – Scale	42.669	42.669	42.669	37.93
M2 – Location	330.821	330.821	330.821	331.28
N of AN Items	27	27	27	24
AN Removed	None	None	None	Seq57
Correlation: New vs. Old IRT A	0.906	0.906	0.906	N/A
Correlation: New vs. Old IRT B	0.937	0.937	0.937	N/A

*based on population.

2009 FCAT Grade 4 Reading
Calibration Result Summary

1. Recommended Solution across Parties

	HumRRO Solution B	CTB	FDOE	2008 Results
N-count	17,278	17,278	17,278	192,727*
Mean Scale Score	323.93	325	323.93	319*
SD Scale Score	51.26	N/A	51.26	56.28*
% ACL 3 and above	74	75	74	70*
% ACL 1	12.49	12	12.49	17*
% ACL 2	13.21	13	13.21	13*
% ACL 3	35.47	35	35.47	32*
% ACL 4	29.99	30	29.99	29*
% ACL 5	8.84	9	8.84	9*
M1 – Scale	37.930	37.951	37.930	41.783
M2 – Location	331.275	332.065	331.275	329.784
N of AN Items	24	25	24	23
AN Removed	Seq 57	None	Seq 57	Seq 56

*based on population.

2008 FCAT Grade 4 Reading
Calibration Result Summary

Recommended Solution across Parties

	HumRRO	**CTB**	**FDOE**	**2007 Results**
N-count	37,993	37,993	37,993	196,512*
Mean Scale Score	321.153	321	321.153	316*
SD Scale Score	55.411	N/A	55.411	57.68*
% ACL 3 and above	70.61	71	70.61	68*
% ACL 1	16.32	16	16.32	18*
% ACL 2	13.07	13	13.07	14*
% ACL 3	31.73	32	31.73	33*
% ACL 4	29.81	30	29.81	27*
% ACL 5	9.07	9	9.07	8*
M1 – Scale	41.783	41.347	41.783	42.716
M2 – Location	329.784	330.382	329.784	326.402
N of AN Items	23	24	23	24
AN Removed	seq 56	None	seq 56	Seq57

*based on population.

2010 FCAT Grade 5 Reading
Calibration Result Summary

Recommended Solution across Parties

	HumRRO Solution A	Pearson Solution A	FDOE	2009 Result
N-count	195,696	195,696	195,696	192,410*
Mean Scale Score	309.99	310	309.992	311*
SD Scale Score	59.99	59.98	59.993	58*
% ACL 3 and above	69.35	69	69.35	71*
% ACL 1	15.38	15	15.38	14*
% ACL 2	15.27	15	15.27	14*
% ACL 3	32.73	33	32.73	34*
% ACL 4	27.81	28	27.81	29*
% ACL 5	8.81	9	8.81	8*
M1 – Scale	46.557	46.557	46.557	45.42
M2 – Location	317.817	317.817	317.817	318.65
N of AN Items	29	29	29	24
AN Removed	None	None	None	None
Correlation: New vs. Old IRT A	0.900	0.900	0.872	N/A
Correlation: New vs. Old IRT B	0.970	0.970	0.961	N/A

*based on population.

2009 FCAT Grade 5 Reading
Calibration Result Summary

1. Recommended Solution across Parties

	HumRRO	CTB	FDOE	2008 Results
N-count	16,921	16,921	16,921	195,392*
Mean Scale Score	310.83	311	310.83	305*
SD Scale Score	57.98	N/A	57.98	57.31*
% ACL 3 and above	70	71	70	67*
% ACL 1	14.36	14	14.36	17*
% ACL 2	14.66	15	14.66	16*
% ACL 3	34.35	34	34.35	34*
% ACL 4	28.43	28	28.43	26*
% ACL 5	8.20	8	8.20	7*
M1 – Scale	45.424	45.424	45.424	44.437
M2 – Location	318.654	318.654	318.654	314.663
N of AN Items	24	24	24	31
AN Removed	None	None	None	None

*based on population.

2008 FCAT Grade 5 Reading
Calibration Result Summary

1. Recommended Solution across Parties

	HumRRO	CTB	FDOE	2007 Results
N-count	38771	38,771	38771	192,289*
Mean Scale Score	306.747	303	306.747	310*
SD Scale Score	56.677	N/A	56.677	56.22*
% ACL 3 and above	68.21	65	68.21	72*
% ACL 1	16.01	18	16.01	14*
% ACL 2	15.78	17	15.78	15*
% ACL 3	34.52	33	34.52	36*
% ACL 4	26.98	25	26.98	29*
% ACL 5	6.70	7	6.70	6*
M1 – Scale	44.437	46.451	44.437	42.871
M2 – Location	314.663	311.252	314.663	318.705
N of AN Items	31	29	31	26
AN Removed	None	56, 60	None	Seq46, 55

*based on population.