



## ***Progress Testing***

**FCAT *Test Maker* Item Bank**  
**RELIABILITY AND VALIDITY REPORT**

**2012**

**RELIABILITY**

Internal Reliability

**VALIDITY**

Content Validity

Concurrent Validity

# Progress Testing Reliability and Validity

## FCAT TESTMAKER ITEM BANK and CITRUS COUNTY SCHOOLS' BENCHMARK ASSESSMENT PROGRAM

Progress Testing's FCAT Test Maker Item Bank is a formative assessment bank of more than 30,000 of the highest quality Reading, Mathematics, Writing, Science, and EOC items for use in Florida's schools, grades 3-12. All items have been specifically written for Florida's Next Generation Sunshine State Standards/Grade Level Expectations, and built to the FCAT 2.0 Test Item Specifications.

John Mullen, a program specialist in Planning, Research and Evaluation, uses the FCAT Test Maker Item Bank, published by Progress Testing, for the Citrus County Benchmark Assessment Test (CBAT) program. The CBAT is administered 3 times each school year to all students enrolled in Citrus County schools, grades 3-10. CBAT results were then analyzed to yield reliability and validity data on the FCAT Test Maker Item Bank.

### RELIABILITY

Reliability tells whether a test is likely to yield the same results if administered to the same group of test-takers multiple times. The Citrus County data presented below demonstrates that the FCAT Test Maker Item Bank is highly reliable — you can have confidence in **all** items in our bank. If parallel tests are developed using items across the bank, the relative scores of students will show little change.

### Internal Reliability

Internal reliability is a measure of the internal consistency of the items in a test. Internal consistency estimates should be thought of as revealing the extent to which items on the test are internally consistent with one another — that is, the extent to which the items are homogeneous within a test. The Kuder-Richardson (K-R20) method was used to determine the internal reliability of the CBATs administered during the past four school years and made using the FCAT Test Maker Item Bank. The following tables illustrate the K-R20 reliability for FCAT Test Maker items, from the 2007-2008 to the 2010-2011 school year.

Math CBAT made with <i>FCAT Test Maker</i> items						
Grade Level	N			Reliability (K-R20)		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
3	985	980	1028	0.76	0.80	0.85
4	1030	1059	1089	0.81	0.84	0.88
5	1026	1057	1081	0.78	0.85	0.87
6	1080	1088	1092	0.84	0.83	0.86
7	1088	1090	1100	0.82	0.83	0.87
8	1095	1083	1101	0.80	0.83	0.86

Table 1: 2010-2011 Reliability of Math CBAT made with FCAT Test Maker Item Bank

4140 NW 27th Lane, Suite G, Gainesville, FL 32606 Phone: 800-930-TEST or 800-930-8378 Fax: 352-336-3782

For reliable results, use *FCAT Test Maker* items

# Progress Testing Reliability and Validity

Reading CBAT made with <i>FCAT Test Maker</i> items						
Grade Level	N			Reliability (K-R20)		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
<b>3</b>	984	995	1033	0.80	0.84	0.86
<b>4</b>	1034	1059	1089	0.84	0.83	0.82
<b>5</b>	1026	1045	1080	0.85	0.83	0.84
<b>6</b>	1079	1089	1092	0.84	0.86	0.89
<b>7</b>	1089	1092	1101	0.82	0.84	0.87
<b>8</b>	1097	1089	1105	0.84	0.85	0.88
<b>9</b>	937	852	N/A	0.81	0.85	N/A
<b>10</b>	1145	1009	N/A	0.85	0.87	N/A

Table 2: 2010-2011 Reliability of Reading CBAT made with FCAT Test Maker Item Bank

Science CBAT made with <i>FCAT Test Maker</i> items				
Grade Level	N		Reliability (K-R20)	
	Test 1	Test 2	Test 1	Test 2
<b>5</b>	1079	1082	0.78	0.79
<b>8</b>	1100	1115	0.80	0.81

Table 3: 2010-2011 Reliability of Science CBAT made with FCAT Test Maker Item Bank

4140 NW 27th Lane, Suite G, Gainesville, FL 32606 **Phone:** 800-930-TEST or 800-930-8378 **Fax:** 352-336-3782

For reliable results, use *FCAT Test Maker* items

# Progress Testing Reliability and Validity

Math CBAT made with <i>FCAT Test Maker</i> items						
Grade Level	N			Reliability (K-R20)		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
<b>3</b>	994	992	1018	0.79	0.80	0.84
<b>4</b>	1029	1058	1078	0.82	0.83	0.87
<b>5</b>	1024	1048	1078	0.79	0.82	0.84
<b>6</b>	1045	1078	1087	0.83	0.83	0.86
<b>7</b>	1088	1092	1104	0.82	0.82	0.87
<b>8</b>	1098	1087	1005	0.81	0.83	0.85

Table 4: 2009-2010 Reliability of Math CBAT made with FCAT Test Maker Item Bank

Reading CBAT made with <i>FCAT Test Maker</i> items						
Grade Level	N			Reliability (K-R20)		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
<b>3</b>	1000	998	1026	0.81	0.84	0.83
<b>4</b>	1023	1045	1087	0.80	0.83	0.84
<b>5</b>	1024	1035	1078	0.84	0.82	0.83
<b>6</b>	1089	1092	1092	0.86	0.86	0.89
<b>7</b>	1088	1092	1005	0.84	0.84	0.85
<b>8</b>	1048	1092	1015	0.84	0.85	0.84
<b>9</b>	971	902	N/A	0.82	0.85	N/A
<b>10</b>	1135	1045	N/A	0.86	0.87	N/A

Table 5: 2009-2010 Reliability of Reading CBAT made with FCAT Test Maker Item Bank

4140 NW 27th Lane, Suite G, Gainesville, FL 32606 **Phone:** 800-930-TEST or 800-930-8378 **Fax:** 352-336-3782

For reliable results, use *FCAT Test Maker* items

# Progress Testing Reliability and Validity

Science CBAT made with FCAT Test Maker items				
Grade Level	N		Reliability (K-R20)	
	Test 1	Test 2	Test 1	Test 2
5	1050	1062	0.74	0.78
8	1089	1092	0.79	0.81

Table 6: 2009-2010 Reliability of Science CBAT made with FCAT Test Maker Item Bank

Math CBAT made with FCAT Test Maker items						
Grade Level	N			Reliability (K-R20)		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
3	1140	1139	1133	0.82	0.83	0.86
4	1088	1069	1065	0.76	0.82	0.85
5	1142	1128	1169	0.85	0.84	0.85
6	1117	1153	1136	0.87	0.79	0.80
7	1212	1202	1135	0.84	0.81	0.80
8	1141	1143	1114	0.83	0.85	0.85

Table 7: 2008-2009 Reliability of Math CBAT made with FCAT Test Maker Item Bank

4140 NW 27th Lane, Suite G, Gainesville, FL 32606 Phone: 800-930-TEST or 800-930-8378 Fax: 352-336-3782

For reliable results, use *FCAT Test Maker* items

# Progress Testing Reliability and Validity

Reading CBAT made with <i>FCAT Test Maker</i> items						
Grade Level	N			Reliability (K-R20)		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
<b>3</b>	1140	1135	1133	0.85	0.84	0.83
<b>4</b>	1088	1068	1080	0.81	0.82	0.81
<b>5</b>	1111	1126	1140	0.82	0.79	0.84
<b>6</b>	1170	1185	1093	0.87	0.87	0.80
<b>7</b>	1217	1216	1144	0.79	0.85	0.86
<b>8</b>	1190	1197	1139	0.87	0.88	0.88
<b>9</b>	1147	1101	997	0.89	0.89	0.89
<b>10</b>	1198	1048	1086	0.87	0.91	0.87

Table 8: 2008-2009 Reliability of Reading CBAT made with FCAT Test Maker Item Bank

Science CBAT made with <i>FCAT Test Maker</i> items						
Grade Level	N			Reliability (K-R20)		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
<b>5</b>	1112	1122	1126	0.77	0.79	0.74
<b>8</b>	1176	1185	1066	0.76	0.83	0.70

Table 9: 2008-2009 Reliability of Science CBAT made with FCAT Test Maker Item Bank

4140 NW 27th Lane, Suite G, Gainesville, FL 32606 **Phone:** 800-930-TEST or 800-930-8378 **Fax:** 352-336-3782

For reliable results, use *FCAT Test Maker* items

# Progress Testing Reliability and Validity

Math CBAT made with <i>FCAT Test Maker</i> items						
Grade Level	N			Reliability (K-R20)		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
<b>3</b>	1042	1122	1076	0.81	0.82	0.85
<b>4</b>	1085	1119	1131	0.81	0.83	0.86
<b>5</b>	1054	1126	1129	0.79	0.83	0.82
<b>6</b>	1181	1252	669	0.87	0.84	0.78
<b>7</b>	1251	1252	797	0.78	0.84	0.83
<b>8</b>	1167	1208	676	0.88	0.86	0.87

Table 10: 2007-2008 Reliability of Math CBAT made with FCAT Test Maker Item Bank

Reading CBAT made with <i>FCAT Test Maker</i> items						
Grade Level	N			Reliability (K-R20)		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
<b>3</b>	1032	1126	1081	0.81	0.83	0.78
<b>4</b>	1074	1120	1109	0.79	0.79	0.81
<b>5</b>	1078	1140	1008	0.79	0.79	0.83
<b>6</b>	1192	1189	1138	0.86	0.84	0.80
<b>7</b>	1251	1266	1279	0.78	0.86	0.86
<b>8</b>	1217	1190	1164	0.87	0.87	0.88
<b>9</b>	1079	981	570	0.88	0.86	0.88
<b>10</b>	956	840	102	0.90	0.91	0.75

Table 11: 2007-2008 Reliability of Reading CBAT made with FCAT Test Maker Item Bank

4140 NW 27th Lane, Suite G, Gainesville, FL 32606 **Phone:** 800-930-TEST or 800-930-8378 **Fax:** 352-336-3782

For reliable results, use *FCAT Test Maker* items



# Progress Testing Reliability and Validity

## VALIDITY

Validity measures tell whether the test is measuring what it claims to measure. Citrus County used two forms of measurement to determine validity of the FCAT Benchmark Item Bank as used in the CBAT: Content Validity and Concurrent validity.

### Content Validity

Content-related evidence of Validity determines whether the test deals with the content it's supposed to be measuring.

### CCS Process for Determining Content Validity

Items were analyzed to determine the degree to which FCAT Test Maker's items were representative. The process for evaluating content validity included the following:

1. Content domain of the test defined in clear, unambiguous language
2. Items based on current Florida Sunshine State Standards
3. Items meet current FCAT Test Item Specifications
4. Identification of Webb's Cognitive Complexity for each item is accurate
5. Content Focus Reports and trend analysis of assessed items
6. Prioritization and Mapping of Benchmarks
7. Items mapped to Instructional Timeline (e.g. pacing chart)
8. Test sent through a series of post facto judgments about the representativeness of its content
9. Reviewed by a panel of experts (CBAT Teams)
10. Statistical review process conducted for test improvement (point-biserial and p-value)

### Concurrent Validity

Concurrent validity is demonstrated when a test correlates well with a measure that has previously been considered valid. Concurrent validity was used to conduct analysis of the correlation between the CBAT with FCAT Test Maker items and the FCAT. The correlations are displayed for Math, Reading, and Science in Tables 12 through 15.

Correlation of CBAT Reading Scores with FCAT Reading Scores			
Grade	Math	Reading	Science
3	0.81	0.84	
4	0.83	0.83	
5	0.85	0.82	0.80
6	0.76	0.80	
7	0.78	0.81	
8	0.73	0.80	0.80
9		0.73	
10		0.70	

Table 12: 2010 -2011 Correlation of CBAT to FCAT Scores



# Progress Testing Reliability and Validity

Correlation of CBAT Scores with FCAT Scores								
Grade Level	Reading			Math			Science	
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3	Test 1	Test 2
3	0.79	0.80		0.77	0.83			
4	0.76	0.82		0.78	0.82			
5	0.79	0.82		0.83	0.86		0.75	0.79
6	0.76	0.81		0.85	0.81			
7	0.75	0.79		0.76	0.81			
8	0.81	0.86		0.81	0.85			
9	0.73	0.75	0.80	0.71	0.73	0.77		
10	0.71	0.74	0.80	0.69	0.74	0.79		

Table 13: 2009-2010 Correlation of CBAT to FCAT Scores\*

Correlation of CBAT Scores with FCAT Scores			
Grade	Math	Reading	Science
3	0.8	0.8	
4	0.8	0.8	
5	0.9	0.8	0.8
6	0.8	0.8	
7	0.8	0.8	
8	0.8	0.8	0.8
9		0.8	
10		0.8	

Table 14: 2008-2009 Correlation of CBAT to FCAT Scores

Correlation of CBAT Scores with FCAT Scores		
Grade	Math	Reading
3	0.84	0.83
4	0.82	0.83
5	0.82	0.84
6	0.85	0.84
7	0.79	0.84
8	0.84	0.84
9		0.75
10		0.63

Table 15: 2007-2008 Correlation of CBAT to FCAT Scores

\*Due to budget constraints in 2009-2010, fewer tests were administered

4140 NW 27th Lane, Suite G, Gainesville, FL 32606 Phone: 800-930-TEST or 800-930-8378 Fax: 352-336-3782

For reliable results, use **FCAT Test Maker** items

# Progress Testing Reliability and Validity

## THE FCAT TEST MAKER ITEM BANK GIVES YOU RESULTS YOU CAN RELY ON

As seen in the Citrus County study, the FCAT Test Maker Item Bank has a very high degree of reliability and validity. Using the FCAT Test Maker Item Bank, educators can be certain:

- that their tests measure the Next Generation Sunshine State Standards,
- that any item they choose from the bank will yield similar results, and
- those results are highly predictive of FCAT scores.

Using the FCAT Test Maker Item Bank, assessment results can be reliably used to determine student growth and student performance relative to statewide grade-level achievement expectations. Assessments can be also be reliably used to guide classroom instruction and identify individual student needs for remediation, intervention and/or acceleration.



“Citrus County has been using **FCAT Test Maker** items for our districtwide assessments for 4 years now. The reliability and validity of their content are predictive of FCAT scores and aid us in identifying student strengths and weaknesses so that they may be built upon or addressed before the actual FCAT. Because of their commitment to customer service and the quality of their product, our district is happy to continue our working partnership with **FCAT Test Maker**, and look forward to a continued partnership for many years to come.”

- John Mullen

*Program Specialist - Planning/Research/Evaluation  
Research and Accountability Department  
Citrus County School Board*

“Three years ago, it was recommended that we look at **FCAT Test Maker** as a source for creating our FCAT simulation tests. We started using **FCAT Test Maker** way back then, but over the next two years, as we moved into data driven planning and instruction, we relied on **FCAT Test Maker** even more — for test items and reports. After creating tests and administering them, we could analyze the data and make decisions on reteaching, remediation, tutoring, and enrichment. We tested frequently with smaller assessments and used an entire FCAT simulation test generated by **FCAT Test Maker** each nine weeks. This frequent assessment and availability of data to drive instruction helped us move from an “F” to an “A” in one year! Now the entire district is using **FCAT Test Maker!**”

*- Felecia Moss, Principal, Metcalfe Elementary,  
Alachua County Schools*



4140 NW 27th Lane, Suite G, Gainesville, FL 32606 **Phone:** 800-930-TEST or 800-930-8378 **Fax:** 352-336-3782

For reliable results, use **FCAT Test Maker** items