



# TEST PREP SYSTEMS

**FCAT TestMaker Item Bank**

**RELIABILITY AND VALIDITY REPORT**

**2011**

**RELIABILITY**

**Internal Reliability**

**VALIDITY**

**Content Validity**

**Concurrent Validity**

# TEST PREP SYSTEMS RELIABILITY AND VALIDITY

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

Test Prep Systems' FCAT TestMaker Item Bank is a formative assessment bank of over 15,000 of the highest-quality Reading, Mathematics, Writing, and Science items for use in Florida's schools, grades 3-12. All items have been specifically written for Florida's Sunshine State Standards/Grade Level Expectations, and Next Generations Sunshine State Standards (when applicable) and built to the FCAT Test Item Specifications.

Dr. Gina Tovine, Director of Research and Accountability, selected the FCAT TestMaker Item Bank, published by Test Prep Systems, 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 TestMaker 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 TestMaker Item Bank is highly reliable-- you can have confidence in ALL the items in our bank. If parallel test 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 2007-2008 school year and made using the FCAT TestMaker Item Bank exclusively. The results are displayed in Tables 1 and 2.

MATH CBAT made with FCAT TestMaker 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 1: Reliability of Math CBAT made with FCAT TestMaker Item Bank

# TEST PREP SYSTEMS RELIABILITY AND VALIDITY

READING CBAT made with FCAT TestMaker 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 2: Reliability of Reading CBAT made with FCAT TestMaker Item Bank

## VALIDITY

Validity measures tell whether the test is measuring what it purports 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 Test Prep Systems' items were representative (representativeness).

The process for evaluating content validity included the following:

- 1) Content domain of the test defined in clear, unambiguous language
- 2) Items based on the Florida Sunshine State Standards – FCAT
- 3) Items meet the 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)

# TEST PREP SYSTEMS RELIABILITY AND VALIDITY

## Concurrent Validity

Concurrent validity is demonstrated where a test correlates well with a measure that has previously been validated. Concurrent validity was used to conduct analysis of the correlation between CBAT and FCAT. The correlations are displayed for Math and Reading in Tables 3 and 4.

Correlation of CBAT Reading Scores with FCAT Reading Scores				
Grade Level	N	Correlation	N	Correlation
	2007 Total Test		2008 Total Test	
<b>3</b>	903	0.80	1094	0.81
<b>4</b>	860	0.80	1103	0.82
<b>5</b>	904	0.82	1114	0.81
<b>6</b>	1035	0.83	1183	0.82
<b>7</b>	1088	0.79	1224	0.81
<b>8</b>	988	0.74	1184	0.80
<b>9</b>	267	0.70	971	0.74
<b>10</b>	NA	NA	805	0.66

Table 3: Correlation of CBAT to FCAT Reading Scores

Correlation of CBAT Math Scores with FCAT Math Scores				
Grade Level	N	Correlation	N	Correlation
	2007 Total Test		2008 Total Test	
<b>3</b>	880	0.84	1090	0.83
<b>4</b>	837	0.82	1102	0.83
<b>5</b>	855	0.82	1105	0.84
<b>6</b>	949	0.85	1185	0.84
<b>7</b>	902	0.79	1228	0.84
<b>8</b>	892	0.84	1154	0.84
<b>9</b>	523	0.80	984	0.75
<b>10</b>	NA	NA	641	0.63

Table 4: Correlation of CBAT to FCAT Math Scores

# TEST PREP SYSTEMS RELIABILITY AND VALIDITY

## THE FCAT TESTMAKER ITEM BANK GIVES YOU RESULTS YOU CAN RELY ON

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

- that their tests measure the 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 TestMaker 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 re-teaching, intervention and/or acceleration.

### **What's the most affordable option for your district?**

- One-time purchase
- Five-year license
- FCAT TestMaker Item Bank and software for use in all of your schools

**Call: 800-930-TEST**

"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 TestMaker!!!"

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

"*FCAT Test Maker* helped Gilchrist District Schools maintain an 'A' district grade and achieve the Florida Department of Education's distinction of being Florida's number one ranked school district for several years in a row. Before we purchased FCAT TestMaker, we had huge costs and headaches associated with test item creation. Now our teachers can focus on teaching while the team at Test Prep Systems focuses on writing high-quality test items for us."

**James E. "Buddy" Vickers, Superintendent of Schools, Gilchrist County**

## **The FCAT TestMaker Item Bank. Because what you use matters.**

