FCRR Technical Report #1

The Relationship Between Performance on a Measure of Oral Reading Fluency and Performance on the Florida Comprehensive Assessment Test

Dr. Julie Buck and Dr. Joseph Torgesen
Florida Center for Reading Research
Tallahassee, Florida

Purpose of the Study

There is an extensive body of research establishing measures of oral reading fluency as valid and reliable predictors of important reading outcomes and performance on high stakes tests (Good, Simmons, & Kame'enui, 2001). The present study was conducted to determine whether these findings also apply to the Florida Comprehensive Assessment Test as a measure of reading outcomes and achievement of Florida's reading standards for third graders. Specifically, we wanted to determine whether performance on brief, one-minute measures of oral reading fluency are predictive of achievement in reading as measured by the reading portion of the Florida Comprehensive Assessment Test-Sunshine State Standards (FCAT-SSS). If these measures reliably predict performance on the FCAT-SSS, then teachers might use them as "early warning signs" that students may, or may not, succeed on the FCAT-SSS.

Method

Thirteen schools from one Florida school district provided data, which included curriculum-based measures of oral reading fluency (ORF) and FCAT scores for third grade students (N=1102). ORF scores were obtained in May 2002, and the FCAT was administered in April 2002. The Standard Reading Passages: Measures for Screening and Progress Monitoring from Children's Educational Services, Inc. were used to determine ORF scores. Three one-minute ORF scores were obtained for each child. However, the ORF scores used in this study were each child's median score for those three passages.

Although, the focus of this study was on the reading comprehension portion of the FCAT-SSS, we also evaluated the relationship between ORF and performance on the math section of the FCAT-SSS and the reading comprehension section for the FCAT norm-referenced test (FCAT-NRT). Scale scores on the FCAT-SSS and the FCAT-NRT were used for these analyses. In the tables presented in this report, reading FCAT-SSS scores were also separated into categories based on achievement levels, which range from 1 to 5. Scores below level three indicate reading performance that is below grade level expectations.

Forty-nine percent of the students were female, and 83% were classified as white, 7% were classified as Black/African American, and 6% were classified as Hispanic. Only 1% of the students were considered limited English proficient (LEP), and 19% were

identified as exceptional students under IDEA. Almost half of the students received free or reduced lunch (46%), indicating low socioeconomic status.

Results

There was a significant correlation between ORF scores and reading FCAT-SSS scores (r=.70, p<.001), math FCAT-SSS scores (r=.53, p<.001), and reading scores on the FCAT-NRT (r=.74, p<.001). These correlations are very similar to the correlation that Good, Simmons, and Kame'enui (2001) found between ORF scores and the Oregon Statewide Assessment (r=.67, p<.001).

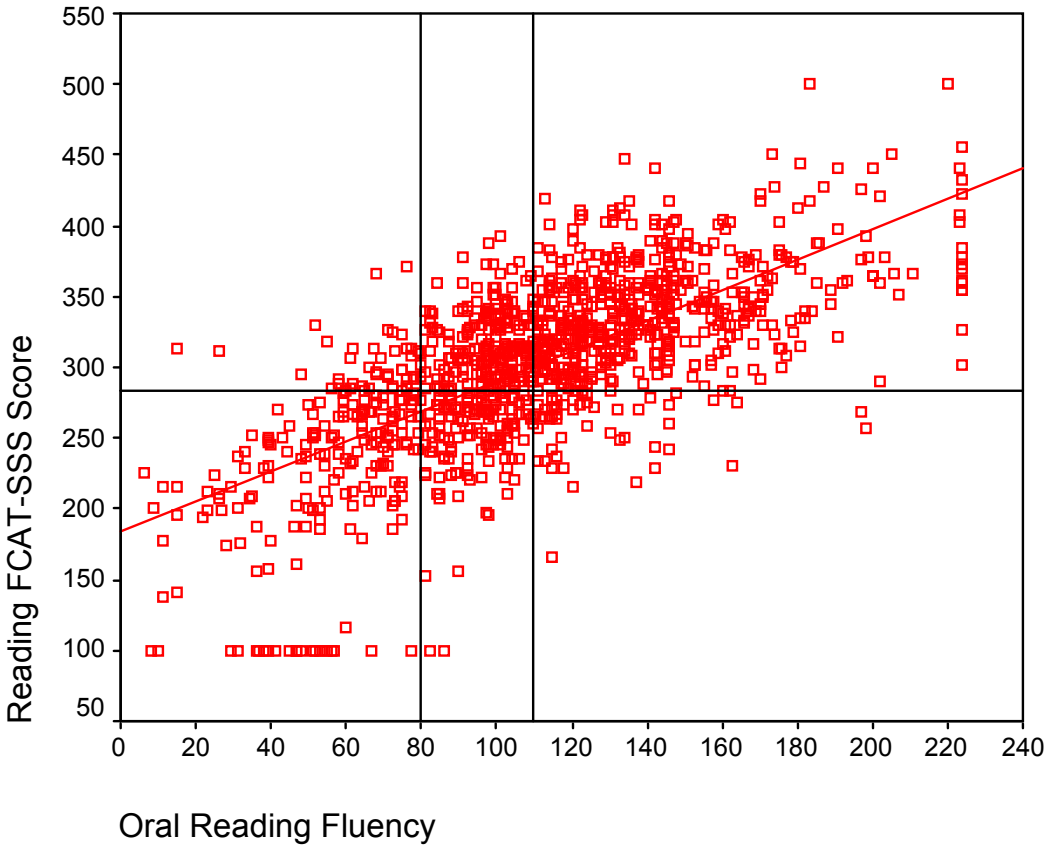

Figure 1: The relationship between ORF and reading FCAT-SSS scores using the ORF benchmarks described by Good, Simmons, Kame'enui, Kaminski, & Wallin (2002).

Figure 1 depicts the relationship between reading scores on the FCAT-SSS and oral reading fluency scores of students in spring of third grade. Students scoring at or above the horizontal line were at or above acceptable grade level in reading (FCAT-SSS achievement level 3 or above). The vertical lines represent ORF benchmarks chosen by Good, Simmons, Kame'enui, Kaminski, and Wallin (2002). As noted by Good et al. (2002), students with ORF scores at or above 110 are considered to be at low risk to show reading skills below grade level on a comprehensive measure of reading comprehension. Students scoring between 80 and 109 correct words per minute are considered to be at some risk. Students scoring below 80 correct words per minute are at high risk for

achieving a below grade level score on reading comprehension measures. Table 1 provides the breakdown of students by FCAT-SSS performance classification and ORF performance classification. Ninety-one percent of the students who read at or above the 110 cwpm benchmark achieved adequate performance on the reading section of the FCAT-SSS. Of the high risk students (students reading < 80 words per minute), 81% performed below level 3 on the reading section of the FCAT-SSS. These percentages are fairly similar to the percentages found by Good et al. (2001). Good and colleagues found that using these benchmarks, 96% of low risk students were correctly classified and 72% of high risk students were correctly classified. Usually, these types of analyses do not make predictions about students classified as at "some risk" because there is little accuracy of prediction for students with these mid level scores on oral reading fluency measures. In other words, a student who scores in this mid range on the ORF measure is considered equally likely to achieve satisfactory or unsatisfactory performance on the FCAT.

Table1

Predicting reading FCAT-SSS scores from ORF scores

	Oral Reading Fluency Classification			
FCAT Performance	High Risk (<80 cwpm)	Some Risk (80-109 cwpm)	Low Risk (≥110 cwpm)	Total
Adequate (FCAT Achievement Level 3-5)	42	188	511	741 67%
Inadequate (FCAT Achievement Level 1-2)	178	130	53	361 33%
Total	220 20%	318 29%	564 51%	1102

Sensitivity = .77
Specificity = .92

Note. Sensitivity and specificity calculations did not include students whose ORF scores were in the midrange (i.e., students at some risk).

In summary, Table 1 indicates that students performing in the midrange on the ORF measures were, in fact, almost equally likely to attain either an adequate or inadequate score on the FCAT-SSS. In contrast, those achieving scores at or above 110 correct

words per minute were very likely to attain scores at level 3 or above, and those scoring below 80 correct words per minute were very likely to obtain FCAT-SSS scores at level 1 or 2.

Data describing the classification of white, African American, and Hispanic students using ORF scores and reading FCAT-SSS scores are reported in Tables 2, 3, and 4. Although the correlations between ORF and FCAT-SSS performance are roughly similar across racial/ethnic groups ($\underline{r}$=.70 for white students; $\underline{r}$=.62 for African American students; $\underline{r}$=.78 for Hispanic students), scores above 110 correct words read per minute are slightly less predictive of success for minority groups, while scores below 80 cwpm are more predictive of failure for these groups. This relationship likely attains because children from minority groups are typically weaker than white students in other reading skills, such as vocabulary, which also predict reading comprehension. Thus, for minority children (particularly African Americans), achieving the benchmark of 110 words per minute was not as strong a guarantee of success on the FCAT as it was for the white students in this study. Conversely, if a minority student attains an oral reading score below 80, there is a greater likelihood of performance at level 1 or 2 on the FCAT than for the white students in this study. However, it is important to note that the data relating to race/ethnicity must be regarded as very preliminary because of the small number of students in minority classifications.

Table 2

Predicting reading FCAT-SSS scores from ORF scores for white students

FCAT Performance	Oral Reading Fluency Classification			
	High Risk (<80 cwpm)	Some Risk (80-109 cwpm)	Low Risk (≥110 cwpm)	Total
Adequate (FCAT Achievement Level 3-5)	37	163	437	637 70%
Inadequate (FCAT Achievement Level 1-2)	131	94	42	267 30%
Total	168 19%	257 28%	479 53%	904

Sensitivity = .76
Specificity = .92

Note. Sensitivity and specificity calculations did not include students whose ORF scores were in the midrange (i.e., students at some risk).

Table 3

Predicting reading FCAT-SSS scores from ORF scores for African American students

FCAT Performance	Oral Reading Fluency Classification			
	High Risk (<80 cwpm)	Some Risk (80-109 cwpm)	Low Risk (≥110 cwpm)	Total
Adequate (FCAT Achievement Level 3-5)	2	7	24	33 38%
Inadequate (FCAT Achievement Level 1-2)	23	25	5	53 62%
Total	25 29%	32 37%	29 34%	86

Sensitivity = .82
Specificity = .92

Note. Sensitivity and specificity calculations did not include students whose ORF scores were in the midrange (i.e., students at some risk).

Table 4

Predicting reading FCAT-SSS scores from ORF scores for Hispanic students

FCAT Performance	Oral Reading Fluency Classification			
	High Risk (<80 cwpm)	Some Risk (80-109 cwpm)	Low Risk (≥110 cwpm)	Total
Adequate (FCAT Achievement Level 3-5)	2	12	27	41 58%
Inadequate (FCAT Achievement Level 1-2)	20	6	4	30 42%
Total	22 31%	18 25%	31 44%	71

Sensitivity = .83
Specificity = .93

Note. Sensitivity and specificity calculations did not include students whose ORF scores were in the midrange (i.e., students at some risk).

Multiway frequency analyses were conducted to determine whether the interaction between racial/ethnic background (African American vs. white) and predictive accuracy for ORF scores documented in Tables 2 and 3 was statistically reliable. This analysis analyses included FCAT-SSS reading scores, ORF scores, and race/ethnicity to determine if ORF scores were equally predictive of adequate and inadequate performance on the FCAT for white versus African-American students. FCAT-SSS scores were divided into a dichotomous variable (i.e., level 1-2 and level 3-5). ORF scores were also divided into a dichotomous variable (i.e., high-risk students versus low-risk students). Students who were at some risk were not included in this analysis. A categorical variable was also created for ethnicity which included white and African-American students (N= 701). The interaction between these variables was not significant, $\chi^2(1) = 0.209$, $p = .65$, indicating that the predictive relationship of ORF on FCAT-SSS scores was not significantly different for African-Americans and white students. However, there was a significant interaction between FCAT-SSS and ORF, $\chi^2 (1) = 372.11$, $p < .0001$, supporting our prior conclusion that ORF scores significantly predict reading FCAT-SSS scores.

Similarly a multiway frequency analysis was conducted using ORF, FCAT-SSS scores, and ethnicity (i.e., Hispanic students and white students) to determine if the predictive

relationship between ORF and FCAT-scores were significantly different depending upon whether the student was Hispanic or white. A significant effect was not found, $\chi^2(1) = 0.45$, $p = .50$, suggesting that ORF scores predict FCAT-SSS scores equally well for white and Hispanic students. We anticipate being able to test these interactions in predictive accuracy between the ORF and FCAT more rigorously when we have data from a more diverse sample of students.

We also calculated the correlation between ORF and FCAT-SSS scores for students based on free or reduced lunch status (see Tables 5 and 6). The correlations between the FCAT-SSS and ORF scores are almost identical for students regardless of whether or not they are receiving free or reduced lunch (r=.69 for students who do not receive free or reduced lunch; r=.70 for students who receive free or reduced lunch). Students in the lower SES group, as determined by lunch status, are more likely to achieve level 1 or 2 performance if they have an oral reading score below 80 correct words per minute than are higher SES students. However, the relationship to satisfactory performance on the FCAT-SSS for students with ORF scores above 110 is identical for both groups.

Table 5

Predicting reading FCAT-SSS scores from ORF scores for students who do not receive free or reduced lunch

	Oral Reading Fluency Classification			
FCAT Performance	High Risk (<80 cwpm)	Some Risk (80-109 cwpm)	Low Risk (≥110 cwpm)	Total
Adequate (FCAT Achievement Level 3-5)	22	108	310	440 74%
Inadequate (FCAT Achievement Level 1-2)	74	49	32	155 26%
Total	96 16%	157 26%	342 58%	595

Sensitivity = .70
Specificity = .93

Note. Sensitivity and specificity calculations did not include students whose ORF scores were in the midrange (i.e., students at some risk).

Table 6

Predicting reading FCAT-SSS scores from ORF scores for students who receive free or reduced lunch

FCAT Performance	Oral Reading Fluency Classification			
	High Risk (<80 cwpm)	Some Risk (80-109 cwpm)	Low Risk (≥110 cwpm)	Total
Adequate (FCAT Achievement Level 3-5)	20	80	201	301 59%
Inadequate (FCAT Achievement Level 1-2)	104	81	21	206 41%
Total	124 24%	161 32%	222 44%	507

Sensitivity = .83
Specificity = .91

Note. Sensitivity and specificity calculations did not include students whose ORF scores were in the midrange (i.e., students at some risk).

A multiway frequency analysis was conducted with FCAT-SSS reading scores, ORF scores, and free or reduced lunch status to determine if ORF scores were equally predictive of adequate and inadequate performance on the FCAT for students who received free or reduced lunch and those who did not. As would predicted based on the previously discussed correlations, the interaction between these variables was not significant, χ^2 (1) = 0.877, p = .35, indicating that free/reduced lunch status did not significantly influence the relationship between ORF and FCAT-SSS scores.

Conclusions

This initial study demonstrates that, for a large heterogeneous group of third graders, performance on brief oral reading fluency measures can quite accurately predict whether or not a given students will attain a score at level 3 or above on the FCAT reading test. In the total sample, 91% of the students who attained oral reading fluency scores at or above 110 words per minute also achieved performance at level 3 or above on the FCAT. For the total sample, the lower cut point for oral reading fluency used in this study did not predict unsatisfactory performance on the FCAT as well as the upper cut point predicted

satisfactory performance. For example, only 81% of students with ORF scores below 80 actually achieved scores at level 1 or 2 on the FCAT.

When predictive relationships were studied separately for students from different ethnic groups, ethnicity did not significantly affect the relationship between ORF and FCAT-SSS scores, possibly due to the small minority population represented in this study. The predictive relationship between ORF and FCAT scores was quite similar for students who received free/reduced lunch and those who did not.

References

Good, R.H., Simmons, D.C., & Kame'enui, E. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*, 257-288.

Good, R. H., Simmons, D., Kame'enui, E., Kaminski, R. A., & Wallin, J. (2002). *Summary of decision rules for intensive, strategic, and benchmark instructional recommendations in kindergarten through third grade* (Technical Report No. 11). Eugene, OR: University of Oregon.

NOTE: Questions regarding this report should be directed to Dr. Julie Buck at the Florida Center for Reading Research. Her e-mail address is: jbuck@fcrr.org.