



Center for Research & Evaluation
Office of Accountability

DIBELS
Predictive Effectiveness

RESEARCH REPORT

October 2009

Reach Further.

Global competitiveness starts here.





**DIBELS Oral Reading Fluency
Predictive Effectiveness for 3rd-grade
Reading Performance in Charlotte-Mecklenburg Schools**

Oct. 12, 2009

Prepared by:

Samantha Kane Salvador, M.A.

Center for Research and Evaluation Sr. Analyst

Jason A. Schoeneberger, M.S.

Research and Evaluation Contractor

Lynne Tingle, Ph.D.

Director of the Center for Research and Evaluation

**Center for Research & Evaluation
Office of Accountability
Charlotte-Mecklenburg Schools**

For more information, contact:
**Center for Research & Evaluation
(980) 343-6242**

Dr. Lynne Tingle
*Director of the Center
for Research and Evaluation*

Robert Avossa
Chief Accountability Officer

Dr. Peter Gorman
Superintendent

TABLE OF CONTENTS

Executive Summary	1
Introduction	3
Method.....	5
Results.....	6
Relationship between DIBELS ORF Scores and 3rd Grade EOG Scores	6
Predictive Power of DIBELS ORF Scores vs. other Factors.....	7
Relationship between Growth on the DIBELS ORF Scale and 3rd Grade EOG Scores	9
Relationship between DIBELS Classification and EOG Proficiency.....	10
Stakeholder Opinions of DIBELS	13
Discussion	14
References.....	15
Appendices	16

EXECUTIVE SUMMARY

Research suggests that oral reading fluency, or the ability to read written passages aloud with minimal errors, is an indicator of future comprehensive reading mastery. Past studies have found that correlations between the DIBELS Oral Reading Fluency (ORF) scores and state reading tests range from .61-.80 but that this range increases when student scores are disaggregated by subpopulation (e.g., ethnicity, special accommodations, economic status). Past studies have also examined the relationship between DIBELS ORF categorical classifications (“At Risk”, “At Some Risk”, “Low Risk”) and proficiency on state reading tests. These studies found that between 77% and 87% of those students that failed state reading tests were classified as “At Risk” on the DIBELS whereas between 78% and 92% of those who passed state reading tests were classified as “Low Risk” on the DIBELS.

The purpose of this study is to verify the DIBELS predictive capability for CMS students as well as the accuracy of DIBELS student classification in predicting proficiency on the 3rd grade EOG Reading test. This report includes the results from the analysis of 2nd grade DIBELS ORF scores for 2007-2008 and 3rd grade EOG Reading scores for 2008-2009. However, six additional analyses were run, analyzing data from 2006-2007 to present. The results of these analyses are included in the footnotes and appendices. We examined the strength of the relationship (i.e., the correlation) between DIBELS ORF scores and 3rd grade EOG Reading scores, the predictive power of the DIBELS ORF scores after controlling for other predicting factors, the relationship between growth on the DIBELS ORF scale and 3rd grade EOG Reading scores, and the accuracy of the DIBELS categorical classification in predicting proficiency on the EOG. Stakeholder interviews were also utilized to ascertain what DIBELS users throughout the district believed were the positive and negative aspects of the DIBELS test.

DIBELS ORF scores and EOG Reading scores were moderately correlated ($r=.66$), meaning that lower DIBELS ORF scores were associated with lower EOG Reading scores and higher DIBELS ORF scores were associated with higher EOG Reading scores. This relationship is within the range ($r=.61-.80$) that has been found in other studies. However, ORF scores account for less than half of the variance in EOG scores. With the exception of American Indian/Alaskan Native students ($r=.42$), the relationship between DIBELS ORF scores and EOG Reading scores remained stable ($r \geq .60$) across ethnic groups, FRL status, LEP status, and EC status. However, it should be noted that compared to other ethnic groups represented in CMS data, there are very few American Indian/Alaskan Natives in this sample.

To further assess the strength of the relationship between DIBELS ORF scores and 3rd grade EOG Reading scores, we conducted a hierarchical multiple regression to assess the amount of variability in 3rd grade EOG Reading scores that was accounted for by DIBELS ORF scores after controlling for demographic characteristics. The regression showed that, of the variables included in the model, the DIBELS ORF score was the strongest predictor of 3rd grade EOG Reading, followed by African American ethnicity. However, the amount of variance accounted for was less than might be expected for an assessment of early reading skills and the predicted change on EOG scores was small relative to ORF score changes (i.e., a 6 point change in ORF scores is associated with a 1 point predicted change in EOG scores).

Growth on the DIBELS was not strongly associated with 3rd grade EOG Reading scores. On average, students gained 43.81 points on the DIBELS ORF scale from the beginning of the school year to the end. However, though significant, the correlation between DIBELS ORF growth and 3rd grade Reading EOG scores was low. A second hierarchical multiple regression showed that growth on the DIBELS ORF scale was *not* a strong predictor of 3rd grade EOG scores. However, when analyses were conducted separately for “At Risk” students versus “Low Risk” students, ORF growth scores were associated with a greater change in EOG performance for “At Risk” students.

Over 90% of students classified as “At Risk” based on their DIBELS ORF scores in 2007-2008 also failed the 3rd grade Reading EOG in 2008-2009. Sixty-eight percent of students classified as “At Some Risk” and 28% of those classified as “Low Risk” also failed the 3rd grade Reading EOG. Sensitivity refers to the number of students who failed the EOG that were correctly classified as “At Risk” or “At Some Risk” on the DIBELS. Specificity, on the other hand, pertains to the number of students who passed the EOG that were correctly classified as “Low Risk”. The sensitivity of the DIBELS in correctly classifying students who failed the EOG in 2008-2009 from their 2007-2008 DIBELS ORF scores was .47, meaning that of all the students who failed, only 47% of them were classified as “At Risk” or “At Some Risk” on the DIBELS ORF scale. Specificity was .91, meaning that 91% of the students who passed the EOG were correctly identified as “Low Risk”.

The proportion of false positives (i.e., those students identified as “At Risk” who then passed the EOG) were low (.09) meaning that of students who passed the Reading EOG, few had been classified as “At Risk/Some Risk” on the DIBELS. The proportion of false negatives (i.e., those students identified as “Low Risk” who then failed the EOG) were high (.53) meaning that a high proportion of students who were expected to pass based on their DIBELS ORF classification, actually failed the EOG. Compared to other studies that utilized sensitivity and specificity measures, sensitivity was *lower* for CMS students (.47 for CMS students vs .77-.87 found in other studies). Specificity was high for CMS students (.91 for CMS students vs .78-.92 found in other studies). In other words, a significant number of students who were classified as “Low Risk” by the DIBELS actually failed the 3rd grade Reading EOG.

The low sensitivity of the DIBELS in capturing students likely to fail the 3rd grade Reading EOG indicates a need to increase the cut score for categorizing students as “Low Risk”. Though doing so will increase the number of false positives (i.e., those students identified as “At Risk” who then passed the EOG), this is preferable to the currently high proportion of false negatives (i.e., those students identified as “Low Risk” who then failed the EOG).

On the other hand, stakeholders indicated that once students reach a very high level of fluency, they believe that repeated DIBELS ORF testing is no longer necessary. Stakeholders also indicated that concepts assessed by the DIBELS ORF scale are not the same as those assessed on the 3rd grade Reading EOG (Fluency vs. Comprehension). Correlation and regression analyses support this assertion; the amount of variance in EOG scores left unaccounted for indicates that, conceptually, there is a gap between the current reading fluency assessment and reading comprehension performance. Thus, rather than continued ORF testing, an assessment that gauges reading comprehension will likely fill this gap and better predict future Reading EOG performance.

INTRODUCTION

The DIBELS is a set of measures used to evaluate students' acquisition of literacy skills (*General Information about DIBELS*, 2008). The purpose of this assessment is to provide educators with a method for evaluating students' progress towards developmentally-appropriate reading goals. The tool consists of five measures of literacy skills: Initial Sounds Fluency (ISF), Phonemic Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), Oral Reading Fluency (ORF), and Word Use Fluency (WUF). DIBELS is not a measure of reading ability, but rather a measure of students' mastery of the skills required for reading comprehension. (For an extensive explanation of each measure, see dibels.uoregon.edu.) For the purpose of this study, we are using only the DIBELS ORF measure for 2nd-3rd grade students to predict their reading comprehension readiness for the third grade.

By the time students reach second and third grade, they should have mastered reading and understanding written passages. Theoretically, the DIBELS ORF assesses if students will be ready for this task. DIBELS ORF gauges a student's ability to read aloud three passages with a minimal amount of errors. Trained staff individually administers the test to a student. As a student reads a passage, the staff member records the number of times the student omits a word, replaces a word with another word, or takes longer than three seconds to proceed with reading a word. The staff member then categorizes the student's reading-skills development. A student can be categorized as "Low Risk", "At Some Risk", or "At Risk". These risk levels translate into odds in favor or against a student's readiness for the next level of reading mastery. "Low Risk" students have the requisite reading fluency to reach the next reading goal. DIBELS predicts that students at "Some Risk" have a 50% chance of being prepared. For "At Risk" students, however, the odds are against them in terms of their 3rd-grade reading comprehension readiness. Again, the measure does not assess a student's reading comprehension; however, research suggests that reading fluency is an indicator of future comprehensive reading mastery (*General Information about DIBELS Measure*, 2008). Given a student's categorization, educators can decide what type of academic support he or she needs in order to prepare him for the next stage of reading skills mastery. In the case of CMS 2nd-grade students, the next level of reading mastery is measured by the North Carolina Department of Public Instruction's (NCDPI) Reading Pretest-Grade 3. Ultimately, CMS' goal is for at least 80% of its 2nd-grade students to have the skills considered essential for 3rd-grade reading excellence (*Charlotte-Mecklenburg Schools Strategic Plan 2010*). Given the state's revisions to the reading test, CMS needs to examine if the DIBELS ORF measure serves as an effective diagnostic tool and if the measure's diagnostic effectiveness holds for targeted subpopulations within the district.

Research suggests that a strong positive relationship exists between students' oral reading fluency mastery and reading scores on standardized tests (Barger, 2003; Buck & Torgesen, 2003; Carlisle, Schilling, Scott, & Zeng, 2004; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008; Shaw & Shaw, 2002; Vander Meer, Lentz, & Stollar, 2005; Uribe-Zarain, 2006; Wilson, 2005). Studies to date have examined the relationship between DIBELS ORF and standardized reading tests in two primary ways. Researchers have examined the correlation between the DIBELS ORF scores and various state tests of reading (e.g., Delaware, North Carolina, Illinois, Ohio, Arizona, and Florida) for 3rd-grade students. The correlation between DIBELS ORF scores and state reading test scores is moderate to strong, ranging from .61-.80. Many of these studies, however, did not have ethnically or socio-economically

diverse samples. DIBELS was developed originally on a sample that does not represent the diversity of students found in CMS; however, the authors of DIBELS contend that the measures have been made generic for wider public use. One study did examine this correlation disaggregated by ethnicity, FRL, and special education status for 652 3rd grade students (Uribe-Zarain, 2006). Results showed that the correlations between DIBELS ORF and reading test scores were all significant, though the range of correlation coefficients was far larger than for studies in which groups were not disaggregated. The correlation was lowest (.35) for special education students and highest for Hispanic students (.72). CMS serves a diverse population and administers DIBELS in all of its schools; therefore, policymakers and educators need to verify that a DIBELS diagnostic capability applies to all of the students in the district.

The second method in which past studies have examined the relationship between DIBELS ORF scores and 3rd grade reading scores is to look at the DIBELS categorical classifications (i.e., “Low Risk”, “At Some Risk”, “At Risk”) of students and whether these students passed or failed state reading tests. Overall, 81-97% of 3rd-grade students who were classified as “At Risk” on the DIBELS ORF scale also failed state reading tests. Of the “Low Risk” students, 65-91% also passed state reading tests. Those students categorized as “At Some Risk” were equally likely to pass or fail state reading tests (Buck & Torgesen, 2003; CPS, 2005; Vander Meer et al., 2005; Wilson, 2005). Two useful statistical concepts for measuring the effectiveness of DIBELS in predicting whether students will pass or fail state reading tests are *sensitivity* and *specificity*. Sensitivity measures the proportion of actual positives that are correctly identified as such (i.e., those students who fail state reading tests and were classified as “At Risk” on the DIBELS). Specificity measures the proportion of correctly identified negatives (i.e., those students who pass state reading tests and were classified as low risk on the DIBELS). Several studies calculated the sensitivity and specificity of the DIBELS ORF measure in predicting performance on state reading tests (Buck & Torgesen, 2003; CPS, 2005; Uribe-Zarain, 2006). Sensitivity calculations ranged from .77-.87 and specificity calculations ranged from .78-.92. In other words, these studies found that between 77% and 87% of those students that failed state reading tests were classified as “At Risk” on the DIBELS whereas between 78% and 92% of those who passed state reading tests were classified as “Low Risk” on the DIBELS.

The purpose of this study is to verify the DIBELS predictive capability for CMS students. In order to do so, this evaluation assesses whether the DIBELS ORF score is a reliable predictor of CMS students’ reading performance on the North Carolina End of Grade (EOG) tests. The relationship between DIBELS and EOG scores will also be examined for students from different ethnic and socioeconomic backgrounds, as well as for students receiving special accommodations (e.g., exceptional children, students with limited English proficiency). The accuracy of student classifications (“At Risk”, “At Some Risk”, “Low Risk”) in predicting whether students pass or fail the Reading EOG will also be examined.

Research Questions

1. What is the relationship between DIBELS ORF scores and 3rd grade EOG Reading scores?
 - a. What is the relationship between these two measures when students are disaggregated by ethnicity, FRL status, EC status, and LEP status?
2. How well do DIBELS ORF scores predict 3rd grade EOG Reading scores in comparison to other predicting factors (such as ethnicity, FRL status, EC status, LEP status)?
3. What is the relationship between growth on the DIBELS ORF scale and 3rd grade EOG Reading scores?
4. Of those students classified as “At Risk”, “At Some Risk”, and “Low Risk”, at the end of 2nd grade, what percent were classified as proficient on the 3rd grade EOG Reading test?
 - a. What are the sensitivity and specificity of these classifications?
5. What do DIBELS users throughout the district believe are the positive and negative aspects of the test?

METHOD

Procedures

DIBELS ORF scores were obtained from Wireless Generation for the 2006-2007, 2007-2008, and 2008-2009 school years for students in grades 2 and 3. All second grade students were tested three times per year: Beginning of Year (BOY), Middle of Year (MOY), and End of Year (EOY). In third grade, only those students already categorized as “At Risk” were mandated to take the DIBELS assessments, though some schools continued to assess all third grade students. Students’ 3rd grade End of Grade (EOG) Reading test scores for 2007-2008 and 2008-2009 were obtained from the CMS data warehouse. Demographic data (i.e., ethnicity, FRL status, LEP status, and EC status) were also obtained from the CMS data warehouse.

This report includes the results from the analysis of 2nd grade DIBELS EOY ORF scores for 2007-2008 and 3rd grade EOG Reading scores for 2008-2009 (DIBELS 0708→EOG 0809). District leadership requested that additional comparisons be conducted and the results of these comparisons are included in the report footnotes. Six additional analyses were conducted. The first additional comparison (DIBELS 0607→EOG 0708) included those students with 2nd grade DIBELS EOY ORF scores for 2006-2007 and 3rd grade EOG Reading scores for 2007-2008. The second additional comparison (DIBELS 0708→PRE 0809) included students with 2nd grade DIBELS EOY ORF scores for 2007-2008 and 3rd grade EOG Reading Pretest scores for 2008-2009. The third additional comparison (DIBELS 0809→PRE 0809) included 3rd grade students with DIBELS BOY ORF scores and 3rd grade EOG Reading Pretest scores for 2008-2009. The fourth additional comparison (DIBELS EOY 0708→EOG 0708) included 3rd grade students with DIBELS EOY ORF scores and 3rd grade EOG Reading scores 2007-2008. The fifth additional comparison (DIBELS BOY 0708→EOG 0708) included 3rd grade

students with 2007-2008 DIBELS BOY ORF scores and 3rd grade EOG reading scores for 2007-2008. The final additional comparison (DIBELS EOY 0809→EOG 0809) included 3rd grade students with 2008-2009 DIBELS EOY ORF scores and 3rd grade EOG reading scores for 2008-2009. Significant findings from these additional comparisons are included in the report footnotes.

In order to assess the relationship between growth on the DIBELS scale and EOG reading scores, the difference between 2nd grade DIBELS BOY and EOY ORF scores for 2007-2008 were compared with 3rd grade EOG Reading scores for 2008-2009 (DIBELS GROWTH 0708→EOG0809).

Stakeholder interviews were utilized to ascertain what DIBELS users throughout the district believe are the positive and negative aspects of the DIBELS test, opinions regarding current cut scores, how comprehension should be assessed, and suggestions for improving the overall K-3 Intensive Reading program. Discussions and interviews were held with 24 stakeholders, including the Director of PreK-12 Literacy, the Specialist in Elementary Reading, three Area Superintendents, seven principals, two assistant principals, six literacy facilitators, two statisticians from outside agencies, one area support coordinator, and one math facilitator.

RESULTS

Relationship between DIBELS ORF Scores and 3rd Grade EOG Scores

DIBELS ORF and EOG Reading scores were moderately correlated ($r = .66$)¹ with DIBELS ORF scores accounting for 44% of the variance in 3rd grade EOG reading scores ($r^2 = .44$). When disaggregated by ethnicity, the relationship between DIBELS ORF scores and EOG Reading varied by ethnic group. Correlations ranged from .42 for Indian/Alaskan Native students to .65 for Asian students². The DIBELS ORF scale was similarly correlated with EOG Reading scores for FRL ($r = .62$) and non-FRL students ($r = .61$)³. ORF scores were slightly more strongly related to EOG Reading scores for non-LEP students ($r = .66$) than for LEP students ($r = .61$)⁴. When disaggregated by EC status, ORF scores were slightly more correlated with EOG scores for EC students ($r = .67$) than non-EC students ($r = .65$)⁵. See Table 1 for results by student classification status.

¹ Correlations for additional analyses ranged from .64 (DIBELS EOY 0809→EOG 0809) to .71 (DIBELS 0809→PRE 0809).

² Correlations for additional analyses ranged from .47 for Indian/Alaskan Native students (DIBELS 0809→PRE 0809) to .75 for Asian students (DIBELS 0809→PRE 0809).

³ Correlations for additional analyses ranged from .62 for FRL students (DIBELS EOY 0809→EOG 0809) to .70 for non-FRL students (DIBELS 0809 →PRE 0809).

⁴ Additional analyses showed that ORF scores were also more strongly related to EOG Reading scores for non-LEP students (correlations ranged from .64-.70) than for LEP students (correlations ranged from .58-.65).

⁵ For additional analyses, the relationship between DIBELS ORF scores and EOG Reading varied depending on the comparison. Correlations ranged from .57 to .70. When restricted to the comparisons with the largest sample sizes (DIBELS 0607→EOG 0708 and DIBELS 0708 →PRE 0809), the relationship between ORF scores and EOG Reading scores was slightly stronger for EC students (.70 for both comparisons) than non-EC students (.68 for both comparisons).

Table 1.
DIBELS and Reading EOG Descriptive Statistics

	DIBELS EOY 0708 --> EOG 0809*					
	Correlation	N	Mean EOG 0809	SD	Mean DIBELS 0708	SD
Ethnicity						
White	0.60	3300	347.54	10.00	124.83	37.66
Black	0.63	3746	335.59	10.80	100.35	36.88
Hispanic	0.64	1615	335.10	11.00	95.52	35.78
Asian	0.65	405	343.06	11.34	124.93	42.42
Multi Racial	0.65	452	341.83	10.56	111.02	36.97
Indian	0.42	44	340.27	10.19	105.14	35.64
FRL Status						
Free/Reduced	0.62	4710	334.86	10.81	96.52	36.55
Paid	0.61	4852	345.51	10.66	122.36	37.58
LEP Status						
LEP	0.61	1541	333.84	10.58	93.18	35.56
Not LEP	0.66	8021	341.49	11.84	112.60	39.15
EC Status						
EC	0.67	891	334.22	12.87	80.65	43.11
Not EC	0.65	8671	340.88	11.71	112.92	37.21
TOTAL	0.66	9562	340.26	11.98	109.33	39.25

* See Appendix A for descriptive data for all analyses run.

Predictive Power of DIBELS ORF Scores vs. other Factors

A hierarchical multiple regression was conducted to assess the amount of variability in 3rd grade EOG Reading scores that was accounted for by DIBELS ORF scores after controlling for demographic characteristics. Variables were entered into the regression in two steps; step 1 included race, FRL status, EC status, and LEP status and step 2 added DIBELS ORF scores. The results are reported in Table 2. The results of step 1 indicated that race, FRL status, EC status, and LEP status accounted for 31% of the variance in 3rd grade EOG Reading scores⁶. Adding DIBELS ORF scores in the 2nd step accounted for an additional 23% of the variance⁷. All variables together accounted for 54% of the variance in 3rd grade EOG Reading scores⁸. With the exception of Asian ethnicity, all variables individually predicted 3rd

⁶ Across the six additional analyses run, R^2 for step 1 ranged from .21-.31.

⁷ Across the six additional analyses run, ΔR^2 for step 2 ranged from .23-.30.

⁸ Across the six additional analyses run, R^2 for the entire model ranged from .49-.57.

grade EOG reading, with DIBELS ORF scores being the strongest predictor ($\beta=.529$). The unstandardized beta (B) shows the change in EOG reading scores for each unit change in the predictor variable. For example, being African American is associated with a loss of 8.703 points on the 3rd grade EOG Reading exam. Alternatively, every one point higher a student scores on the DIBELS ORF scale is associated with a .172 point higher score on the 3rd grade EOG Reading exam. In other words, every 6 points on the DIBELS ORF scale is associated with an increase of 1 point on EOG Reading. The standardized beta (β) standardizes variables so that their relative contribution to the model can be assessed. Doing so shows that of the variables entered into the regression, DIBELS ORF scores are the strongest predictor ($\beta=.529$) followed by African American ethnicity (.355)⁹. However, it should be noted that almost half of the variance in EOG Reading scores is not accounted for by these variables.

Table 2.
Hierarchical Regression Evaluating Predictors of 3rd Grade EOG Reading Scores.

Variables	Adj. R ²	ΔR^2	B	β	
<i>Step 1</i>	0.310				
FRL			-5.568	-0.232	***
African American			-8.703	-0.355	***
Hispanic			-3.811	-0.119	***
Asian			-0.673	-0.011	
American Indian/Alaskan Native			-5.359	-0.030	***
Multi			-4.096	-0.073	***
EC			-7.101	-0.172	***
LEP			-6.706	-0.206	***
<i>Step 2</i>	0.542	0.231			
DIBELS ORF EOY			0.172	0.529	***

Note. Betas reported are those from the step at which the variable was entered into the equation *** $p < .000$

⁹ Across the six additional analyses run, DIBELS ORF scores were the strongest predictor of EOG Reading scores, followed by African American ethnicity.

Relationship between Growth on the DIBELS ORF Scale and 3rd Grade EOG Scores

On average, second grade students gained 43.81 points on the DIBELS ORF Scale from the beginning of the school year to the end of the school year in 2007-2008. The range of ORF growth for students identified as “At Risk” at the beginning of their 2nd grade year was -4 to 149; for BOY “Low Risk” students, the range of ORF score growth was -46 to 151. Of the students who had negative growth, it should be noted that 94% of them passed the EOG.

Though statistically significant, growth on the DIBELS is not highly correlated with 3rd grade EOG Reading scores ($r=.038$). A second hierarchical multiple regression was conducted to assess the amount of variability in 3rd grade EOG Reading scores that was accounted for by DIBELS ORF score growth after controlling for demographic characteristics. Variables were entered into the regression in two steps; step 1 included race, FRL status, EC status, and LEP status and step 2 added DIBELS ORF score growth. The results are reported in Table 3. The results of step 1 indicated that race, FRL status, EC status, and LEP status accounted for 31% of the variance in 3rd grade EOG Reading scores. Adding DIBELS ORF score growth in the 2nd step added less than 1% of the variance accounted for. The unstandardized beta (B) showed that every one point of growth on the DIBELS ORF scale is associated with a .025 point increase on the 3rd grade EOG Reading exam. The standardized beta (β) showed that DIBELS ORF score growth are not the strongest predictor ($\beta=.040$) of 3rd grade EOG Reading scores. Because of the difference in ORF score growth ranges among “Low Risk” and “At Risk” students, two additional regression analyses were performed. The unstandardized beta (B) for BOY Low Risk students was .006; for At Risk students it was .18. This indicates that ORF score growth for 2nd grade BOY “At Risk” students results in a greater change in EOG performance when compared to those classified as “Low Risk.”

Table 3.
Hierarchical Regression Evaluating Predictors of 3rd Grade EOG Reading Scores.

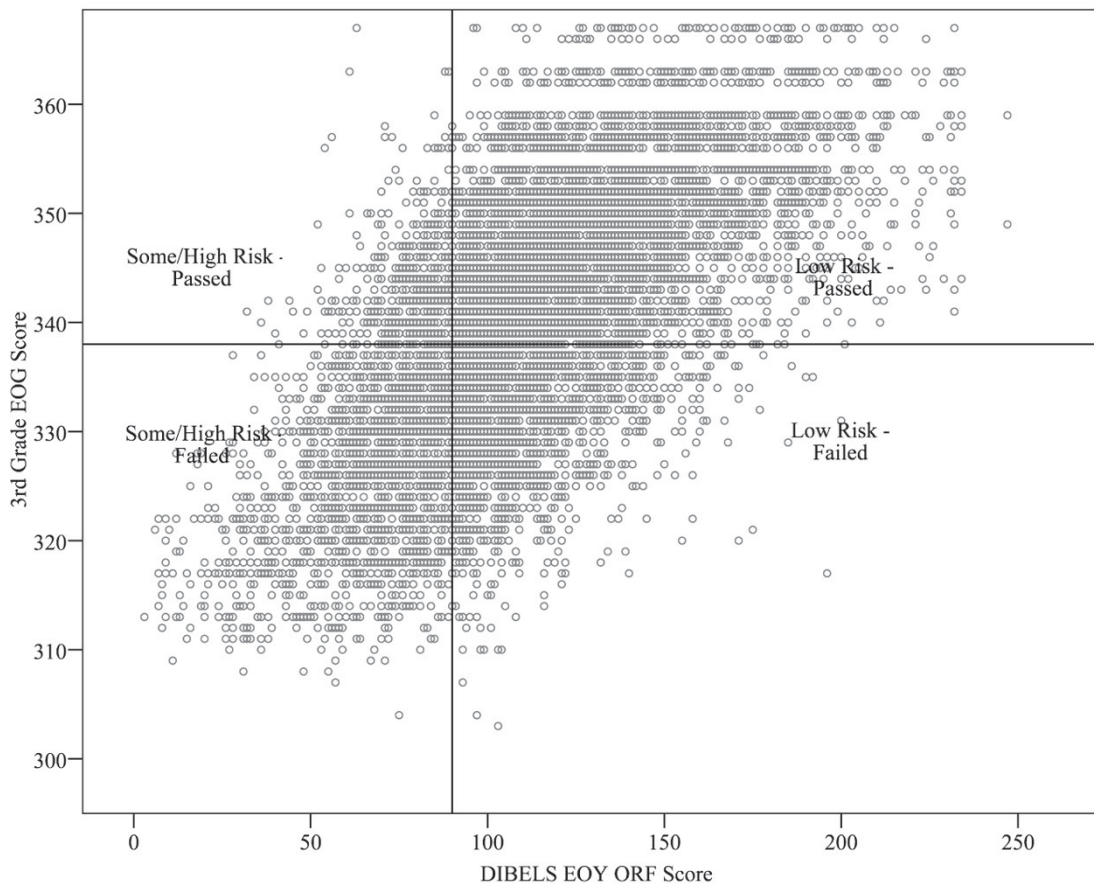
Variables	Adj. R ²	Δ R ²	B	β	
<i>Step 1</i>	0.310				
FRL			-5.568	-0.232	***
African American			-8.703	-0.355	***
Hispanic			-3.811	-0.119	***
Asian			-0.673	-0.011	
American Indian/Alaskan Native			-5.359	-0.030	***
Multi			-4.096	-0.073	***
EC			-7.101	-0.172	***
LEP			-6.706	-0.206	***
<i>Step 2</i>	0.316	0.002			
DIBELS ORF Growth			0.025	0.040	***

Note. Betas reported are those from the step at which the variable was entered into the equation
*** $p < .000$

Relationship between DIBELS Classification and EOG Proficiency

Over 90% of 2nd grade students classified as “At Risk” based on their DIBELS ORF scores in 2007-2008 also failed the 3rd grade Reading EOG in 2008-2009. Sixty-eight percent of second grade students classified as “At Some Risk” on the DIBELS ORF scale in 2007-2008 also failed the 3rd grade Reading EOG in 2008-2009¹⁰. However, being classified as “Low Risk” did not accurately predict performance on the EOG. Almost 30% of students classified as “Low Risk” on the DIBELS ORF scale in 2007-2008 also failed the 3rd grade Reading EOG in 2008-2009 (Table 4)¹¹. Figure 1 illustrates the distribution of EOG Reading scores based on DIBELS ORF scores.

Figure 1.
Distribution of EOG Reading Scores based on DIBELS ORF Scores.



¹¹ Over the six additional comparisons run, between 90.8% and 99.3% of students classified as “At Risk” also failed the 3rd grade Reading EOG.

¹² Over the six additional comparisons run, between 40.2% and 72.7% of students classified as “Low Risk” passed the EOG.

Sensitivity is a measure of the actual positives that are correctly identified by a diagnostic test. In this case, the sensitivity of the DIBELS relates to the number of students who failed the EOG that were classified as “At Risk” or “At Some Risk” on the DIBELS. Specificity, on the other hand, is a measure of the actual negatives that are correctly identified by a diagnostic test. In this case, specificity pertains to the number of students who passed the EOG that were classified as “Low Risk”. The sensitivity of the DIBELS in correctly classifying students who failed the EOG in 2008-2009 from their 2007-2008 DIBELS ORF scores was .47, meaning that of all students who failed, only 47% of them were classified as “At Risk” or “At Some Risk” on the DIBELS ORF scale. Specificity was .91, meaning that 91% of the students who passed the EOG were correctly identified as “Low Risk” (Table 4). The proportion of false positives (i.e., those students identified as “At Risk/At Some Risk” who then passed the EOG) were low (.09) while the proportion of false negatives (i.e., those students identified as “Low Risk” who then failed the EOG) were high (.53)¹².

Table 4.
Student Proficiency Rates on the EOG by DIBELS ORF Classification.

DIBELS EOY 0708 --> EOG 0809							
	% Not Proficient	N	% Proficient	N	Sensitivity	Specificity	Accuracy
At Risk	90.5%	904	9.5%	95	0.47*	0.91	.75 (105)†
At Some Risk	67.8%	884	32.2%	420			
At Risk/At Some Risk	77.6%	1788	22.4%	515			
Low Risk	27.7%	2012	72.3%	5247			

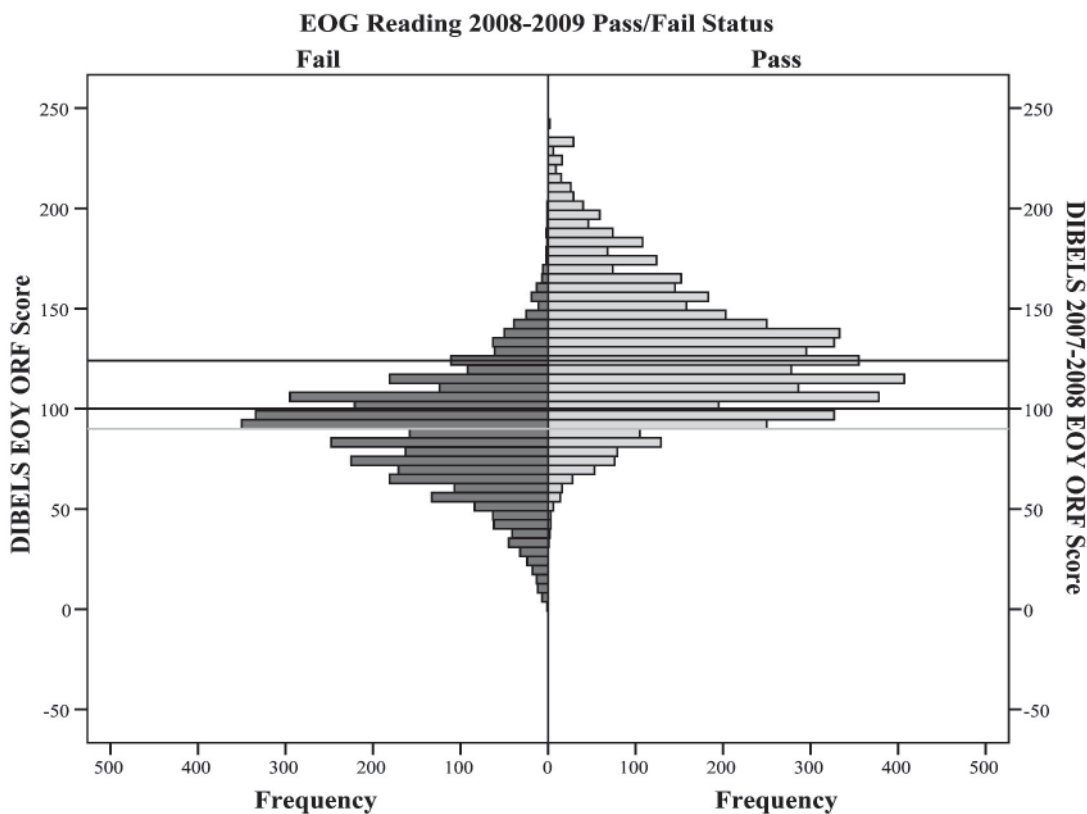
*The At Risk and At Some Risk students were grouped together to calculate sensitivity, specificity, and accuracy.

†Cut score at which highest accuracy is achieved

¹² Across the six additional analyses run, sensitivity scores ranged from .38-.69, meaning that between 38% and 69% of the students who failed the EOG were correctly identified as “At Risk”. Specificity scores ranged from .75-.97, meaning that between 75% and 97% of the students who passed the EOG were correctly identified as “Low Risk”. The proportion of false positives (i.e., those students identified as “At Risk” who then passed the EOG) ranged from .03-.25 while the number of false negatives (i.e., those students identified as “Low Risk” who then failed the EOG) ranged from .31-.62).

Figure 2 illustrates the number of students that passed and failed the EOG by DIBELS ORF scores. The bottom reference line is the current DIBELS ORF scale cut score (90) for 2nd graders, at which a student is considered “Low Risk”. The top reference line represents a higher cut score (124) which captures more of the students who are currently categorized as “Low Risk” on the DIBELS but then fail the EOG. The middle reference line (105) represents the cut score at which the most students will be accurately identified (74%)¹³.

Figure 2.
3rd Grade EOG Reading Pass/Fail Frequencies by 2nd Grade DIBELS ORF Scores



¹³ The highest possible accuracy rate for additional analyses ranged from 73-81%. In order to increase accuracy, ORF cut scores needed to increase between 7 and 50 points.

Stakeholder Opinions of DIBELS

Overall stakeholders like the DIBELS test. They report that it is useful in assessing beginning reading skills and identifying students who need extra help. For those students who are identified as “At Risk”, the DIBELS helps to build specific goals for students, is useful in red-flagging certain students and filtering them into intensive programs, and possesses a good monitoring system and indicator of achievement. Staff find that the M-Class system PDAs are not time consuming, that graphs and charts easy to interpret, and that the core reading assessments are very useful.

However, stakeholders report that the concepts assessed by the DIBELS test do not match those assessed by the EOG (i.e., fluency vs. comprehension) and another assessment that gauges comprehension is needed. Test corruption, in which teachers teach students specific skills (i.e., nonsense words) to improve DIBELS scores rather than the underlying concepts (i.e., phonics) that DIBELS assesses, is also a problem. Additionally, stakeholders reported operator errors in scoring and reduced consistency when multiple people administer the same test. Some suggested that once students reach a very high level of fluency they should no longer be required to take the DIBELS assessment. Furthermore, it was suggested that once students start comprehending, their speed may slow down in order to process the information. This was verified with the finding of 94% of the students who slowed down from BOY to EOY, actually passed the EOG. Some acknowledged that benchmark scores are currently set too low, however many strive for higher numbers than the formal cut score.

DISCUSSION

As illustrated by the correlational data and regression analyses, there is a moderate relationship between DIBELS ORF scores and 3rd grade EOG Reading scores. This relationship is within the range (.61-.80) that has been found in other studies that have assessed the relationship between DIBELS ORF scores and state standardized reading assessments. Though small differences were found when students were disaggregated by ethnicity, FRL, LEP, and EC status, the correlation remained stable ($r \geq .60$) across subgroups, with the exception of American Indian/Alaskan Natives. However, it should be noted that compared to other ethnic groups represented in CMS data, there are very few American Indian/Alaskan Natives in this sample.

As the regression analyses show, of the variables included in the model (ethnicity, FRL status, EC status, LEP status, DIBELS EOY ORF score, and DIBELS ORF growth), a student's DIBELS EOY ORF score was the strongest predictor of his/her 3rd grade EOG Reading scores. However, the amount of variance accounted for was less than might be expected for an assessment of early reading skills and the predicted change on EOG scores was small relative to ORF score changes (i.e., a 6 point change in ORF scores is associated with a 1 point predicted change in EOG scores).

Analyses showed that the relationship between the categorical classification (“At Risk”, “At Some Risk”, and “Low Risk”) and proficiency on the 3rd grade Reading EOG is weak for those students who fail the Reading EOG. Of the students who passed the 3rd grade Reading EOG, it was rare to find a student classified as “At Risk/At Some Risk”. However, of all students who failed the 3rd grade Reading EOG, only 47% of them were classified as “At Risk” or “At Some Risk” on the DIBELS ORF scale. **The low sensitivity of the DIBELS in capturing students likely to fail the 3rd grade Reading EOG indicates a need to increase the cut score for categorizing students as “Low Risk”.** Though doing so will increase the number of false positives (i.e., those students identified as “At Risk” who then passed the EOG), this is preferable to the currently high proportion of false negatives (i.e., those students identified as “Low Risk” who then failed the EOG).

On the other hand, stakeholders indicated that once students reach a very high level of fluency, they believe that repeated DIBELS ORF testing is no longer necessary. Stakeholders also indicated that concepts assessed by the DIBELS ORF scale are not the same as those assessed on the 3rd grade Reading EOG (Fluency vs. Comprehension). Correlation and regression analyses support this assertion; the amount of variance in EOG scores left unaccounted for indicates that, conceptually, there is a gap between reading fluency assessment and reading comprehension performance. Thus, rather than continued ORF testing, an assessment that gauges reading comprehension will likely fill this gap and better predict future Reading EOG performance.

REFERENCES

- Barger, J. (2003). *Comparing the DIBELS oral reading fluency indicator and the North Carolina end of grade reading assessment* (Technical Report). Asheville, NC; North Carolina Teacher Academy.
- Buck, J., & Torgesen, J. (2003). *The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test (FCRR Technical Report No. 1)*. Tallahassee, FL: Florida Center for Reading Research.
- Carlisle, J. F., Schilling, S. G., Scott, S. E., & Zeng, J. (2004). *Do fluency measures predict reading achievement? Results from the 2002-2003 school year in Michigan's Reading First schools* (Technical Report No. 1). Ann Arbor, MI: University of Michigan, Evaluation of Reading First in Michigan.
- Charlotte-Mecklenburg Schools Strategic Plan 2010, *Educating students to compete locally, nationally, and internationally*. Charlotte, NC.
- Good, R. H., III., Simmons, D. C., & Kame'enui, E. J. (2001). The importance of decision-making utility of a continuum of fluency-based indicators of foundational reading skills of 3rd-grade high-stakes outcomes. *Scientific Studies of Reading*, 3, 257-288.
- North Carolina Testing Program. (2008, May). *Pretest-Grade 3 multiple-choice test for Fall 2008* (Vol, 12 No.1). Raleigh, NC: NCDPI Division of Accountability Services.
- North Carolina Testing Program. (2009, January). *Understanding the individual student report for the North Carolina Pretest-Grade 3* (Stock No. 12175). Raleigh, NC: NCDPI Division of Accountability Services.
- Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology*, 4, 343-366.
- Shaw, R., & Shaw, D. (2002). *DIBELS oral reading fluency-based indicators of third grade reading skills for Colorado State Assessment (CSAP; Technical Report)*. Eugene, OR; University of Oregon.
- University of Oregon Center on Teaching and Learning. (2008, February). *General information about DIBELS measures* [Electronic version]. Retrieved February 3, 2009, from <https://dibels.uoregon.edu/dibelsinfo.php>
- Uribe-Zarain, X. (2006). *Relationship between student performance on DIBELS oral reading fluency and third grade reading DSTP* (Publication T06.013.01). Newark, DE; University of Delaware, Delaware Research & Development Center.
- Vander Meer, C. D., Lentz, F. E., & Stollar, S. (2005). *The relationship between oral reading fluency and Ohio proficiency testing in reading* (Technical Report). Eugene, OR; University of Oregon.
- Wilson, J (2005). *The relationship of Dynamic Indicators of Basic Literacy Skills (DIBELS) oral reading fluency to performance on Arizona Instrument to Measure Standards (AIMS)*. Tempe, AZ: Assessment and Evaluation Department, Tempe School District No. 3.

APPENDIX A.

Descriptive Statistics and Correlations by Comparison Type.

	DIBELS 0607--> EOG 0708					DIBELS 0708--> PRE 0809						
	Correlation	N	Mean EOG 0708	SD	Mean DIBELS 0607	SD	Correlation	N	Mean EOG 0708	SD	Mean DIBELS 0607	SD
Ethnicity												
White	0.62	3272	347.04	10.42	126.40	37.69	0.66	3397	336.47	12.70	124.77	37.61
Black	0.64	3633	334.73	11.12	100.72	34.54	0.65	3939	324.10	12.33	100.38	36.87
Hispanic	0.69	1462	334.51	11.20	94.95	36.64	0.66	1712	322.92	11.66	95.49	35.76
Asian	0.66	397	343.91	11.84	126.96	38.39	0.69	426	332.75	13.73	124.94	42.22
Multi Racial	0.63	422	341.00	11.29	112.46	37.02	0.68	479	329.62	12.85	111.02	36.97
Indian	0.54	38	340.39	9.59	102.76	32.89	0.61	46	326.02	13.69	105.14	35.64
FRL Status												
Free/Reduced	0.64	4503	333.89	11.04	96.76	35.13	0.65	5010	322.95	11.96	96.52	36.55
Paid	0.64	4721	345.37	10.89	123.78	37.23	0.66	4989	334.54	13.04	122.27	37.53
LEP Status												
LEP	0.58	821	328.90	9.36	78.47	32.16	0.62	1639	321.45	11.09	93.18	35.52
Not LEP	0.67	8403	340.83	12.12	113.73	37.79	0.70	8360	330.16	13.81	112.60	39.12
EC Status												
EC	0.70	879	333.56	14.15	91.02	42.03	0.70	1084	319.95	13.63	80.67	43.10
Not EC	0.68	8345	340.42	11.99	112.65	37.70	0.68	8915	329.80	13.42	112.91	37.18
TOTAL	0.69	9224	339.76	12.38	110.59	38.65	0.70	9999	328.74	13.78	109.33	39.21

APPENDIX A.

Descriptive Statistics and Correlations by Comparison Type.

	DIBELS 0809 --> PRE 0809				DIBELS EOY 0708 --> EOG 0708							
	Correlation	N	Mean EOG 0708	SD	Mean DIBELS 0607	SD	Correlation	N	Mean EOG 0708	SD	Mean DIBELS 0607	SD
Ethnicity												
White	0.71	833	333.69	13.79	104.51	38.30	0.66	826	339.92	11.61	112.65	36.38
Black	0.67	2341	323.05	11.96	82.73	32.13	0.63	2437	331.16	10.66	101.96	34.84
Hispanic	0.68	970	321.70	11.57	76.76	31.52	0.67	993	331.42	10.60	100.01	36.05
Asian	0.75	192	328.51	13.72	99.96	44.95	0.62	155	336.70	11.26	114.69	39.09
Multi Racial	0.68	231	328.42	12.22	94.27	33.61	0.65	224	336.83	11.29	113.35	33.76
Indian	0.47	25	323.20	13.09	80.38	29.08	0.74	22	335.36	9.97	99.30	36.08
FRL Status												
Free/Reduced	0.67	3144	322.24	11.84	79.58	32.38	0.63	3241	330.96	10.53	100.19	35.33
Paid	0.70	1448	331.61	13.30	102.30	36.32	0.65	1416	338.47	11.60	114.34	35.11
LEP Status												
LEP	0.65	881	319.37	10.63	72.17	31.34	0.57	675	327.52	9.09	88.72	33.92
Not LEP	0.70	3711	326.58	13.21	90.20	35.26	0.64	3982	334.22	11.47	107.35	35.43
EC Status												
EC	0.62	532	315.50	10.55	59.73	35.39	0.57	696	326.53	11.07	76.28	37.55
Not EC	0.70	4060	326.46	12.83	90.19	33.69	0.63	3961	334.43	11.04	110.45	32.40
TOTAL	0.71	4592	325.19	13.07	86.67	35.26	0.65	4657	333.25	11.40	104.37	35.85

APPENDIX A.

Descriptive Statistics and Correlations by Comparison Type.

	DIBELS BOY 0708 --> EOG 0708				DIBELS EOY 0809 --> EOG 0809							
	Correlation	N	Mean EOG 0708	SD	Mean DIBELS 0607	SD	Correlation	N	Mean EOG 0809	SD	Mean DIBELS 0809	SD
Ethnicity												
White	0.64	531	345.44	10.45	106.52	36.27	0.67	1110	342.24	10.91	122.65	36.47
Black	0.66	1161	334.03	11.02	83.61	33.76	0.62	2928	333.29	10.49	109.51	33.83
Hispanic	0.69	477	332.81	10.93	74.33	35.89	0.64	1189	333.12	10.79	109.34	33.39
Asian	0.64	120	340.53	11.27	106.17	40.48	0.63	221	338.05	12.05	125.03	40.52
Multi Racial	0.60	128	341.02	10.58	97.24	32.83	0.55	287	337.95	10.06	115.90	34.26
Indian	0.61	10	334.10	10.14	77.70	22.43	0.68	26	337.35	8.09	109.50	30.65
FRL Status												
Free/Reduced	0.65	1504	333.20	10.86	78.82	34.06	0.64	1896	340.28	11.13	121.01	36.02
Paid	0.65	923	343.12	11.04	104.37	36.06	0.62	3893	333.00	10.55	108.68	34.13
LEP Status												
LEP	0.61	331	328.73	9.48	62.55	31.14	0.59	1085	331.11	10.19	105.22	33.75
Not LEP	0.68	2096	338.28	11.77	92.99	36.03	0.65	4660	336.42	11.24	114.87	35.00
EC Status												
EC	0.68	269	328.56	11.78	58.04	36.92	0.61	626	328.63	10.87	84.61	38.70
Not EC	0.67	2158	338.03	11.54	92.93	34.81	0.63	5119	336.25	11.01	117.35	32.20
TOTAL	0.70	2427	336.98	11.94	88.49	36.96	0.64	5789	335.38	11.27	112.67	35.23

APPENDIX B.

Student Proficiency Rates on the EOG by DIBELS ORF Classification.

	DIBELS 0607 --> EOG 0708				DIBELS 0708 --> PRE 0809									
	% Not Proficient	N	% Proficient	N	Sensitivity	Specificity	Accuracy	% Not Proficient	N	% Proficient	N	Sensitivity	Specificity	Accuracy
At Risk	90.8%	1112	9.2%	113				99.3%	1201	0.7%	9			
At Some Risk	68.4%	911	31.6%	420	0.53	0.9	.76 (97)	92.6%	1254	7.4%	100	0.38*	0.97	.78 (124) †
Risk	79.1%	2023	20.9%	533				95.7%	2455	4.3%	109			
Low Risk	27.3%	1820	72.7%	4847				53.3%	3963	46.7%	3472			

*The At Risk and At Some Risk students were grouped together to calculate sensitivity, specificity, and accuracy.

†Cut score at which highest accuracy is achieved

Student Proficiency Rates on the EOG by DIBELS ORF Classification, Cont.

	DIBELS 0809 --> PRE 0809				DIBELS EOY 0708 --> EOG 0708									
	% Not Proficient	N	% Proficient	N	Sensitivity	Specificity	Accuracy	% Not Proficient	N	% Proficient	N	Sensitivity	Specificity	Accuracy
At Risk	99.7%	687	0.3%	2				94.0%	802	5.2%	44			
At Some Risk	95.7%	1110	4.3%	50	0.52	0.96	.81 (127)	77.6%	1277	22.4%	369	0.69*	0.75	.75 (126) †
Risk	97.2%	1797	2.8%	52				83.4%	2079	16.6%	413			
Low Risk	59.8%	1639	40.2%	1104				42.4%	919	57.6%	1246			

*The At Risk and At Some Risk students were grouped together to calculate sensitivity, specificity, and accuracy.

†Cut score at which highest accuracy is achieved

APPENDIX B.

Student Proficiency Rates on the EOG by DIBELS ORF Classification, Cont.

	DIBELS BOY 0708 --> EOG 0708					DIBELS EOY 0809 --> EOG 0809								
	% Not Proficient	N	% Proficient	N	Sensitivity	Specificity	Accuracy	% Not Proficient	N	% Proficient	N	Sensitivity	Specificity	Accuracy
At Risk	95.7%	331	4.3%	15	0.61	0.88	.76 (85)	94.5%	657	5.5%	38	0.61*	0.81	.73 (122) †
At Some Risk	75.9%	415	24.1%	132				75.5%	1355	24.5%	439			
Risk	83.5%	746	16.5%	147				80.8%	2012	19.2%	477			
Low Risk	31.4%	481	68.6%	1053				39.3%	1279	60.7%	1977			

*The At Risk and At Some Risk students were grouped together to calculate sensitivity, specificity, and accuracy.

†Cut score at which highest accuracy is achieved

An Evaluation Report Prepared by the
CENTER FOR RESEARCH & EVALUATION
OFFICE OF ACCOUNTABILITY



October 2009

In compliance with federal law, Charlotte-Mecklenburg Schools administers all education programs, employment activities and admissions without discrimination against any person on the basis of gender, race, color, religion, national origin, age or disability.