

Do intermediate students—grades 3, 4, and 5—who take the 4Sight assessment show improved performance on the PSSA?

(Intersections: School Processes with Student Learning with Demographics)

Mark Schlosser

Oswayo Valley Elementary School
Oswayo Valley School District

Mrs. Carolyn Fugate, Principal/On-site Supervisor

Abstract

The focus project studied the following question, “Do intermediate students—grades 3, 4 and 5—who take the 4Sight assessment show improved performance on the PSSA?” Multiple measures intersected were Student Learning by School Processes by Demographics. Student assessment data from the 4Sight assessments and PSSA (eMetric) were analyzed. The preliminary results seem to indicate that although the students tend to score lower on the 4Sight assessments than the PSSA, the regular progress monitoring serves as instructional benchmarks to increase student learning.

Problem

The enactment of the “No Child Left Behind” ushered in an era of heretofore unprecedented degree of accountability. The typical yearly assessment utilized by many districts no longer proved adequate measures of student achievement. The legislation brought a parallel emphasis on progress monitoring. Schools turned to benchmark assessment systems to monitor student progress, identifying specific gaps in content and/or learning which needed emphasis. The 4Sight Assessment, marketed through the “Success for All Foundation”, provides districts a mechanism to assess, adjust, and—hopefully—achieve.

The Oswayo Valley School Elementary School administers the 4Sight Assessment to grades 3, 4, and 5 a minimum of 4 times throughout the school year. In the years since introducing the 4Sights, no study determined the validity of the assessments as predictor of student performance on the Pennsylvania System of School Assessment (PSSA) end of year evaluations. It has been assumed the time, energy, and money invested in administering the 4Sights has produced results; this study purposed to verify that assumption.

Bernhardt (2009) describes multiple measures of data used “to understand the whole experience of school.” The study incorporated Student Learning as evidenced by PSSA results. Secondary levels of achievement lay in the results of the 4Sights; however, those results must be seen as intermediary guides rather than final goals. The administration and interpretation of the 4Sights represent School Processes. The analysis across grade levels (3-5) and cohort years (2008-2011) reflects Demographics.

Method

The initial information needed for the study laid in the 4Sight and PSSA Mathematics and Reading results. The specific goal of the question required examination of data from three different levels: what have the scores over time shown (results for all students over all years); how have cohorts progressed (specific classes over all years); and how have individual students tested (specific students over specific years). The web-based tools available through “Data Interaction for Pennsylvania Student Assessments”, eMetric, provided the necessary historical data for comparison of the PSSA results; the web-based tools available through “Success for All Foundation” provided the necessary historical data for comparison of the 4Sight results.

In order to establish greater validity, it was necessary to monitor results over time. A longitudinal study allows for a larger sampling number as well as marking trends. Individual “snapshots” of class or student performance, while helpful, may not fully establish patterns. The chart below illustrates the levels of sample selection.

	2008	2009	2010	2011
Class of 2015	Grade 5			
Class of 2016	Grade 4	Grade 5		
Class of 2017	Grade 3	Grade 4	Grade 5	
Class of 2018		Grade 3	Grade 4	Grade 5
Class of 2019			Grade 3	Grade 4
Class of 2020				Grade 3

Results

The first step for the project required entering raw data from the 4Sight assessments and the PSSA into a workable spreadsheet. Though time-consuming, the process created a comparative reference point for all students grades 3-5 who have taken both the 4Sight and PSSA since 2008. The grade level norms for each particular 4Sight imply the lowest scores should be the baseline (i.e. a third grader testing in September would be expected to perform lower than the same student testing in the following Spring of the same year); the results in both Mathematics (see Appendix A) and Reading (see Appendix B) bear this out—the baseline shows the lowest level of performance. However, as each successive test is given, collective performance rises as the percentage of “Advanced” and “Proficient” increases (by 48.5 and 23.8 percentage points respectively for Mathematics; by 18.9 and 12.4 percentage points respectively for Reading) and the percentage of “Basic” and “Below Basic” decreases (by 22.1 and 46.5 percentage points for Mathematics; by 15.9 and 15.3 percentage points respectively for Reading). In its most general sense, the 4Sight accurately gauges how students progress throughout the academic year.

On a more specific level, the results of the two assessment mechanisms should correlate. As a precursor to the PSSA, how accurately does the 4Sight predict student performance? In order to answer this question the data were organized to include only those records in which students completed the 4Sight 4th Session assessments (as mentioned, this test should be the most reliable predictor) and the PSSA. In the years studied, 292 records met this criterion. In Mathematics the following results occurred when comparing the 4Sight to the PSSA:

- ◆ Students who raised their performance level: 123 (42%);
- ◆ Students who maintained their performance level: 148 (51%); and
- ◆ Students who lowered their performance level: 21 (7%). (See Appendix C)

In Reading the following results occurred when comparing 4Sight to the PSSA:

- ◆ Students who raised their performance level: 75 (26%);
- ◆ Students who maintained their performance level: 171 (59%); and
- ◆ Students who lowered their performance level: 46 (16%). (See Appendix D)

The results show that as a predictor the 4Sight overwhelmingly indicates comparable or better performance on the actual PSSA (93% for Mathematics and 85% for Reading). A clear pattern exists. It is necessary to emphasize these numbers track not general school-wide results, but specific student results. In other words, the data reflect how each student who scored “Advanced” on the PSSA scored on the 4Sight, how each student who scored “Proficient” on the PSSA scored on the 4Sight, etc.

Similarly, a comparison can be made with the 3rd Session assessments. In the years studied, fewer 3rd Session assessments were given—just 193. In Mathematics the following results occurred when comparing the 4Sight to the PSSA:

- ◆ Students who raised their performance level: 101 (52%);
- ◆ Students who maintained their performance level: 89 (46%); and
- ◆ Students who lowered their performance level: 3 (2%). (See Appendix E)

In Reading the following results occurred when comparing 4Sight to the PSSA:

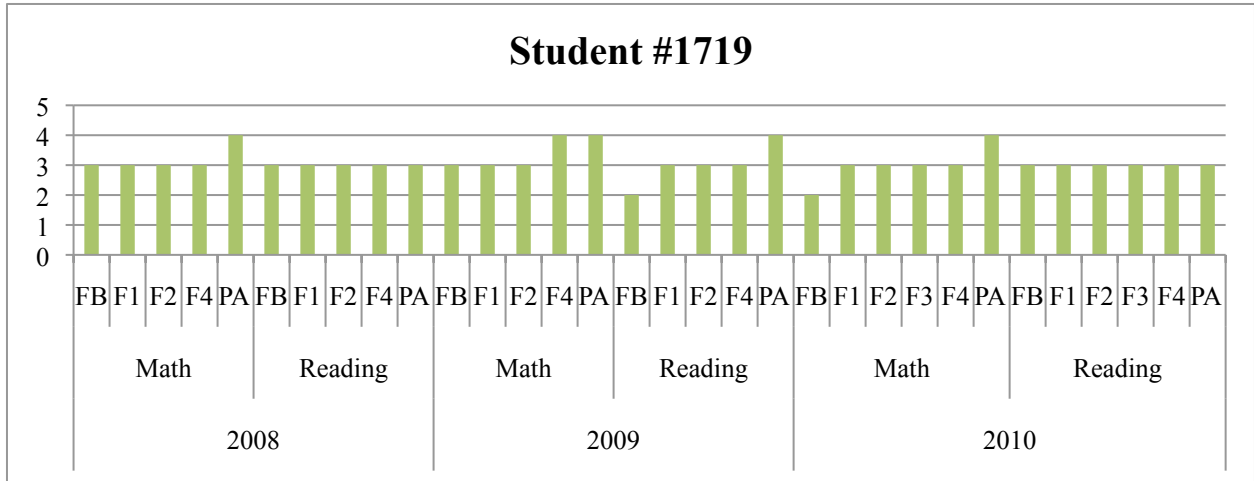
- ◆ Students who raised their performance level: 53 (27%);
- ◆ Students who maintained their performance level: 110 (57%); and
- ◆ Students who lowered their performance level: 30 (16%). (See Appendix F)

The results are strikingly similar to those of the 4th Session comparison. Students perform at comparable levels in both Mathematics (98%) and Reading (84%). Once again, the 4Sight serves as a valid predictor of student performance.

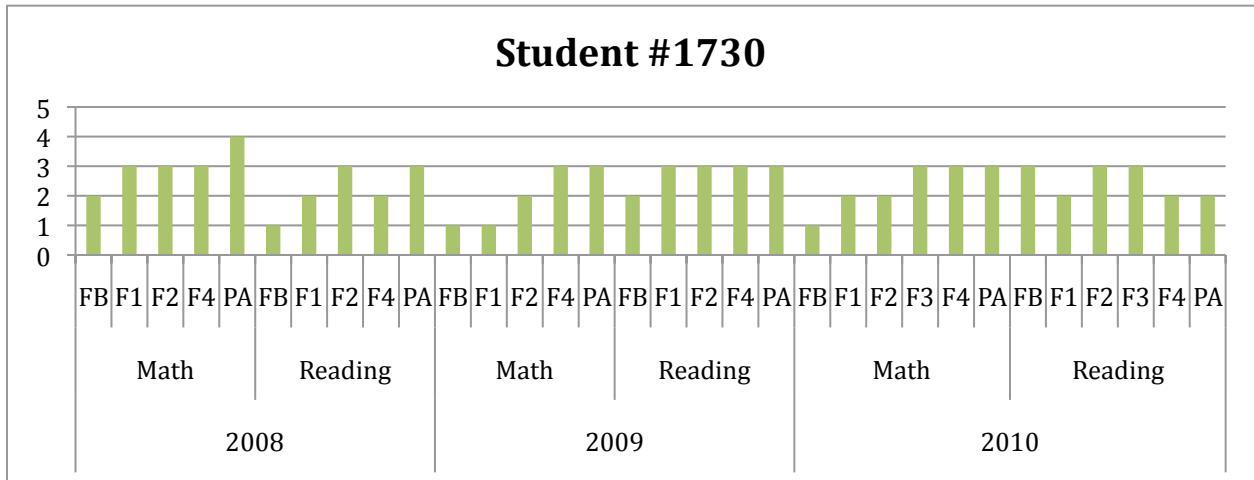
The study clearly indicates students progress throughout the year. The data, however, reveal the 4Sight’s inability to predict increased achievement from year to year. In the student sample selected, two cohorts—the classes of 2017 and 2018—completed assessments for the testing period. (See table on page 4) Of those students, forty-nine had PSSA scores for all three years studied. The results, as illustrated in the following table, show no significant increase in successive years.

PSSA Scores in Successive Years (Grades 3, 4, 5)				
	Increased Scores (e.g. P,P,A; BB,B,P)	Identical Scores	Decreased Scores (e.g. P,P,B; A,P,B)	Varied Scores (e.g. P,A,P; B,BB,B)
Mathematics	33%	76%	8%	14%
Reading	20%	37%	18%	22%

A closer examination of two sample students illustrates the 4Sights as legitimate predictors of progress within a testing year but not necessarily in successive years.



This student’s scores show that progress within the year may not represent sustained improvement. (Reading PSSA fluctuates from 2008 to 2010) . It also illustrates the tendency to perform higher on the PSSA than on the 4Sight. The second student similarly shows improvement within a single year marked by inconsistent year to year performance.



Discussion and Conclusion

As mentioned previously, the 4Sight serves as an accurate predictor of student performance on the PSSA. The results suggest that consistent progress on the 4Sight tests prepares the student for success. Particularly at the lower grade levels, the use of the 4Sight serves an introductory role to acclimate the student to standardized testing.

A variety of additional conclusions stand out from the study:

- ◆ The study creates a platform for including additional data. The independent reporting systems of the two assessments must be merged to provide sufficient comparison. The database used for this study can easily accommodate additional years’ data to extend the analysis into future years.
- ◆ The study justifies the time and money spent administering the tests and evaluating the results. Districts face budgetary restraints; testing costs money. The retention of the 4Sights plays a significant role in progress monitoring and curriculum alignment.
- ◆ The study suggests the full testing battery of baseline plus the four subsequent assessments may not be necessary. The close connection between the 3rd and 4th session results indicate that students may reach a plateau in their performance on the 4Sights. Though the testing prepares the students structurally for taking the PSSA, “testing burnout” perhaps can be minimized by scheduling only three assessments instead of the prescribed four.

- ◆ The study illustrates performance differences between the two content areas. There appears to be stronger results , both in progress and final results, for Mathematics than Reading. This study did not examine curriculum issues, but the reporting data from 4Sight could be utilized for such an investigation.

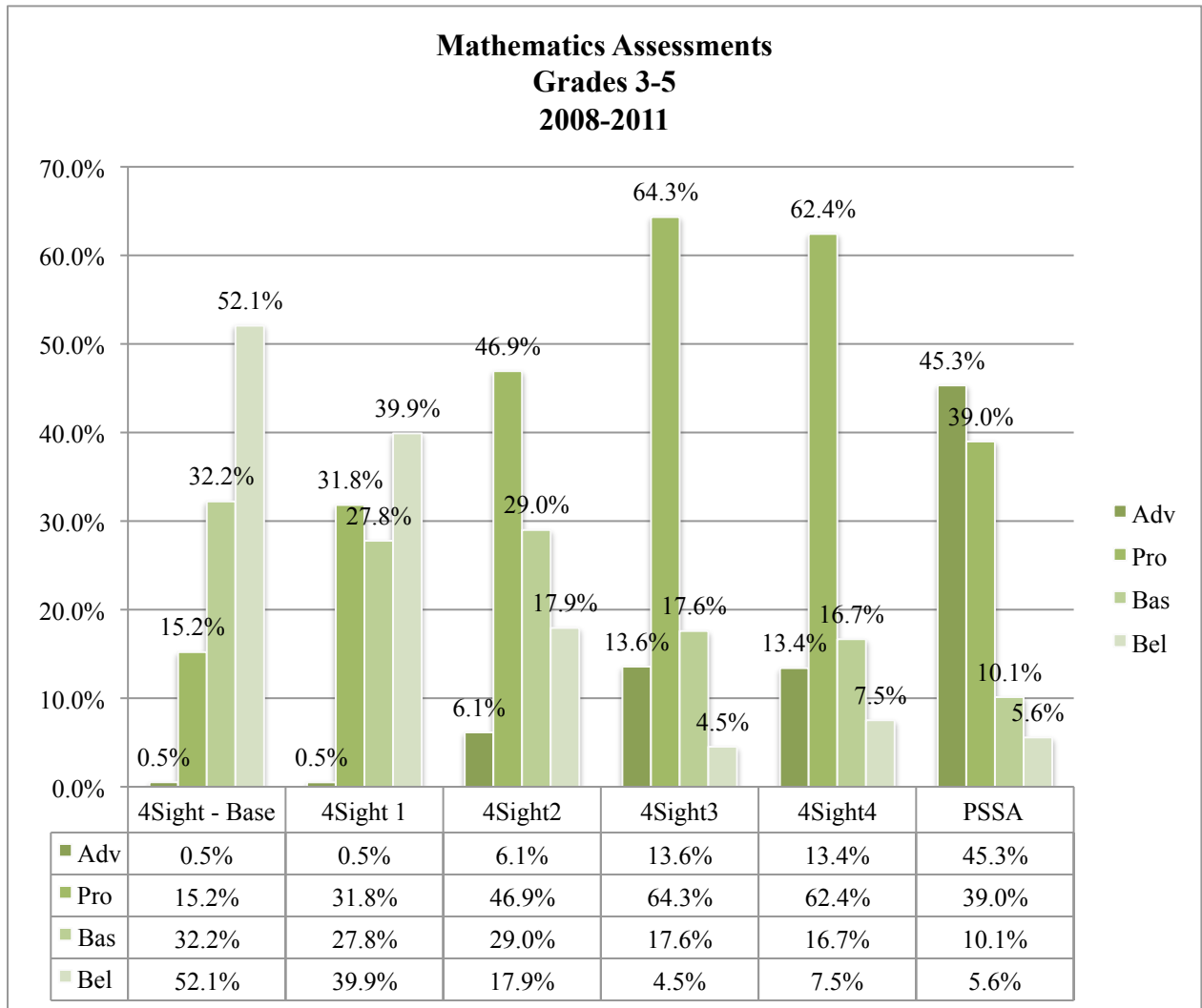
Finally, it is important to recognize that the results from this study apply only to the Oswayo Valley School district. The sample size, particularly in the longitudinal studies of cohort progress, is limited. Further examination of additional cohorts in future years will be necessary to verify the preliminary conclusions reached in this study. In addition, the study addressed only the testing results; a correlation between the two assessments appears to exist. However, multiple variables—that were not considered—influencing test performance preclude reaching a causation effect created solely by the 4Sights.

This focus project, rather than being a conclusion, represents a beginning of professional evaluation on school processes for advancing student learning.

References

- Bernhardt, Victoria L. (2009). *Data, data, everywhere*. Larchmont, NY: Eye on Education.
- Classroom reports (RP-3711)*. Success for All Foundation. Retrieved June and July 2011 from <http://members.successforall.net>.
- Fugate, Carolyn. (2011, June and July). Personal interviews.
- Individual performance results*. Data Interaction for Pennsylvania System for School Assessment. Retrieved July, 2011 from <http://solutions1.emetric.net>.

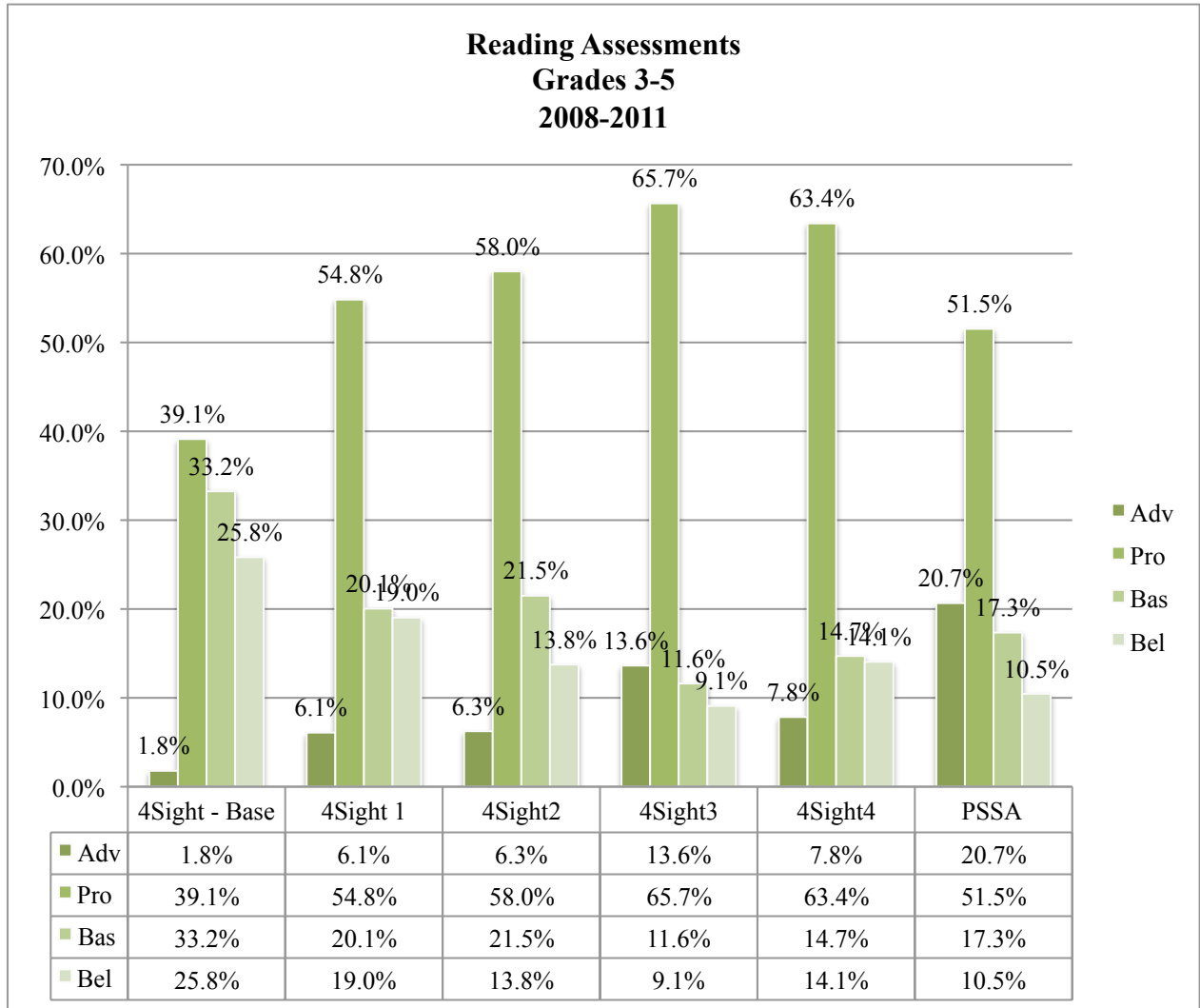
Appendix A



Number of Assessments Administered

4Sight Baseline N=388
 4Sight Test 1 N=396
 4Sight Test 2 N=407
 4Sight Test 3 N=199
 4Sight Test 4 N=306
 PSSA N=395

Appendix B

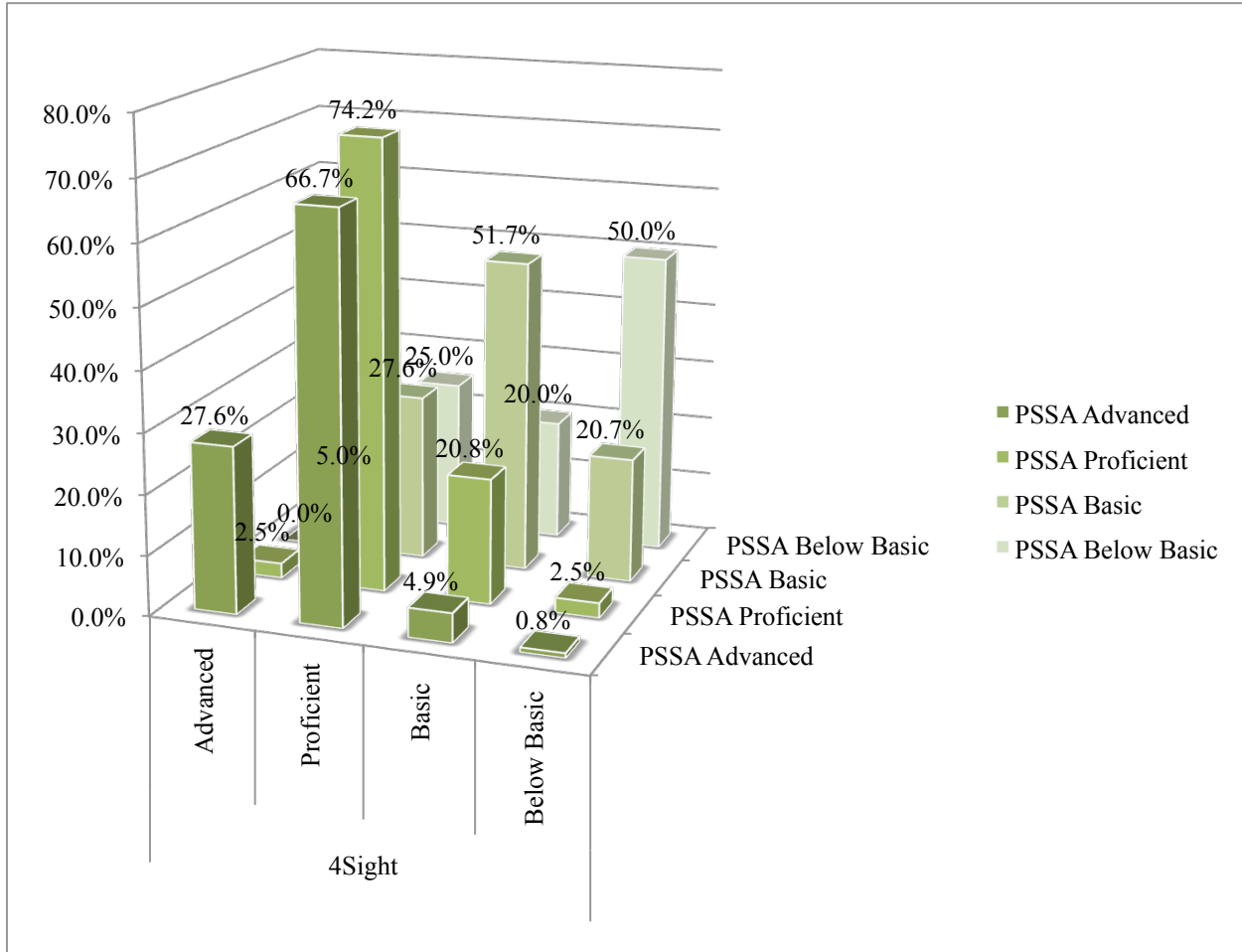


Number of Assessments Administered

4Sight Baseline N=391
 4Sight Test 1 N=394
 4Sight Test 2 N=400
 4Sight Test 3 N=198
 4Sight Test 4 N=36
 PSSA N=392

Appendix C

4th Session 4Sight as Predictor of PSSA Performance
Mathematics Assessment, Grades 3-5, 2008-2011

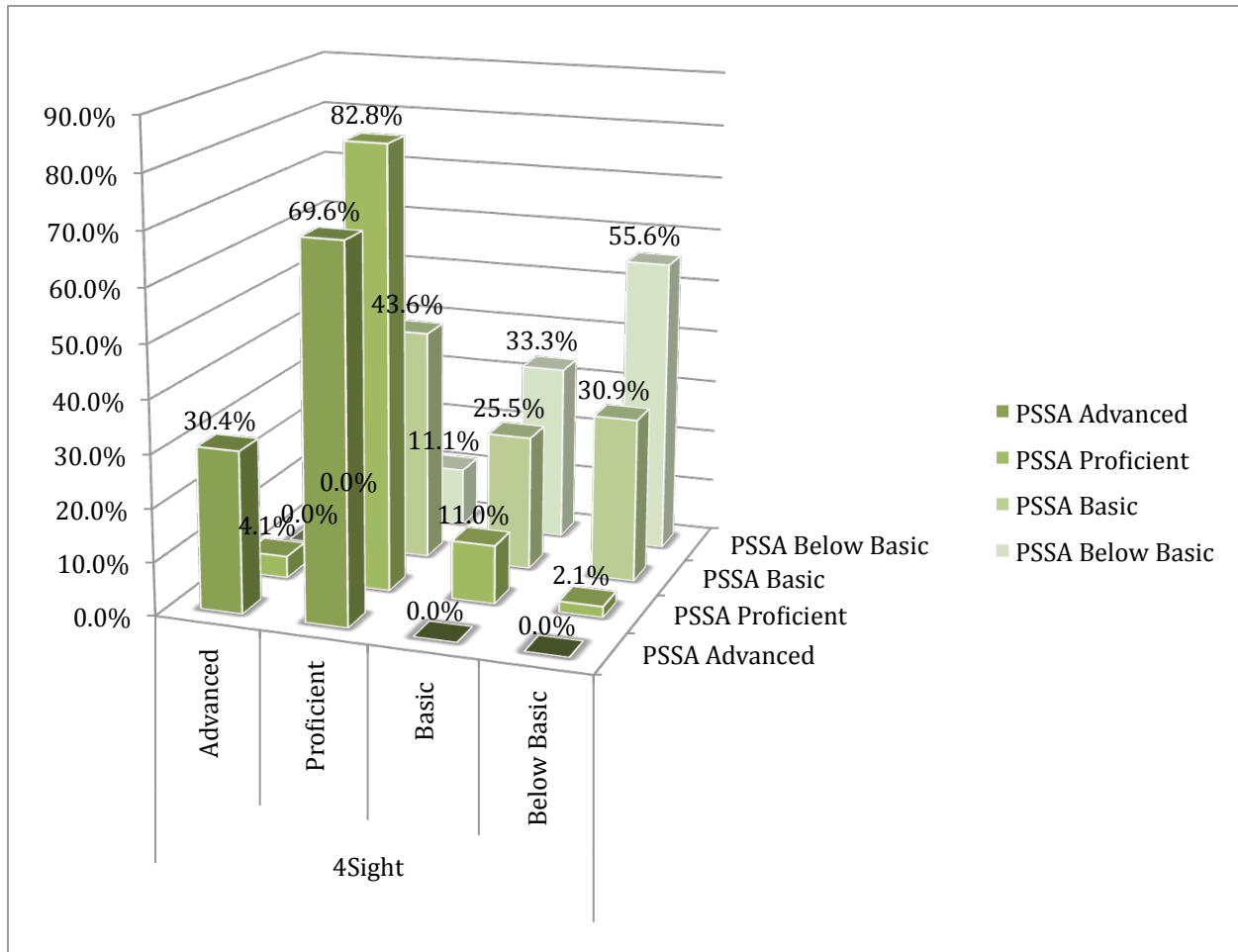


N=292

		4Sight 4 th Session			
		Advanced	Proficient	Basic	Below Basic
PSSA	Advanced (N=123)	27.6% (N=34)	66.7% (N=82)	4.9% (N=6)	0.8% (N=1)
	Proficient (N=120)	2.5% (N=3)	74.2% (N=89)	20.8% (N=25)	2.5% (N=3)
	Basic (N=29)	0.0% (N=0)	27.6% (N=8)	51.7% (N=15)	20.7% (N=6)
	Below Basic (N=20)	5.0% (N=1)	25.0% (N=5)	20.0% (N=4)	50.0% (N=10)

Appendix D

4TH Session 4Sight as Predictor of PSSA Performance
Reading Assessment, Grades 3-5, 2008-2011

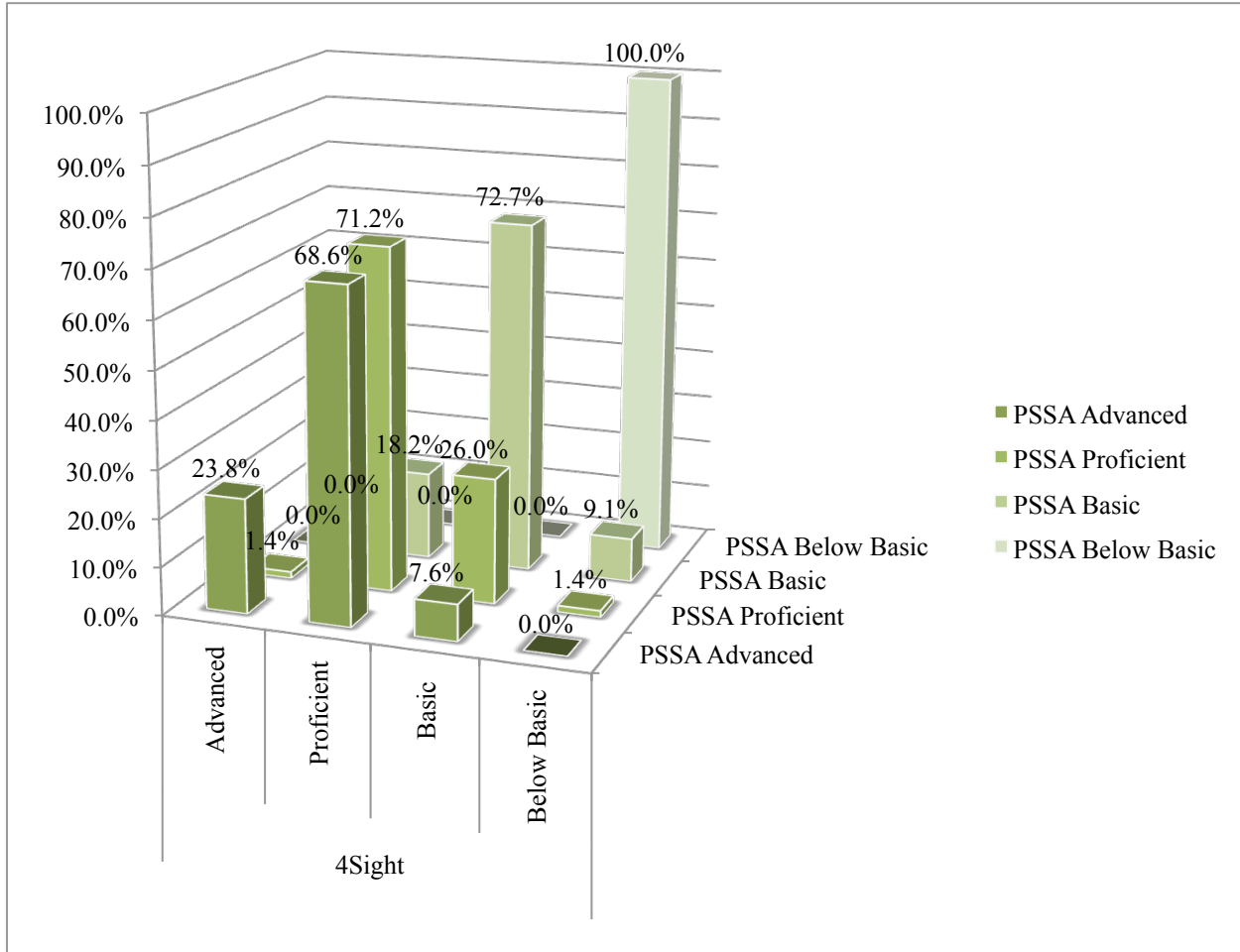


N=292

		4Sight 4 th Session			
		Advanced	Proficient	Basic	Below Basic
PSSA	Advanced (N=56)	30.4% (N=17)	69.6% (N=39)	0.0% (N=0)	0.0% (N=0)
	Proficient (N=145)	4.1% (N=6)	82.8% (N=120)	11.0% (N=16)	2.1% (N=3)
	Basic (N=55)	0.0% (N=0)	43.6% (N=24)	25.5% (N=14)	30.9% (N=17)
	Below Basic (N=36)	0.0% (N=0)	11.1% (N=4)	33.3% (N=12)	55.6% (N=20)

Appendix E

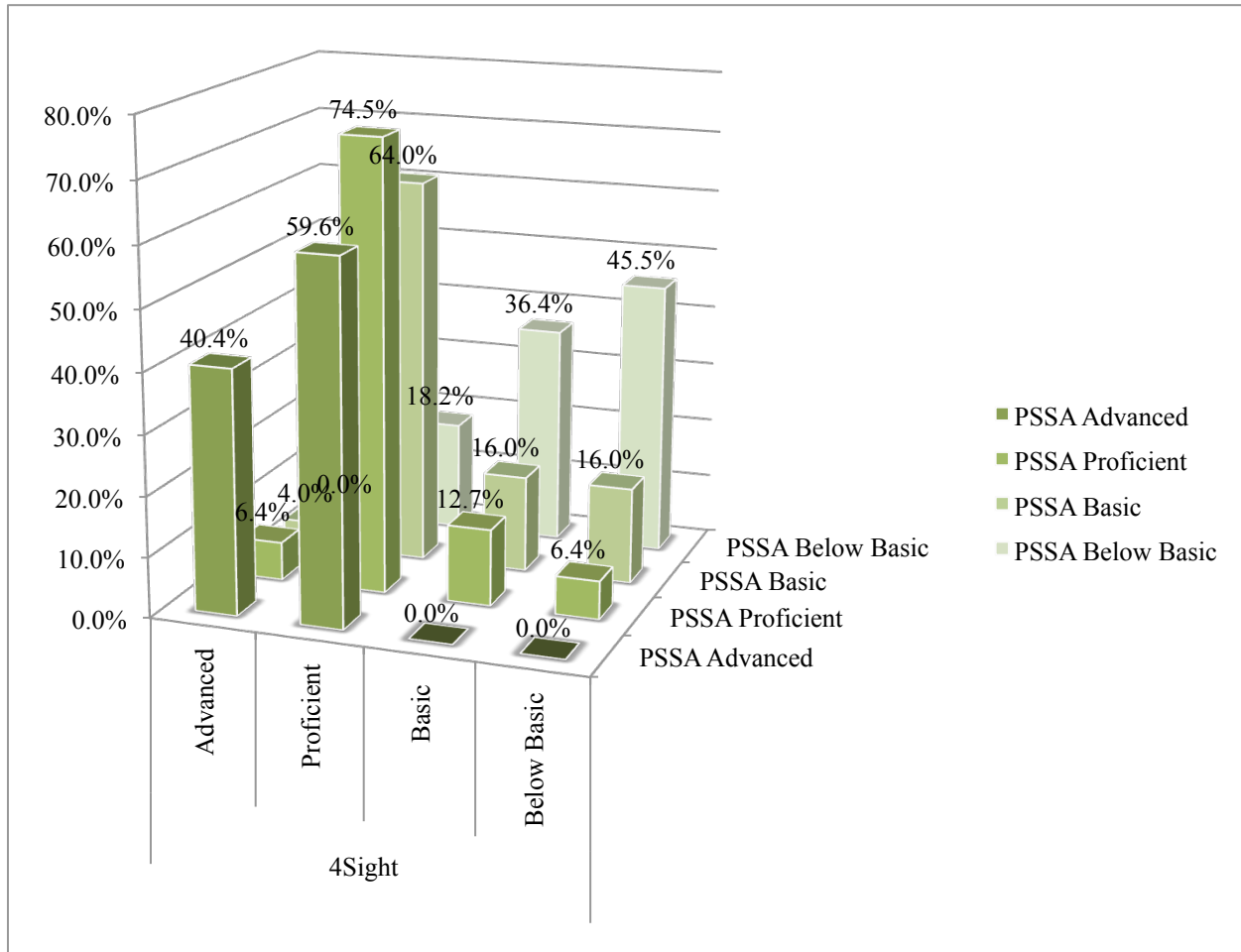
3rd Session 4Sight as Predictor of PSSA Performance
 Mathematics Assessment, Grades 3-5, 2008-2011



		4Sight 3 rd Session			
		Advanced	Proficient	Basic	Below Basic
PSSA	Advanced (N=105)	23.8% (N=25)	68.6% (N=72)	7.6% (N=8)	0.0% (N=0)
	Proficient (N=73)	1.4% (N=1)	71.2% (N=52)	26.0% (N=19)	1.4% (N=1)
	Basic (N=11)	0% (N=0)	18.2% (N=2)	72.7% (N=8)	9.1% (N=1)
	Below Basic (N=4)	0.0% (N=0)	0% (N=0)	0% (N=0)	100% (N=4)

Appendix F

3rd Session 4Sight as Predictor of PSSA Performance
Reading Assessment, Grades 3-5, 2008-2011



N=193

		4Sight 3 rd Session			
		Advanced	Proficient	Basic	Below Basic
PSSA	Advanced (N=47)	40.4% (N=19)	59.6% (N=28)	0% (N=0)	0.0% (N=0)
	Proficient (N=110)	6.4% (N=7)	74.5% (N=82)	12.7% (N=14)	6.4% (N=7)
	Basic (N=25)	4.0% (N=1)	64.0% (N=16)	16.0% (N=4)	16.0% (N=4)
	Below Basic (N=11)	0.0% (N=0)	18.2% (N=2)	36.4% (N=4)	45.5% (N=5)