



**Florida Alternate
Assessment
Student Growth Study
Summary of Results
2010–2011**



Prepared by Measured Progress for the
Florida Department of Education

TABLE OF CONTENTS

CHAPTER 1	OBJECTIVE.....	1
CHAPTER 2	RESEARCH QUESTIONS	3
CHAPTER 3	METHOD	5
3.1	<i>Design</i>	5
3.1.1	Design Specifics.....	5
3.2	<i>Analyses: Phase 1</i>	5
3.2.1	Overview.....	5
3.2.2	Details	5
3.3	<i>Results: Phase 1</i>	8
3.4	<i>Analyses: Phase 2</i>	10
3.4.1	Overview.....	10
3.4.2	Details	10
3.5	<i>Results: Phase 2</i>	11
CHAPTER 4	DISCUSSION	13
APPENDIX A	15
APPENDIX B	25

CHAPTER 1 OBJECTIVE

One goal that the Florida Department of Education (DOE) has for the Florida Alternate Assessment (FAA) is to develop a growth model that can be used to monitor and evaluate growth for students taking the alternate assessment. In order to develop the best model that meets the needs of the DOE, it is necessary to explore a variety of options for measuring growth with this population of students. This study has addressed this goal by using an exploratory approach. Existing FAA data have been examined for students for whom multiple years of scores are available. This report addresses the data analysis portion of the study.

For this study, our analyses are divided into two phases. In Phase 1, a variety of growth models with varying degrees of complexity were analyzed to learn more about the effect of these complexities on estimation of growth. In Phase 2, based on the results of Phase 1 and consultation with Florida DOE and the Technical Advisory Committee (TAC), the simplest model from Phase 1 was compared with another more complex model that incorporates the most attractive features from all the complex models.

CHAPTER 2 RESEARCH QUESTIONS

1. How is growth defined for students who are classified into a Performance Level (PL) lower than 4?
2. For students who are classified into a PL lower than 4, how likely are they to exhibit growth?
3. For students who are classified into a PL of 4 or higher, how likely are they to maintain the growth they have already attained?

CHAPTER 3 METHOD

3.1 Design

Using data obtained from the 2007–08, 2008–09, and 2009–10 testing years, student records were matched across administrations, and trends in students’ performance were analyzed. To do this, it was necessary to obtain data files containing one record for each student that includes item-level scores (i.e., 0/1 scores for each level of the item—participatory, supported, and independent) for each of the three years (or two of the three). Piedra Data Services conducted the matching of the student records and provided the needed data files. In addition, they supplied a crosswalk that enabled Measured Progress Data Analysis and Psychometric staff to correctly identify the items in the data file. Data included reading scores and mathematics scores for all grade levels. To ensure consistency in PL classifications across years, the most recent cut scores were applied to the sum of the item scores in all the years used in the study.

3.1.1 Design Specifics

All students with two or three consecutive years of scores (i.e., 2007–08, 2008–09, and 2009–10 or 2008–09 and 2009–10) were included in the analyses. For each study, the students were separated into two groups: students who were classified as PL 3 or lower (below-proficient students) in the first year of the study, and students who were classified as PL 4 or higher (proficient students) in the first year (the percentages of below-proficient and proficient students for each year, content area, and grade level are given in Appendix A). For example, in studying growth from 2007–08 to 2008–09, the two groups of students were those classified as PL 3 or lower in 2007–08 and those classified as PL 4 or higher in 2007–08.

3.2 Analyses: Phase 1

3.2.1 Overview

In this phase of the analysis, various methods are used to define growth; these conceptualizations then can be compared to see how they affect the results. Briefly, one model looks at growth as measured over a three-year span, while the other four models measure growth over two-year spans. The two-year models differ by how they measure growth for students who are below proficient in the first year. These different models are defined in more detail below.

3.2.2 Details

Here, the details of the five models of growth are described. In all of the two-year models, a student who is classified as PL 4 or higher in the first year (and is thus considered to have attained proficiency) is counted as having exhibited positive growth if the student either maintains his or her current PL or moves up one or more levels in the following year. The differences in the two-year models are in how they measure the

growth of students who are below proficient in the first year of the study, and our Phase 1 results for the two-year models focus on this group of students.

1. Model 1 (across years 2007–08 to 2008–09 and 2008–09 to 2009–2010).

This model was chosen because of its simplicity in adhering to the basic idea that a below-proficient student exhibits growth by either becoming proficient or by moving toward proficiency by advancing one or more PLs.

- a. A growth score of 1 was applied to below-proficient students who moved up at least one PL.
- b. A growth score of 0 was applied to below-proficient students who did not move up or down at least one PL.
- c. A growth score of -1 was applied to below-proficient students who moved down at least one PL.
- d. Growth scores were aggregated across the grade levels within each content area for these years for all below-proficient students.

2. Model 2 (across three years: 2007–08, 2008–09, and 2009–10).

This model extends the basic idea of Model 1 to a three-year approach. That is, a below-proficient student is considered to be exhibiting growth only if the student becomes proficient and stays proficient *or* if the student makes progress every year toward becoming proficient by moving up one PL per year.

- a. A growth score of 1 was given to each below-proficient student who moved up at least one PL each year *or* moved up to or above PL 4 and stayed at or above PL 4 for subsequent years. Proficient first-year students were given a growth score of 1 so long as they stayed at or above PL 4.
- b. A growth score of 0 was applied to all other below-proficient students (for example, students who do not move up at least one PL each year).
- c. Growth scores were aggregated at the grade and subject level.
- d. Growth scores were aggregated across the grade levels with each content area for these years for all below-proficient students.

3. Model 3 (across years 2007–08 to 2008–09 and 2008–09 to 2009–2010).

A concern that can be raised in regard to Model 1 is that a below-proficient student may exhibit growth within a PL by scoring low within the PL one year and then scoring high within it the next year. Model 3 addresses this concern by dividing each below-proficient PL into two pieces and then simply treating these two pieces as separate PLs as in Model 1.

- a. Each below-proficient PL from 1 to 3 was divided into two PL*s ("PL*" denotes the new set of below-proficient PLs) using cut scores that were halfway between the lowest and highest scores corresponding to each PL.
- b. A growth score of 1 was applied to below-proficient students who moved up at least one PL*.
- c. A growth score of 0 was applied to below-proficient students who did not move up or down at least one PL*.
- d. A growth score of -1 was applied to below-proficient students who moved down at least one PL*.
- e. Growth scores were aggregated across the grade levels within each content area for these years for all below-proficient students.

4. Model 4 (across years 2007–08 to 2008–09 and 2008–09 to 2009–2010).

While Model 1 takes a simple approach by modeling growth of below-proficient students in terms of changes in PL, Model 4 takes a different simple approach by modeling growth as measured directly by change in test score. This method takes advantage of the fact that, for every grade level, the assessment instrument has the same range of score points, the same raw score at the proficiency cut, and the same interpretation of proficiency relative to the skills at that grade level. Thus, extrapolating from these score similarities, Model 4 is based on the assumption that the scores on different instruments are comparable in terms of difficulty level. Changes in a below-proficient student's test scores are evaluated to determine growth by comparing the difference score to a standard error (SE) estimate, which takes measurement error into account.

- a. Calculate score change statistic, $Z = (\text{raw score in 2nd year} - \text{raw score in 1st year})/\text{SE}$ of the difference in the two test scores.
- b. If $Z > 1$, then growth score = 1.
- c. If $Z < -1$, then growth score = -1.
- d. If $-1 < Z < 1$, then growth score = 0.
- e. Calculate SE of the difference of two test scores. Here we employ the classical test theory assumption that the error in a student's first-year test score is uncorrelated with the student's error in the second-year test score. The standard error of measurement (SEM) for each test (SEM1 and SEM2 for the 1st year and 2nd year tests, respectively) is estimated separately from the reliability of each test (see Appendix B for a detailed description of the equations for calculating SEM). Thus, the SE of the difference score is given by $\text{SQRT}(\text{SEM1} * \text{SEM1} + \text{SEM2} * \text{SEM2})$.
- f. Growth scores were aggregated across the grade levels within each content area for these years for all below-proficient students.

5. Model 5 (across years 2007–08 to 2008–09 and 2008–09 to 2009–2010).

One problem with Model 1 that was mentioned above is the possibility of below-proficient students attaining meaningful growth within a PL but not having it count because they do not grow enough to move up to the next PL. The flip side of this problem is when below-proficient students attain a very small amount of growth, but have it count as significant simply because the students were next to the boundary between two PLs. Model 5 is a more complicated version of Model 1 that takes these two concerns into account. To identify students who attain growth within a PL, Model 5 calculates the difference of how far they have moved on a proportional basis within the PL (e.g., a student at Year 1 may have a score that is one-third of the way through the PL and move at Year 2 to a score that is two-thirds of the way through) and compares that score difference to the standard error of the difference. To identify a student who has moved up a level but has not exhibited significant growth, Model 5 calculates the score difference for the student relative to the raw-score cuts (one for Year 1 and the other for Year 2) for the boundary between the two PLs. For example, a student at Year 1 may be within one point of the cut score for the next highest PL, and at year 2 may be just two points greater than that cut score (which may be a different score in Year 2). Model 5 would calculate this as a score difference of three points. These score differences are compared to an appropriate standard error to determine if they are substantial enough to be counted as growth.

- a. For students who moved up or down two or more PLs, growth score was calculated as described in Model 1.
- b. For students who stayed in the same PL across both years, a Z statistic was calculated as follows:
 1. $Z = A/\text{SE}$
 2. $A = ((\text{raw score in 2nd year} - \text{low cut in PL})/(\text{upper cut in PL} - \text{lower cut in PL})) - ((\text{raw score in 1st year} - \text{low cut in PL})/(\text{upper cut in PL} - \text{lower cut in PL}))$

3. SE of the difference of the two test scores = $\text{SQRT} ((\text{SEM1}*\text{SEM1}/(\text{upper cut in PL} - \text{lower cut in PL})^2) + (\text{SEM2}*\text{SEM2}/(\text{upper cut in PL} - \text{lower cut in PL})^2))$
- c. For students who moved up or down one PL across both years, the Z statistic was calculated as follows:
 1. $Z = A/SE$
 2. $A = (\text{raw score in 2nd year} - \text{cut which was crossed}) - (\text{raw score in 1st year} - \text{cut which was crossed})$
 3. SE of the difference of the two test scores: $\text{SQRT} (\text{SEM1} * \text{SEM1} + \text{SEM2} * \text{SEM2})$
- d. Once the Z statistic was calculated, the growth scores were determined in the same manner as in Model 4.
- e. Growth scores were aggregated across the grade levels within each content area for these years for all below-proficient students.

3.3 Results: Phase 1

The Phase 1 results for the below-proficient students for the two-year models are summarized in Table 1, aggregated across all the grade levels. The results across the individual grade levels did not show any consistent differences, but the grade-level results can be found in Appendix A. First, note that Table 1 shows that the results for Models 4 and 5 were nearly identical. Because Model 5 was the most complicated model, the closeness of its results with those of Model 4 clearly indicate that the added complexity of Model 5 was not necessary.

Table 1.
Summary of Phase 1 results for two-year student growth models for below-proficient students.

Content Area	Model	Decreased Years		Change Years		% Grew Years	
		1-2	2-3	1-2	2-3	1-2	2-3
Math	1	14.5	12.7	48.6	46.8	36.8	40.4
	3	22.8	19.6	30.1	29.6	47.1	50.8
	4	11.7	8.7	55.4	56.1	32.8	35.2
	5	11.7	8.9	55.5	56.4	32.8	34.7
Reading	1	13.8	11.6	49.4	46.7	36.8	41.7
	3	22.4	18.0	30.1	30.4	47.5	51.6
	4	12.3	9.7	53.5	52.0	34.2	38.3
	5	12.4	9.4	53.4	52.3	34.2	38.2

Note: Years 1, 2, and 3 correspond to 2007–08, 2008–09, 2009–2010, respectively. Results are aggregated across grade levels 3 through 8.

The next focus is comparing Models 1 and 3. Recall that Model 3 doubled the number of below-proficient performance levels; thus, we'd expect to see an increased proportion of students attaining growth as measured by the attainment of the next highest performance level. This is precisely what we observed. For reading, the percentage of students who attained growth increased from 36.8% to 47.5% for when Year 1 was 2007–08, and the percentage increased from 41.7% to 51.6% for when Year 1 was 2008–09. For mathematics, the corresponding results were 36.8% to 47.1% for when Year 1 was 2007–08, and 40.4% to 50.8% for when Year 1 was 2008–09. On average, the increase in those attaining growth was about 10 percentage points. However, there was also the expected concomitant increase in those going *down* a performance level. The average increase in those displaying negative growth was about 8 percentage points.

Now, attention is turned to Model 4 which used simple comparisons of test score differences relative to appropriate standard errors to measure growth, regardless of the PL classifications. The standard errors of the difference scores ranged from 10.9 to 11.6, the mean and median were both 11.2, and the standard deviation was 0.2 (see Appendix A for complete results on standard errors and reliabilities for each grade level and content area). It turned out that the Model 4 results were, perhaps surprisingly, very similar to the results for Model 1. The differences in the aggregate results for percentage of students attaining growth were generally 3 to 5 percentage points, with Model 1 having the higher values. For example, for mathematics for the two-year analysis beginning in 2007–08, Model 1 yielded 36.8% of students attaining growth, while Model 4 yielded 32.8%. The Model 4 results suggest that the increased growth percentages of Model 3 came from test score differences that are mostly not greater than the standard error of those differences. Thus, the growth increases due to splitting the levels in half in Model 3 seem to be due to capitalizing on chance variation and are not reliable.

Regarding the two-year models, Models 1 and 4 were selected for further investigation in Phase 2. The Phase 1 analyses indicated Model 5 was more complex than was necessary, and Model 3 was capitalizing too much on chance rather than indicating real increases in percentages of students attaining growth.

Finally, results are discussed for Model 2, the three-year model that was analyzed. The percentage of below-proficient first-year students attaining growth across all three years was 16.4% for mathematics and 17.9% for reading. As expected, the percent attaining growth across all three years was much less than the percent attaining growth over any two-year span for any of the two-year models (which, recall from Table 1, ranged from about 35% to 52%). This was expected because the requirement that a below-proficient student improves or remains proficient across a three-year span is more stringent than improving or remaining proficient over a two-year span. Although the three-year model yields important information about the nature of student growth, the three-year model was not carried over to Phase 2 for two reasons: (1) the three-year model is not part of the accountability system as the system focuses on growth from one year to the next, and

(2) the Phase 1 results (perhaps because of reason (1)) did not, at this time, result in any further interest in exploring more questions in regard to three-year models.

3.4 Analyses: Phase 2

3.4.1 Overview

After reviewing the Phase 1 results with the Florida DOE and TAC, two two-year models were selected for further investigation in Phase 2. Model 1 was studied in Phase 2 as is, but Model 4 was slightly modified. In particular, Model 4 was modified so that a below-proficient student who increases at least one PL is automatically identified as having attained growth. This new model will be referred to as Model 6. As a matter of fairness, it was deemed important that any student who moves up one (or more) PL deserves to be counted as having positive growth. Thus, the added value of Model 6 is that it also has the ability to detect the growth of students who stay within the same PL from one year to the next by looking at their test score difference and comparing it to an appropriate standard error. It was hoped that, in this way, Model 6 would appropriately estimate an increase in the percentage of students who grew from one year to the next, without capitalizing on chance (as occurred with Model 3) and without penalizing students who increased a PL but with only small increases in test score (a penalty inherent in the original Model 4). Furthermore, the Phase 2 results look only at the percent of students who are estimated as exhibiting growth, without further estimating among the remaining the students the percent who displayed “negative” growth because this concept is not part of the FAA accountability system. For completeness sake, the results will be presented for the below-proficient students as well as for the proficient students.

3.4.2 Details

Here the details of the two models of growth that were studied in Phase 2 are described. In the models below, a student who is classified as PL 4 or higher in the first year (and is thus considered to have attained proficiency) is counted as having exhibited positive growth as long as the student maintains the same or higher PL in subsequent years. The differences in the two models are in how they measure the growth of students who are below proficient in the first year.

Note that Model 1 is slightly modified from Phase 1 since estimation of negative growth is not included in Phase 2.

1. Model 1 (across years 2007–08 to 2008–09 and 2008–09 to 2009–2010).

A below-proficient student exhibits growth by either becoming proficient or by moving toward proficiency by advancing one or more PLs.

- a. A growth score of 1 was applied to students who moved up at least one PL or stayed at or above PL 4.

- b. A growth score of 0 was applied to students below PL 4 who did not move up at least one PL.
- c. Growth scores were aggregated across the grade levels within each content area for these years for all students.

2. Model 6 (across years 2007–08 to 2008–09 and 2008–09 to 2009–2010).

Model 6 takes a closer look at below-proficient first-year students who stay within the same PL in the second year, to see if their scores may indicate growth within that level. This model takes advantage of the fact that, for every grade level, the assessment instrument has the same range of score points, the same raw score at the proficiency cut, and the same interpretation of proficiency relative to the skills at that grade level. Thus, extrapolating from these score similarities, Model 6 is based on the assumption that the scores on different instruments are comparable in terms of difficulty level. Changes in test scores within a given PL are compared to a standard error estimate to take measurement error into account. Growth scores for below-proficient students were calculated as follows:

- a. For students who moved up or down at least one PL, growth score was calculated as described in Model 1.
- b. For students who stayed in the same PL across both years, calculate score change statistic, $Z = (\text{raw score in 2nd year} - \text{raw score in 1st year}) / \text{SE of the difference in the two test scores}$.
- c. If $Z > 1$, then growth score = 1; otherwise, growth score = 0.
- d. SE of the difference of two test scores: $\text{SQRT}(\text{SEM1} * \text{SEM1} + \text{SEM2} * \text{SEM2})$. Here we employ the classical test theory assumption that the error in a student's first-year test score is uncorrelated with the student's error in the second-year test score. The standard error of measurement for each test is estimated from the reliability of each test.
- e. Growth scores were aggregated across the grade levels within each content area for these years for all students.

Detailed step-by-step instructions for implementing Model 6 can be found in Appendix B.

3.5 Results: Phase 2

First, the results did not show any significant variation across grade levels. Thus, only the results aggregated across grade levels are discussed. (The complete results are provided in Appendix A.) Table 2 (below) shows the results for the below-proficient students, as well as for all students, comparing Model 1 with Model 6 for each content area and each of the two sets of years analyzed. No systematic pattern seems to be present for the two different sets of years. Furthermore, the results for mathematics are very similar to the results for reading. The two models, however, do give slightly different results. As expected, when Model 6 takes a closer look at the below-proficient students who stayed within a PL from one year to the next, it finds that some of the students have scores that increased by more than what is expected by chance (as defined by the standard error of the difference score). Thus, Model 6, as hoped for, does indeed identify an increased number of students who have attained growth.

Table 2.
Summary of Phase 2 results for Models 1 and 6 for two-year student growth.

<i>Content Area</i>	<i>Model</i>	<i>Percent Grew</i>			
		<i>Below Proficient Students</i>		<i>All Students</i>	
		<i>Years</i>		<i>Years</i>	
		<i>1–2</i>	<i>2–3</i>	<i>1–2</i>	<i>2–3</i>
Math	1	36.8	40.4	50.2	53.7
	6	41.1	44.0	51.9	55.0
Reading	1	36.8	41.7	52.7	55.6
	6	41.1	46.1	54.3	57.1

Note: Years 1, 2, and 3 correspond to 2007–08, 2008–09, and 2009–10, respectively. Model 1 is the PL model, and Model 2 combines the PL criterion across levels with a standard error criterion within levels.

Next, examine the results for when all students are looked at together (both those who are proficient and those who are below proficient in Year 1). In this case, both models show that a large percentage of the students showed growth simply because a large percentage of the students are proficient in both years and tend to stay proficient over either two-year span. Model 6, as expected, again shows a small increase in the percentage of the students who exhibited growth by taking a closer look at the differences in those student’s within-PL scores and comparing them to the standard error of the difference.

CHAPTER 4 DISCUSSION

By using multiple models with varying assumptions and conditions, Phase 1 of this study revealed that the most statistically accurate growth models (Models 4 and 5) seem to give results that are nearly identical to the simple performance-level based model, Model 1. The Phase 1 results also indicated that the particular three-year model included in the current study places a much greater stringency on student performance that counts as exhibiting growth. Although further examination of this or other possible three-year models was beyond the scope of the current study, further research should be conducted to explore possible modifications.

By combining the performance-level criteria of Model 1 with appropriate standard-error-based within-performance-level testing of Model 4, a Model 6 was developed in Phase 2 of this study that more fully captures student growth. The Phase 2 results showed noteworthy increases in the identification of students who have attained growth, thereby more accurately displaying this aspect of the FAA testing program.

APPENDIX A

Table A.1.
Phase 1 Model 1 results for two-year student growth for below-proficient students.

<i>Content Area</i>	<i>Grade Level</i>	<i>% Decreased</i>		<i>% No Change</i>		<i>% Grew</i>		<i>Sample Size</i>	
		<i>Years</i>		<i>Years</i>		<i>Years</i>		<i>Years</i>	
		<i>1-2</i>	<i>2-3</i>	<i>1-2</i>	<i>2-3</i>	<i>1-2</i>	<i>2-3</i>	<i>1-2</i>	<i>2-3</i>
Math	3	12.9	14.7	44.2	51.5	42.9	33.8	513	456
	4	16.6	11.5	53.9	43.5	29.4	45.0	523	607
	5	16.5	11.5	43.5	48.6	40.0	39.9	648	566
	6	11.7	9.6	47.5	43.9	40.9	46.5	575	512
	7	13.0	12.3	51.9	42.1	35.1	45.6	493	480
	8	16.4	17.1	52.7	52.0	30.9	30.8	457	519
Reading	3	17.0	11.8	46.9	47.1	36.1	41.2	529	476
	4	11.0	11.6	46.9	46.3	42.1	42.2	620	510
	5	17.4	5.7	47.3	41.5	35.2	52.8	579	578
	6	10.2	10.8	48.2	45.7	41.6	43.5	502	446
	7	10.3	16.1	54.3	47.6	35.5	36.3	409	441

Note: Years 1, 2, and 3 correspond to 2007–08, 2008–09, and 2009–10, respectively.

Table A.2.
Phase 1 Model 3 grade-level results for two-year student growth for below-proficient students.

<i>Content Area</i>	<i>Grade Level</i>	<i>% Decreased</i>		<i>% No Change</i>		<i>% Grew</i>		<i>Sample Size</i>	
		<i>Years</i>		<i>Years</i>		<i>Years</i>		<i>Years</i>	
		<i>1–2</i>	<i>2–3</i>	<i>1–2</i>	<i>2–3</i>	<i>1–2</i>	<i>2–3</i>	<i>1–2</i>	<i>2–3</i>
Math	3	21.4	22.1	23.8	31.8	54.8	46.1	513	456
	4	26.4	17.0	33.8	27.8	39.8	55.2	523	607
	5	24.1	16.6	27.6	33.6	48.3	49.8	648	566
	6	18.4	17.0	30.1	26.4	51.5	56.6	575	512
	7	21.1	18.5	32.7	27.1	46.2	54.4	493	480
	8	25.8	27.0	33.7	31.2	40.5	41.8	457	519
Reading	3	26.5	17.4	26.7	30.0	46.9	52.5	529	476
	4	16.9	19.0	29.2	28.8	53.9	52.2	620	510
	5	27.3	10.2	29.2	29.2	43.5	60.6	579	578
	6	17.7	16.8	31.5	28.9	50.8	54.3	502	446
	7	16.4	25.2	35.5	30.4	48.2	44.4	409	441
	8	29.1	21.3	30.3	35.0	40.6	43.8	475	512

Note: Years 1, 2, and 3 correspond to 2007–08, 2008–09, and 2009–10, respectively.

Table A.3.
Phase 1 Model 4 grade-level results for two-year student growth for below-proficient students.

<i>Content Area</i>	<i>Grade Level</i>	<i>% Decreased</i>		<i>% No Change</i>		<i>% Grew</i>		<i>Sample Size</i>	
		<i>Years</i>		<i>Years</i>		<i>Years</i>		<i>Years</i>	
		<i>1–2</i>	<i>2–3</i>	<i>1–2</i>	<i>2–3</i>	<i>1–2</i>	<i>2–3</i>	<i>1–2</i>	<i>2–3</i>
Math	3	8.4	12.3	50.5	61.0	41.1	26.8	513	456
	4	15.7	10.0	58.9	49.6	25.4	40.4	523	607
	5	14.8	5.1	50.2	60.4	35.0	34.5	648	566
	6	6.3	6.3	56.5	54.5	37.2	39.3	575	512
	7	9.7	10.2	61.7	50.8	28.6	39.0	493	480
	8	15.8	9.1	56.2	61.3	28.0	29.7	457	519
Reading	3	12.7	6.9	49.0	55.3	38.4	37.8	529	476
	4	8.1	9.4	52.6	51.2	39.4	39.4	620	510
	5	15.9	4.7	52.5	47.8	31.6	47.6	579	578
	6	9.0	9.4	56.6	52.7	34.5	37.9	502	446
	7	10.5	17.5	56.7	48.5	32.8	34.0	409	441
	8	18.1	11.7	55.2	57.2	26.7	31.1	475	512

Note: Years 1, 2, and 3 correspond to 2007–08, 2008–09, and 2009–10, respectively.

Table A.4.
Phase 1 Model 5 grade-level results for two-year student growth for below-proficient students.

<i>Content Area</i>	<i>Grade Level</i>	<i>% Decreased</i>		<i>% No Change</i>		<i>% Grew</i>		<i>Sample Size</i>	
		<i>Years</i>		<i>Years</i>		<i>Years</i>		<i>Years</i>	
		<i>1-2</i>	<i>2-3</i>	<i>1-2</i>	<i>2-3</i>	<i>1-2</i>	<i>2-3</i>	<i>1-2</i>	<i>2-3</i>
Math	3	8.6	12.1	51.5	61.2	40.0	26.8	513	456
	4	15.1	10.0	59.7	50.4	25.2	39.5	523	607
	5	14.8	6.0	49.8	60.1	35.3	33.9	648	566
	6	6.3	6.4	56.7	54.3	37.0	39.3	575	512
	7	10.1	10.0	61.3	50.4	28.6	39.6	493	480
	8	15.8	9.2	55.6	63.0	28.7	27.7	457	519
Reading	3	14.0	7.6	48.8	54.6	37.2	37.8	529	476
	4	8.1	9.4	52.7	51.0	39.2	39.6	620	510
	5	16.1	4.7	51.8	47.8	32.1	47.6	579	578
	6	9.0	8.5	56.6	53.6	34.5	37.9	502	446
	7	9.8	15.9	57.0	50.1	33.3	34.0	409	441

Note: Years 1, 2, and 3 correspond to 2007–08, 2008–09, and 2009–10, respectively.

Table A.5.
Phase 1 Model 2 grade-level results for three-year student growth for first-year below-proficient students.

Grade Level	Subject			
	Math		Reading	
	% Percent Grew	Sample Size	% Percent Grew	Sample Size
3	15.0	513	19.1	529
4	10.9	523	21.3	620
5	20.5	648	20.9	579
6	21.9	575	21.9	502
7	16.8	493	10.5	409
8	11.2	457	10.5	475

Table A.6.
Phase 2 Model 6 grade-level results for two-year student growth.

Content Area	Grade Level	Percent Grew			
		Below Proficient Students		All Students	
		Years		Years	
		1–2	2–3	1–2	2–3
Math	3	49.1	36.2	60.0	50.0
	4	33.1	49.1	43.0	58.0
	5	43.5	43.1	55.0	50.0
	6	45.0	49.8	54.0	63.0
	7	38.9	48.3	55.0	55.0
	8	35.2	36.2	44.0	54.0
Reading	3	42.2	45.2	53.0	61.0
	4	46.1	47.8	61.0	56.0
	5	38.0	55.7	52.0	64.0
	6	44.6	46.9	63.0	59.0
	7	40.3	40.4	47.0	53.0
	8	33.9	38.7	50.0	50.0

Note: Years 1, 2, and 3 correspond to 2007–08, 2008–09, and 2009–10, respectively.

Table A.7.
Grade-level mathematics proficiency percentages for the corresponding years as needed for studying growth across the 2007–08, 2008–09, and 2009–10 school years.

<i>Grade Level</i>	<i>Year</i>					
	<i>07–08</i>		<i>08–09</i>		<i>09–10</i>	
	<i>% Proficient</i>	<i>Sample Size</i>	<i>% Proficient</i>	<i>Sample Size</i>	<i>% Proficient</i>	<i>Sample Size</i>
3	61.3	1325
4	61.7	1364	65.6	1325	.	.
5	55.7	1462	55.5	1364	63.2	1325
6	59.5	1421	61.3	1462	61.4	1364
7	66.0	1452	64.0	1421	63.5	1462
8	67.8	1420	66.9	1452	69.7	1421
9	.	.	63.5	1420	69.9	1452

Table A.8.
Grade-level reading proficiency percentages for the corresponding years as needed for studying growth across the 2007–08, 2008–09, and 2009–10 school years.

<i>Grade Level</i>	<i>Year</i>					
	<i>07–08</i>		<i>08–09</i>		<i>09–10</i>	
	<i>% Proficient</i>	<i>Sample Size</i>	<i>% Proficient</i>	<i>Sample Size</i>	<i>% Proficient</i>	<i>Sample Size</i>
3	60.4	1336
4	54.8	1372	64.4	1336	.	.
5	60.5	1466	62.8	1372	68.8	1336
6	64.9	1429	60.6	1466	64.8	1372
7	71.9	1457	68.8	1429	69.4	1466
8	66.5	1418	69.7	1457	71.4	1429
9	.	.	63.9	1418	69.2	1457
10	64.1	1418

Table A.9.
Grade-level estimated standard errors and reliabilities for reading.

<i>Grade Level</i>	<i>SEM</i>			<i>SE of Difference</i>		<i>Reliability</i>		
	<i>Years</i>			<i>Years</i>		<i>Years</i>		
	<i>1</i>	<i>2</i>	<i>3</i>	<i>1–2</i>	<i>2–3</i>	<i>1</i>	<i>2</i>	<i>3</i>
3	7.81	7.93	7.72	11.13	11.07	0.96	0.95	0.96
4	7.97	7.97	7.89	11.27	11.21	0.95	0.95	0.96
5	7.94	8.08	7.89	11.33	11.29	0.96	0.96	0.96
6	8.06	7.85	7.67	11.25	10.98	0.95	0.96	0.95
7	7.69	7.79	7.73	10.95	10.97	0.96	0.96	0.96
						0.96	0.96	0.96

Note: Years 1, 2, and 3 correspond to 2007–08, 2008–09, and 2009–10, respectively.

Table A.10.
Grade-level estimated standard errors and reliabilities for math.

<i>Grade Level</i>	<i>SEM</i>			<i>SE of Difference</i>		<i>Reliability</i>		
	<i>Years</i>			<i>Years</i>		<i>Years</i>		
	<i>1</i>	<i>2</i>	<i>3</i>	<i>1–2</i>	<i>2–3</i>	<i>1</i>	<i>2</i>	<i>3</i>
3	8.51	7.85	7.68	11.58	10.98	0.94	0.95	0.95
4	7.69	7.68	7.80	10.87	10.95	0.95	0.95	0.95
5	7.81	7.96	7.87	11.15	11.19	0.96	0.96	0.95
6	8.09	8.05	7.90	11.41	11.28	0.95	0.94	0.94
7	8.05	7.66	7.92	11.11	11.018	0.95	0.95	0.94
8	7.74	8.10	7.93	11.20	11.34	0.95	0.93	0.94

Note: Years 1, 2, and 3 correspond to 2007–08, 2008–09, and 2009–10, respectively.

APPENDIX B

Florida Alternate Assessment “Business Rules”

- I. Students in the first year who are below-proficient count as having exhibited growth in the second year if their scores meet one of the two following criteria:
 1. Their second-year score results in the student moving up one or more performance levels as compared to the first year.
 2. Their second-year score increases relative to their first year score by more than the standard error of the difference of the two scores.

The standard error of the difference is calculated as follows:

1. Calculate the standard error of measurement for the first-year test (SEM1) based on the estimated reliability of the test as calculated from classical test theory.

Cronbach’s α (alpha) is a commonly used estimator of test reliability and is calculated as follows:

$$\alpha \equiv \frac{n}{n-1} \left[1 - \frac{\sum_{i=1}^n \sigma^2_{(Y_i)}}{\sigma_x^2} \right]$$

where

i indexes the item,

n is the total number of items,

$\sigma^2_{(Y_i)}$ represents individual item variance, and

σ_x^2 represents the total test variance.

$$\text{SEM} = \text{square root of } [\sigma_x^2 (1 - \alpha)]$$

2. Calculate the standard error of measurement for the second-year test (SEM2) in the same way using its estimated reliability.
 3. SE of difference = square root of (SEM1*SEM1 + SEM2*SEM2)
- II. Students in the first year who are proficient count as having exhibited growth in the second year if they either stay in the same performance level that they had in the first year or move up one or more performance levels as compared to the first year.