6

Improving Mathematics Accountability

Paul Clopton
Williamson M. Evers

Florida is justly praised for its K–12 accountability system. The state's accountability system is detecting schools where students are not learning enough and where teaching needs to be better. The state's accountability system provides incentives that enhance productivity.[1] Yet accountability is only a part of comprehensive school reform when that reform effort is based on standards, testing, and accountability. And accountability is ultimately hollow if it is not based on rigorous academic standards and tests that measure student mastery of the sub-

The authors wish to thank Paul E. Peterson and Jay P. Greene for their comments and suggestions.

1. See Lance T. Izumi and Williamson M. Evers, "State Accountability Systems," in *School Accountability*, ed. Evers and Herbert J. Walberg (Stanford, Calif.: Hoover Institution Press, 2002), pp. 105–153; Jay P. Greene and Marcus A. Winters, "Competition Passes the Test: Still More Evidence from Florida that Public Schools Improve When Threatened with the Loss of Students and Money," *Education Next*, 4 (Summer 2004): 66–71; Jay P. Greene, "The Looming Shadow: Florida Gets Its F Schools to Shape Up," *Education Next*, 1 (Fall 2001): 76–82. See also Paul E. Peterson, "The A+ Accountability Plan," in this volume.

ject-matter material cataloged in those standards. Hence, to evaluate a school-reform effort, we must also look at testing and at curriculum standards. How sound is the curriculum that forms the basis for teaching in Florida? How tough and thorough are Florida's tests? Could curriculum and testing be improved in ways that would help teachers teach and students learn? Here, we endeavor to address these questions by looking at Florida's mathematics-curriculum standards and its statewide mathematics tests. Our aim is to identify areas in need of change and refinement and to offer insights that may be useful to Florida educators, as well as to those in other states.

Sunshine State Standards

> *The Sunshine State Standards are the centerpiece of a reform effort in Florida to align curriculum, instruction and assessment.*[2]

The standards of learning in Florida, called the *Sunshine State Standards*, were approved by the State Board of Education in 1996. Although the standards are being revised, this report applies to the standards that are currently in force. It is our hope that Florida educators make use of this report during their revision process.

The standards are organized into four grade-spans: PreK–2, 3–5, 6–8, and 9–12. In mathematics,[3] the Sunshine State Standards list learning-objectives in five strands:

- Number Sense, Concepts, and Operations
- Measurement
- Geometry and Spatial Sense
- Algebraic Thinking

2. Sunshine State Standards Grade Level Expectations: Mathematics http://www.firn.edu/doe/curric/prek12/pdf/introma.pdf.

3. The standards are available on the Florida Department of Education web site at www.fldoe.org.

- Data Analysis and Probability

Within each strand are three to five broadly worded standards. For example, Standard 3 in the Data Analysis and Probability strand states: "The student uses statistical methods to make inferences and valid arguments about real-world situations." This standard and the others are repeated across each of the four spans. As such, the standards themselves do not provide more than a conceptual framework for organizing content in a way that is consistent across grades.

Within each grade span, the standards listings include a few statements that provide more detail for that specific grade span. To continue with Data Analysis and Probability Standard 3, in grades pre-K to 2 one of the clarifying statements says that the student "decides what information is appropriate and how data can be collected, displayed, and interpreted to answer relevant questions." This seems to be a rather lofty goal to accomplish by second grade, but its exact meaning is not sufficiently clear. By grades 9 to 12, a clarifying statement for this same standard indicates that the student "explains the limitations of using statistical techniques and data in making inferences and valid arguments." Again, the expectations are not sufficiently clear.

Thus, even with the clarifying statements, the Sunshine State Standards do not stipulate what students are expected to know and be able to do with enough specificity. This means that the standards are not sufficiently detailed to guide curriculum development or testing. This is true at the level of grade spans. The standards are even more inadequate in terms of identifying objectives at particular grade levels.

The standards should therefore be accepted as an outline that provides a structure for organizing content with little of the content actually identified. As an outline that provides a taxonomy for mathematics topics that cuts across all grade spans, the Sunshine State Standards provide a structure that is similar to that in many states.

While a taxonomy may be of some utility, there are two dangers inherent in this design. First, when the outline is filled in with content, it is likely that each grade will be expected to cover a little of each standard in roughly equal proportions, with little depth in any particular area. The unfortunate consequence tends to be a shallow, "mile-wide but inch-deep" coverage of subject-matter topics.[4] The taxonomy approach to standards usually results in a small sampling from each of the offerings at the mathematical buffet each year, largely repeated year after year. This is unlikely to be an effective way to build mathematics achievement.[5]

Second, the taxonomy-oriented structure of Florida's standards de-emphasizes arithmetic in elementary school and algebra in secondary school. Treating the operations of arithmetic as but a part of one strand of the standards in the early grades means that they will receive insufficient attention. Likewise, the topics in algebra are at risk in later grades. This lessening of concentrated attention on arithmetic in the

4. A team of Michigan State University scholars received widespread attention when they looked at American mathematics education using Third International Mathematics and Science Study (TIMSS) data. They received this attention for their claim that the American mathematics curriculum is "a mile wide and an inch deep" and that this broad but thin coverage leads to poor mathematics performance by American students. See William H. Schmidt, Richard Houang, and Leland Cogan, "A Coherent Curriculum: The Case of Mathematics," *American Educator* (American Federation of Teachers), Summer 2002; William H. Schmidt, Curtis C. McKnight, and Senta A. Raizen, *A Splintered Vision: An Investigation of U.S. Science and Mathematics Education* (Boston: Kluwer Academic Publishers, 1997). For an alternative view, challenging both the Michigan State researchers' method of comparative analysis of different countries' curricula and the conclusions they draw from their analysis, see Stan Metzenberg, "Science and Math Testing: What's Right and Wrong with the NAEP and the TIMSS?" in *Testing Student Learning, Evaluating Teaching Effectiveness*, ed. Williamson M. Evers and Herbert J. Walberg (Stanford, Calif.: Hoover Institution Press, 2004), pp. 140–152.

5. E. D. Hirsch Jr., *The Schools We Need and Why We Don't Have Them* (New York: Doubleday, 1996), pp. 26–33; Hirsch, *The Knowledge Deficit: Closing the Shocking Education Gap for American Children* (Boston: Houghton Mifflin, 2006), pp. 115–119. See also R. James Milgram, *The Mathematics Pre-Service Teachers Need to Know*, 2005 http://math.stanford.edu/ftp/milgram/; R. James Milgram and Hung-Hsi Wu, The Key Topics in a Successful Math Curriculum, Dec.19, 2005 http://math.berkeley.edu/~wu/.

early years and algebra later on is harmful to successful mathematics education. Any future revision of the Sunshine State Standards should endeavor to restore an emphasis on arithmetic in elementary school and algebra in secondary school.

Grade Level Expectations

Educators in Florida were aware of the limitations of the Sunshine State Standards and in 1999 began work on Grade Level Expectations for the Sunshine State Standards to fill in the details.[6] For example, the development of "division" across grades includes:

- Grade 3: solves real-world division problems having divisors of one digit, dividends not exceeding two digits, with or without remainders.
- Grade 4: solves real-world division problems having divisors of one digit and dividends of three digits, with or without remainders.
- Grade 5: solves real-world problems involving addition, subtraction, multiplication, and division of whole numbers, and addition, subtraction, and multiplication of decimals, fractions, and mixed numbers using an appropriate method (for example, mental math, pencil and paper, calculator).
- Grade 6: knows the effects of the four basic operations on whole numbers, fractions, mixed numbers, and decimals: uses models or pictures to show the effects of addition, subtraction, multiplication, and division, on whole numbers, decimals, fractions, and mixed numbers.
- Grade 7: knows the effects of the four basic operations on whole numbers, fractions, mixed numbers, and decimals: uses models or pictures to show the effects of addition, subtraction, multiplica-

6. The grade level expectations are available on the web site at www.fldoe.org.

tion, and division on whole numbers, decimals, fractions, mixed numbers, and integers.

This listing provides some features of grade-by-grade learning expectations. In certain cases, they are clear enough, as in the grade 3 and 4 statements. In these cases, the Grade Level Expectations do provide sufficient detail to identify clear and measurable expectations. By grade 5, expectations become less explicit and more confusing because they combine multiple objectives and multiple methods in one statement. Division is limited to whole numbers at this grade level. However, combining so many operations and methods in one statement means that only a small portion of this objective will be tested in any given year. By grade 7, Florida expects its students to know division with fractions only through its "effects" as illustrated with models or pictures. Finally, note that use of electronic calculators is included in the grade 5 expectations. Is the Florida student ever expected to divide by a 2-digit divisor without a calculator? What does it mean to "know the effects" of operations? Is the student expected to learn the standard algorithm for division? Are students ever expected to solve division problems with fractions in the denominator? The grade level expectations do not answer these important questions.

In 2005, the Thomas B. Fordham Foundation updated its review of mathematics standards in the various states.[7] Both the Sunshine State Standards and the Grade Level Expectations were included when Florida was evaluated.[8] The authors gave Florida a final grade of "F"

7. David Klein, Bastiaan J. Braams, Thomas Parker, William Quirk, Wilfried Schmid, and W. Stephen Wilson, *The State of State Math Standards—2005* (Washington, D.C.: Thomas B. Fordham Foundation, 2005) http://www.edexcellence.net/doc/mathstandards05FINAL.pdf.

8. The authors of this chapter agree that the Fordham reviewers were correct to include the Grade Level Expectations in their review. We disagree with Gerry Meisels and Robert Potter, who criticized the Fordham review of Florida's science standards for including the Grade Level Expectations which Meisels and Potter dismiss as "a minor, albeit weak, implementation guide." Without the Grade Level Expectations, even with their shortcomings, the Florida Standards would be operationally useless.

even though they note that the state had received a "D" in earlier evaluations. The review authors were disappointed with the lack of clarity and content in Florida's Standards and Expectations. They also raised other issues. One is the pervasiveness of calculator-use starting in first grade. The Fordham reviewers say that excessive use of calculators may undermine number sense and arithmetic. Indeed, Florida's Standards and Expectations do not explicitly call, as grade levels advance, for mastery of standard algorithms and for computation unassisted by a calculator on the more difficult problems. This suggests that Florida is backing away from expectations of competence in operations. The Fordham reviewers also noted that Florida's Standards and Expectations include multiple expectations dealing with patterns. This emphasis on patterns takes time away from other, more important, content areas. In addition, it is all too often the case that patterns are assumed without any evidence of certainty. Such an assumption flies in the face of the nature of mathematics, which leads to certainty and often proof. Consider the pattern 3, 5, 7, __. What is the next number in the pattern? Is it 9 because the pattern is consecutive odd numbers starting with 3? Or, is it 11 because the pattern is consecutive primes starting with 3? Or is it 3 because the pattern is the continued repetition of the three digits 3, 5, 7? The evaluation by the Fordham reviewers should be seen as a warning flag by Florida policymakers. Revision of the Standards and Grade Level Expectations in light of the Fordham review could only improve these documents. Florida has a great opportunity here to ratchet up rigor as well as grade-level specificity.

Statewide Testing: The Florida Comprehensive Assessment Test (FCAT)

> *The purpose of the FCAT is to make sure that students are learning the basic skills they are supposed to be learning in school. These basic skills*

Compare Meisels and Potter, "Florida's 'F' In Science Reflects Reviewers' Bias," *Tampa Tribune*, March 26, 2006.

are called the Sunshine State Standards, and were created by teachers in the mid-1990's.[9]

Statewide assessment systems typically negotiate a difficult path through the interests of many different groups who see different purposes for the assessments. Some want them for accountability purposes in district or school governance, others want to use student test results to evaluate teachers, still others want the tests to provide diagnostic information to target weaknesses in instruction. Not everyone has the same view of the purpose of the Florida Comprehensive Assessment Test (FCAT). A Member of the U.S. House of Representative from Florida says: ". . . [*T*]*he purpose of the FCAT is to grade our schools and implement high stakes penalties or rewards based on their scores, NOT to see where our students need help to boost their performance*."[10] Our view of state testing is that it should also provide information that helps educators improve student achievement.

The Florida Department of Education's official characterization of the FCAT makes it sound as if the state is saying that the FCAT is a minimum-competency examination system.[11] In other instances, however, the Department refers to the test as "rigorous," implying that success on the test requires more than minimal competence. Here we evaluate the FCAT by looking at sample items and released items provided by the state.[12]

The FCAT Mathematics test is given in grades 3 to 10. It contains items in multiple-choice, response-grid, and written-response format

9. FCAT: The Florida Comprehensive Assessment Test http://www.firn.edu/doe/sas/fcat/pdf/fcatpurpose.pdf.

10. U.S. Representative Jim Davis, Dec. 13, 2001 news release http://www.house.gov/jimdavis/press_releases/pr011213.htm.

11. From 1977 until 1998, when the High School Competency Test (HSCT) was last used, Florida acknowledged that its high-school exit examination tested minimum competency. See Elena Llaudet, "Eight Years of Reform," in this volume. At that time, the pass rate on the mathematics portion of the test was 77%. See www.fldoe.org.

12. Both are available from the Department web site at www.fldoe.org.

in various quantities across grade levels. The state publishes school results and uses them in the state accountability system. A student must pass the grade 10 mathematics examination to graduate from high school with a standard diploma.

FCAT Mathematics Grade Level Estimate

We examined each of the sample items provided on the Florida Department of Education's web site (www.fldoe.org) for the FCAT Mathematics examination.[13] We used the California standards (which have received the highest ratings from the American Federation of Teachers and the Fordham Foundation and are used as a benchmark by Achieve) to estimate the grade level of the FCAT sample items.[14]

Analyzing the sample items, we found that the FCAT items started out close to California's high-level standards. However, the rigor of Florida sample items fell farther and farther behind as grade-levels increased. We conclude from this exercise that the FCAT is best characterized as a minimum competency examination.[15]

13. On the Florida Department of Education's policy on releasing test questions and its costs, see "Real FCATs Released Online," *Tallahassee Democrat*, October 19, 2005; Jim Saunders and Linda Trimble, "Don't Fear FCAT; Try the Test on for Size," *Daytona Beach News-Journal*, September 22, 2005.

14. The California standards can be found at http://www.cde.ca.gov/be/st/ss/mthmain.asp. The authors know these standards well. One author served on the 1996–98 California State Commission for the Establishment of Academic Content and Performance Standards and on that commission's mathematics subcommittee. The other author served on the 1997 California Mathematics Curriculum Framework Committee. Both authors are members of the California Mathematics Assessment Review Panel that monitors the content of statewide standards-based tests in mathematics.

15. In a related matter, Paul E. Peterson and Frederick M. Hess compare proficiency rates on state tests and the National Assessment of Educational Progress (NAEP). They give Florida a "C" grade for 4th- and 8th-grade mathematics because the FCAT results show more students are proficient than the NAEP results do. Peterson and Hess, "Keeping an Eye on State Standards: A Race to the Bottom?" *Education Next*, 6 (Summer 2006): 28–29. See also Peterson and Hess, "Johnny Can Read . . . In Some States: Assessing the Rigor of State Assessment Systems," *Education Next*, 5 (Summer 2005): 52–53.

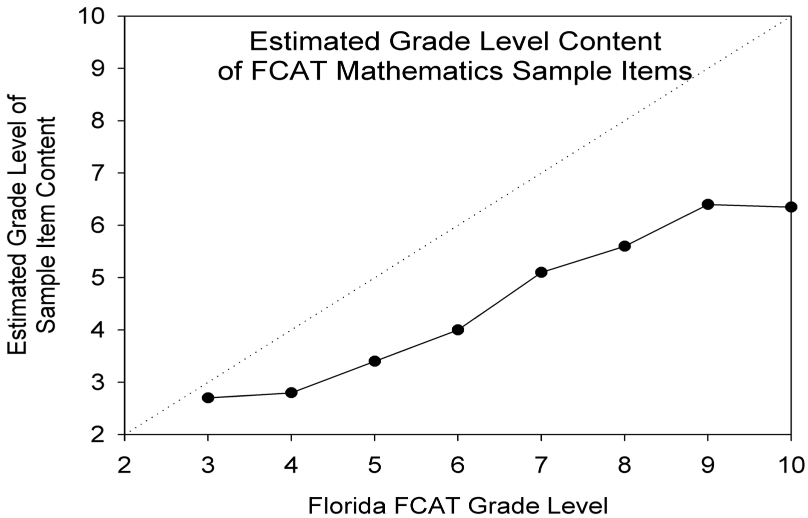


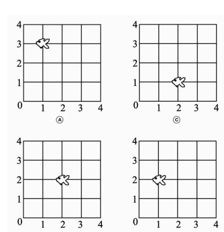
Figure 1. Estimated Grade Level Content of FCAT Mathematics Sample Items

While evaluating these items, we noted instances of similar problems in different grades. As noted above, the Sunshine State Standards set forth achievement goals according to grade spans. Despite the fact that the Grade Level Expectations attempt to provide grade-by-grade learning expectations, the FCAT Mathematics test often fails to adequately differentiate by grade level. We illustrate this inadequate differentiation in the cases of "points in the first quadrant" and "tree diagrams."

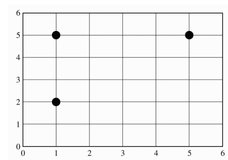
Points in the First Quadrant

Here are sample items addressing point plotting in the first quadrant.

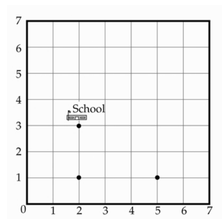
- Grade 3 Item 3
  Which grid shows the fish
  located at (1, 2)?



- Grade 4 Item 14
  Ashley plotted 3 points on a grid, as shown below. These points are 3 of the 4 vertices needed to make a rectangle. Which ordered pair best represents the location of the fourth vertex needed to complete the rectangle?



- Grade 5 Item 9
  The track team warms up for practice by jogging through the neighborhood near the school. A coordinate grid of the neighborhood is shown below. The team runs from the school along a path that forms a rectangle. Three of the corners of the rectangle are shown on the grid. What are the coordinates of the fourth corner of the rectangle?



Very little progress is suggested across these grade levels, particularly between grades 4 and 5. The problems in these two grades are virtually the same, there are just more words in the problem statement for 5th grade students.

Tree Diagrams

A type of question that is increasingly found on tests across the county asks students to do simple combinations. Students are not expected to know the fundamental counting principles that yield the equations for permutations and combinations. They merely elaborate all the possible combinations (often in tree diagrams) and count them, as long as the problem is simple enough. Consider:

- Grade 4 Item 13
  Sierra is designing a poster. She may choose 2 of the following colors.

  RED YELLOW BLACK GREEN

  How many different 2-color combinations can be made from the colors listed?

- Grade 10 Item 13
  The Regal High Debate Club qualified for the state meet. How many possible two-member debate teams can be formed from a pool of six students in the Regal High Debate Club?

In grade 4, students count combinations of two from a set of four. By grade 10 (the level of achievement required in Florida for a high school diploma) students count combinations of two from a set of six. Either the Grade Level Expectations or the process of developing test questions is failing to provide sufficient grade-level differentiation. This can easily send the message that a teacher should repeat the same content grade after grade.

A positive feature of the FCAT system is that the state has produced test-item specifications for each grade level.[16] However, as the items above illustrate, even these specifications have been insufficient to assure that test questions distinguish between achievement levels in a meaningful way from grade to grade.

The 10th grade math examination (together with the 10th grade reading examination) is a requirement for high school graduation in Florida. Students are given five additional chances to pass it if they are not successful on the first attempt.[17] Because this is a high-stakes examination, because it constitutes the highest level of assessment in the Florida K–12 system, and because a complete released examination is available, we are considering it in greater detail.

Characterizing the Grade 10 Mathematics Exam

In 2005, Florida released the Grade 10 Mathematics exam that was used in 2004.[18] Evaluation of a complete test provides a more accurate

16. See, for example, *FCAT Mathematics Test Item Specifications, Grades 3–5, Version 3* (2005) available from www.fldoe.org.

17. See Alan Richard, "Graduation Requirements Put Gov. Bush to the Test," *Education Week*, April 23, 2003; Michelle Galley, "Florida and North Carolina Modify Use Of Tests in School Ratings," *Education Week*, March 21, 2001.

18. The examination and the answer booklet are available at www.fldoe.org.

picture than reviewing sample items from a variety of tests given in grades 3 through 10.[19] The 10th-grade exam is a 50-item test with a total possible score of 60—some items are given more than one point credit and partial credit is possible. Of the one-point items, some use a response grid (students must record the actual numerical answer), some are multiple-choice in format, and some require students to provide answers in written form. Students are allowed to use calculators, and the state supplies a detailed reference sheet, so students don't need to remember formulas and other facts.

The 10th-grade mathematics test also serves as part of the high school graduation requirement, so students are given multiple chances to achieve a passing score. The cut score for a passing grade can vary from year to year (according to test difficulty), but it is typically between 16 and 24 points out of 60.[20] That's roughly between 27% and 40%, meaning students could get most of the test wrong and still pass the test and meet the graduation requirement.

To provide a good sense of the level of mathematics involved in this examination, we will present six of the one-point items from the test together with our estimation of the grade level of the material in the items.[21]

The Most Difficult Items

Here are the most difficult items—that is, those with the lowest numbers of students getting a correct answer.

19. Although the released samples are not necessarily representative of the full set of live test items, many of them appear in the official item-writing *Specifications*.

20. How is the total score on FCAT determined? http://www.fldoe.org/faq/default.asp?Dept=202&ID=661#Q661. For an account of setting the cut scores, see Stephen Hegarty, "Educators Wrestle with FCAT," *St. Petersburg Times*, Sept. 20, 1998.

21. We are presenting one-point items here because the statewide means for these items translate directly to the percentage of students responding correctly to the item.

- Item 38

  A bicycle store owner sells all merchandise at 25% above original cost. A customer bought a bicycle for $140.00. How many dollars did the store owner originally pay for the bicycle?

Only 12% of students got this item correct, which was the lowest rate of the one-point items on the entire test. It is a grid-response item, meaning students must record the answer rather than selecting from among choices. This question elicits a common erroneous response, in which students assume that if they reduce the price by 25%, the result will be the original cost. In fact, the price must be reduced by 20% to get back to the original cost. Think of the current cost as 125/100 times the original price. To get back to the original cost multiply the current cost by 100/125 or 4/5.

The mathematics in this problem requires some experience with percentages, and is not beyond the 6th or 7th grade level. Another solution is to set up the problem as a simple equation and solve it:

$$140 = p \cdot 1.25 \text{ where } p \text{ is the original price.}$$

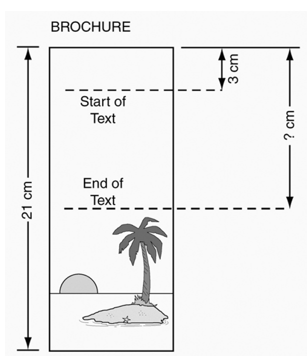
Students merely divide both sides by 1.25 and the problem is solved—and they have a calculator available for the arithmetic. Even this more formal solution is 6th or 7th grade material.

- Item 39

  A graphic artist is creating a brochure according to the following specifications.

  - The length of the brochure is 21 centimeters.
  - The text must start 3 centimeters from the top of the brochure.
  - The total amount of space the text occupies must not exceed one-third of the total length of the brochure.

  What is the maximum distance in centimeters (cm) between the top of the brochure and the end of the text?



This is another grid-response item. It is a multiple step problem that can be solved in a variety of ways. The maximum text length is 21cm divided by 3. That is 7 cm. The bottom of the text is therefore 3cm + 7cm from the top, or 10cm. This is a fairly clear two-step word problem.[22] The material is certainly not above the level expected in 7th grade, if not earlier.

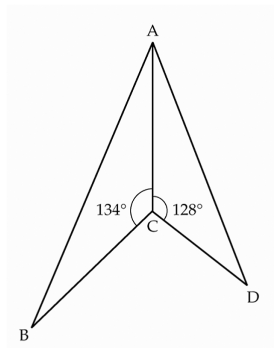
Moderately Difficult Items

Here are two items from the middle of the percent-correct distribution.

- Item 44

  The diagram below shows a design found on a mask from Nigeria. In the diagram, ∠ACB measures 134°, and ∠ACD measures 128°.

  What is the measure, in degrees, of ∠BCD?



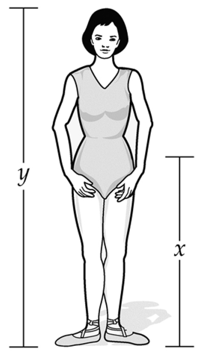
Of course, the solution to this problem is to subtract both 134° and 128° from 360°. The arithmetic is a 2-step subtraction problem, certainly not beyond 6th grade level material. This problem also requires knowing that a full turn or a circle contains 360°. This fact is not always explicitly stated as a standard, but can be found in math materials nationwide from 4th to 7th grade.[23]

- Item 18

22. Some districts have used the presence on the FCAT of multiple-step problems as a justification for imposing discover-learning math instruction (such as the Investigations and Connected Math programs) on their teachers and students. For example, see Tia Mitchell, “Duval Gives ‘New Math’ Good Grade: Supporters Say It Helps Students Understand, Not Memorize, But to Others It’s Just Confusing,” Jacksonville *Florida Times-Union*, June 26, 2005. For an introduction to the drawbacks of discovery-learning math, see Barry Garelick, “An A-Maze-ing Approach to Math,” *Education Next*, 5 (Spring 2005): 28–36.

23. The 4th grade reference can be found in the California Mathematics Framework 2000 edition, while the 7th grade reference can be found in the standards for Seattle Schools at http://www.seattleschools.org/area/acastan/stan/math/Math7.xml.

Artists have traditionally studied human proportions to draw human figures realistically. When drawing a female figure like the one in this picture, the realistic ratio of the distance from the hip to the toe (x) to the height of the woman (y) is 0.613. An artist is creating a 9-inch high drawing of a woman. What should be the approximate distance in inches from the hip to the toe?



A. 0.07 inch
B. 3.5 inches
C. 5.5 inches
D. 14.7 inches

Here students need to find 0.613 of 9 inches, so they need to do the multiplication. Students can use calculators for the arithmetic. But, they really don't need to bother with this. They know the hip-to-toe distance is a little more than half of nine inches, so the answer has to be *C*. This material is not beyond a 6th grade level.

Low Difficulty Items

Here are two items with high numbers of students getting the correct answers.

- Item 49

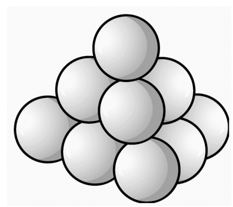
  A salesperson's total salary includes a base pay of $500 per month plus 8.5% of the monthly sales.

  If x = sales per month and y = total salary, which of the following shows how to determine the total salary for any month?

  A. 500 − y = 0.085x
  B. y = 500 + 0.085x
  C. y = −500 − 0.085x
  D. −0.085x + y = −500

This is a problem of rather classic form asking the student to write (select) a simple equation that represents a problem situation. The students have to know something about percentages also. This is 7th grade material in most programs, and 87% of the students got this problem correct.

- Item 28
  The pirates who plundered ships in the Caribbean Sea used cannons in their attacks. Cannonballs were stacked on the deck in a square pyramid like the one shown below. The top layer had 1 ball, the second layer had 4 balls, and the third layer had 9 balls. If the pyramid were 5 layers high, how many cannonballs would be in the fifth layer?



  A. 25
  B. 39
  C. 44
  D. 55

There are various ways to solve this problem. The most sophisticated might approximate 7th grade material, although it could be argued that this is 4th grade material.

Overall Impression

The illustrative items given above are a good representation of the level of mathematics achievement addressed by the 10th grade examination. The vast majority of the items are written at the 6th and 7th grade level.

The test designers are to be commended for including the response-grid format for several items. This format requires students to come up with the actual correct answer rather than relying on test-taking skills, estimation, and ruling-out options that so often can be used on multiple-choice items.

Despite some positive features, however, the most striking feature of the examination is that the level of achievement required by the 10th grade examination is not reflective of 10th grade work, or even

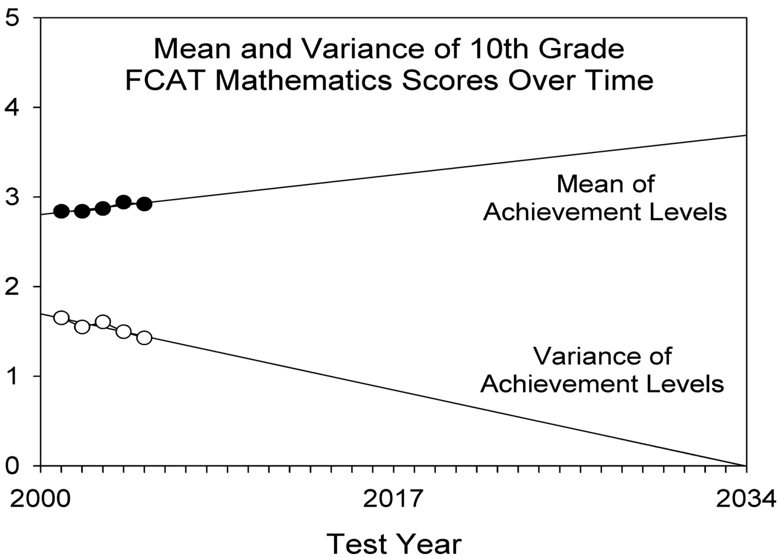


Figure 2. Mean and Variance of 10th Grade FCAT Mathematics Scores Over Time

8th grade work. This is not a test of high school algebra and geometry by any stretch of the imagination.

Consequences

What are the consequences of having a minimum-competency test in high school? Theoretically, it could serve the purpose of backing up the high school diploma with some evidence of competence. However, it is of limited utility beyond that point. An examination at such a low level is unlikely to encourage improved achievement in high school mathematics when high school mathematics itself isn't even tested. Worse yet, if too much attention and course time is devoted to FCAT-level material, students may actually suffer by losing the opportunity for more appropriate, grade-level instruction.

Indeed, the FCAT could encourage greater achievement among the weakest performing students to the detriment of the great majority of students who are achieving closer to grade level if not at or above grade level.[24] Taken to its extreme, the tendency in Florida under the 1996 standards could wind up making sure all 10th grade students end up with mediocre 7th grade abilities as depicted in Figure 2. We computed linear regressions based on 10th grade achievement levels and their variances for recent years, and carried the results forward to the year 2034 to illustrate the point.

College Board SAT Trends

Recent College Board SAT scores provide one indicator of high school mathematics achievement in Florida.[25] As shown in Figure 3, the mean SAT Mathematics score for Florida is decreasing slightly while that for the country as a whole is increasing.[26] Of course, the population of SAT test-takers in Florida is changing, but this is also true in many other states. Over the same period, for example, New York had math score gains equal to that of the U.S. and had a greater increase in participation rates than seen in Florida.[27] This result further reinforces our concerns about high school mathematics in Florida.

24. A Florida parent writes: “It has been numerous years that the FCAT has been well below the education level that my son and his friends are attaining at their public high schools. . . . Much of the high school FCAT material these students knew before middle school. . . . If a public high school student passed algebra in the 7th grade and is taking calculus in the 10th or 11th grade, why is this student losing classroom hours for FCAT-related tasks?” Mary Rapkin, “FCAT Harms Advanced Students, Too, Not Just the Struggling Ones,” Letter to the Editor, *Palm Beach Post*, March 25, 2006.

25. For a different interpretation of these College Board SAT results for Florida, see Paul E. Peterson, “The A+ Plan,” in this volume. One should allow for the possibility that changing demographics may partially lessen the significance of the relative declines in College Board SAT mathematics scores.

26. SAT scores available at http://www.firn.edu/doe/evaluation/pdf/SAT.pdf.

27. Based on College Entrance Examination Board reports, “College-Bound Seniors: 1998 Profile of SAT Program Test Takers,” and “2005 College-Bound Seniors” national and state reports.

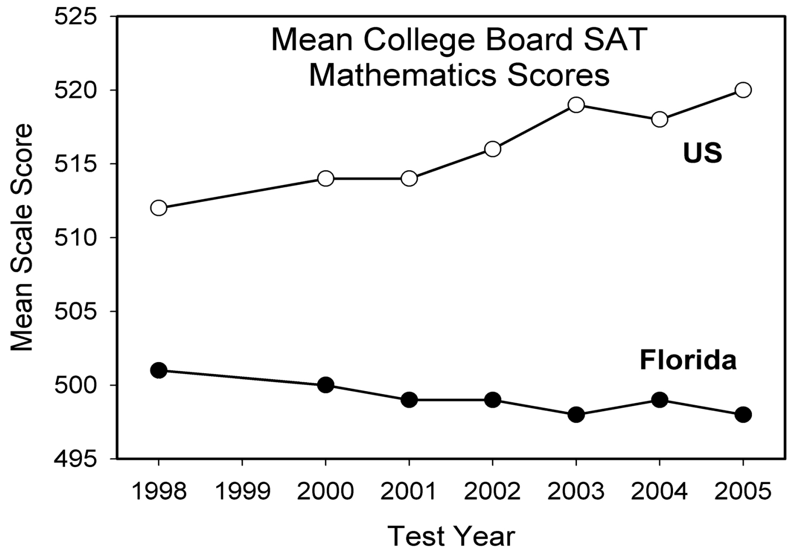


Figure 3. Mean College Board SAT Mathematics Scores

This is not to say that the problem is only present in high school. Students in lower grades may become farther and farther off track for eventually learning high school algebra and geometry as they progress through the grades, and a minimum-competency testing system may contribute to this result.

Conclusion

Florida is currently working on a revision of its math standards. We recognize the political difficulty of revising and strengthening the FCAT mathematics test in a political climate in which opponents of accountability are preying upon the worries of parents of students who might not pass the FCAT. But we are forced to conclude that currently Florida does not have a testing system in place that can measure achievement in high school mathematics. As such, this leaves the state without a real accountability system that applies to high school mathematics. Furthermore, as students approach high school, the standards and testing system is not supportive of greater levels of success in high school mathematics. Therefore, our recommendations for Florida are:

1. Develop a set of specific course-level standards for high school mathematics that are rigorous and truly world-class.[28]
2. Develop end-of-course examinations for algebra and beyond to match these new standards and integrate the examinations into the accountability system.
3. Revise standards and grade level expectations through grade 8 so they align better with high school achievement goals.
4. Revise assessments prior to grade 8 so they will target progressively higher levels of achievement and encourage algebra readiness.

Increasing the rigor of standards and assessments can be difficult politically, yet there are workable options that make use of graduation requirements, diplomas, and transcripts—options that Florida policymakers should consider. In California, for example, the high school course-level examinations are a part of the accountability system for schools, and results are reported in the news media, but course exams are not required for high school graduation. California has a separate, minimum-competency high-school exit examination. Other alternatives include providing differing types of diplomas for various levels of student achievement or having student transcripts that list course-exams passed. We encourage Florida policymakers to adopt one of these alternatives or combine them in a reasonable way. Fixing the fit between Florida's testing system and its requirements for graduation,

28. Florida already has course descriptions for high school mathematics courses. The Thomas B. Fordham Foundation review noted earlier addresses these descriptions briefly, and the comments given there suggest that these course descriptions would not provide an adequate basis for course-level standards.

its diplomas, or its transcripts will allow policy-makers to improve the specificity and rigor of Florida's mathematics Standards and Expectations and allow the state's tests to match these enhanced academic goals. Only then will Florida's academic standards set forth mathematics content that is truly at the high-school level, and Florida's tests become true measures of accountability.