What do the NAEP math tests really measure?

David Klein[1]
July 2010, Revised August 2010

The National Assessment of Educational Progress (NAEP) – often called the nation's report card – includes math tests administered to representative populations of 4th and 8th grade students in each state.[2] Widely cited by education researchers, the scores are used to track the progress (or lack thereof) of U.S. mathematics education, and they have appreciable influence on national education policies. Until recently, the past two decades have seen 4th and 8th grade NAEP math scores increase, albeit slowly. The National Council of Teachers of Mathematics claimed credit for the trend, pointing to its standards documents and policies,[3] but scores were flat from 2007 to 2009. That prompted U.S. Education Secretary Arne Duncan to call for reforms to "accelerate achievement." David Driscoll, chair of the National Assessment Governing Board of NAEP argued that the lack of improvement demonstrated the need for better training of elementary school math teachers [1]. But what do the math NAEP tests actually measure?

Most problems from the NAEP exams are reused, and are therefore kept secret. However, some items each year are retired and made available to the public. The National Center for Educational Statistics suggests that these released items can be used to "supplement classroom instruction" and "provide additional insight into the content of the assessment" [2]. Two of the algebra problems administered in 2009 are displayed in Figs. 1 and 2.

8. Kiara set her beads on a table to make a repeating pattern. Some of the beads rolled off the table.

Here is what was left.

Which of the following should Kiara use to replace the missing beads in the pattern?

A.

B. 

C. 

D. 

Figure 1: Grade 4 NAEP, 2009, Content Classification: Algebra

[1]Department of Mathematics, California State University, Northridge, david.klein@csun.edu

[2]The occassionally administered 12th grade math test is not discussed here, nor the long term trend tests.

[3]See the NCTM webpages at www.nctm.org/news/release.aspx?id=766 and www.nctm.org/news/content.aspx?id=12500.

Write the next two numbers in the number pattern.

1 6 4 9 7 12 10 ____ ____

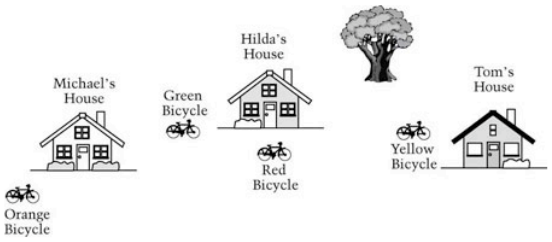
Write the rule that you used to find the two numbers you wrote.

Figure 2: Grade 8 NAEP, 2009, Content Classification: Algebra

The obvious lack of mathematical content in these items is unfortunately not limited to them. Algebra problems constitute 15% of the 4th grade test and 30% of the 8th grade test [3]. In its 2008 report, the National Mathematics Panel (NMP) found that, "At Grade 4, most of the NAEP algebra items relate to patterns or sequences." [4] (pg. 59). The Task Group on Assessment for the NMP reported [5] (pg. 8-9) that,

> While the inclusion of patterns in textbooks or as state curriculum expectations may reflect a view of what constitutes algebra, patterns are not emphasized in the curricula of high-achieving countries. ... The prominence given to patterns at the preschool through Grade 8 level is not supported by comparative analysis of curricula or by mathematical considerations. ... the Task Group strongly recommends that algebra problems involving patterns be greatly reduced in these tests.

Not explicitly pointed out in the NMP reports is that the geometry problems are as devoid of mathematical content as the algebra problems. Examples are shown in Figs. 3 and 4.



1. Jan left her bicycle between Hilda's house and Michael's house. What color is Jan's bicycle?

A. Orange
B. Green
C. Red
D. Yellow

Did you use the calculator on this question?

○ Yes ○ No

Figure 3: Grade 4 NAEP, 2007, Content Classification: Geometry

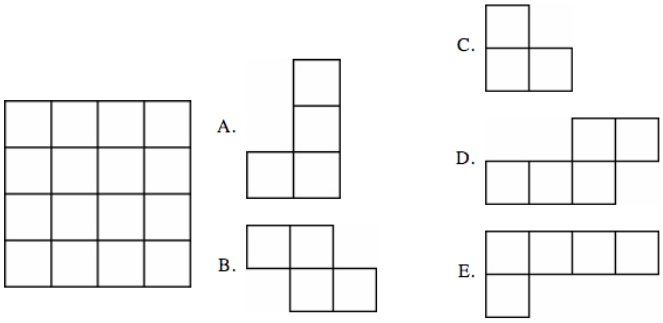


Figure 4: Grade 8 NAEP, 2009, Content Classification: Geometry. The problem says: "Identical puzzle pieces have been put together to form the large square shown [to the left]. Which of the following [shapes to the right] could be the shape of each puzzle piece?"

Overall, the prerequisites for the released NAEP math questions in all categories for 4th and 8th grade are minimal. Some questions test insignificant vocabulary only. Calculators are permitted on a substantial portion of the tests. The NMP Task Group on Assessment identified as "one of its greatest concerns" that "fractions (defined here as fractions, decimals, and related percent) are underrepresented on NAEP." Perhaps most importantly, many of the questions appear to be IQ items, rather than math problems, in the sense that their solutions rely on almost no education or knowledge of mathematical techniques. This is especially the case for those questions that require students to complete a pattern or to fill in geometric shapes with other geometric shapes, like puzzles.

Indeed, NAEP scores have been used by psychologists for the purpose of estimating IQ, state by state [6]. That may be a more plausible use for these tests than their stated purpose, to measure mathematics achievement. In fact, a congressionally mandated report, more recent than the Task Group Assessment concluded that "intended uses of NAEP assessment scores were not clearly defined." The report called for additional research into alignment between NAEP exams and state assessments based on academic content standards [7]

While it is true that some NAEP test questions do include rudimentary mathematical content, many others are as deficient as those displayed in Figs. 1 - 4. What then do marginal differences in NAEP math scores between states really measure? Do they measure relative effectiveness of states' mathematics education programs, as is usually assumed, or do they measure differences in average IQ (whatever that might mean) of the residents? If it is primarily the latter, one would expect NAEP scores to show little if any increase even if school math instruction improved significantly.

The NAEP exam is widely regarded as the yard stick of mathematics achieve-

ment at the 4th and 8th grade levels in the U.S. Diane Ravitch described one of its uses as follows [8],

> NAEP monitors trends; if the state says its scores are rising but its scores on NAEP are flat, then the state reports are very likely inflated. In a choice between the state's self-reported scores and an audit test, the public should trust the audit test.

In the case of some states, the public should trust neither. Regardless of flaws in state assessments, to the extent that the NAEP is an IQ test, it is measuring something different from what state tests are designed to assess. Achievement is not the same as ability.

Until such time as a reliable national mathematics achievement test comes into existence, the plethora of education research articles that base their findings on NAEP math scores should be considered with reservations. More reliable, for the time being, are state administered K-12 mathematics assessments directly tied to the content of credible state standards, as in the case of California.

Acknowledgement. The author thanks Elizabeth Carson, Tom Loveless, Richard Phelps, Hung-Hsi Wu, and Ze'ev Wurman for helpful comments and suggestions.

References

[1] Loveless, T. How well are American students learning? Brown Center Report on American education, Brookings Institution 2009.

[2] U.S. Department of Education, National Center for Education Statistics, NAEP Questions Tool: http://nces.ed.gov/nationsreportcard/itmrlsx/search.aspx?subject=mathematics. According to the National Center for Educational Statistics website, "After each assessment, NAEP releases dozens of sample questions to the public–more than 2,000 questions are currently available. The tools featured here can be used to supplement classroom instruction, provide additional insight into the content of the assessment, and show what students nationally or in your state or district know and can do." http://nces.ed.gov/nationsreportcard/itmrlsx/

[3] National Assessment Governing Board, U.S. Department of Education, Mathematics Framework for the 2009 National Assessment of Educational Progress.

[4] Foundations for success: The Final Report of the National Mathematics Advisory Panel, U.S. Department of Education (2008). http://ed.gov/about/bdscomm/list/mathpanel/reports.html

[5] Chapter 8: Report of the Task Group on Assessment, National Mathematics Advisory Panel, (2008). http://ed.gov/about/bdscomm/list/mathpanel/reports.html

[6] McDaniel, M., Estimating state IQ: Measurement challenges and preliminary correlates, *Intelligence* **34**, p. 607 - 619 (2006) doi:10.1016/j.intell.2006.08.007.

[7] U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Evaluation of the National Assessment of Educational Progress, Study Reports, Washington, D.C., 2009. http://www2.ed.gov/rschstat/eval/other/naep/naep-complete.pdf

[8] Ravitch, D., The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education, Basic Books (2010) p. 162.