

The 25% Solution: Improving Instruction by Improving the NCLB Tests

Dr. Steven L. Kramer
Math Science Partnership of Greater Philadelphia

What is the 25% solution? The 25% solution would be a pool of test items developed nationally, for each grade level tested under the NCLB. Each year, a set of 20-30 items for each grade level would be selected from the pool to form a Common Core of items for that year. Participating states would use the Common Core items to comprise roughly 25% of their test at each grade level.

For which academic subjects would a Common Core of items be developed? A separate pool of items would be developed for each subject tested under NCLB.

Why a Common Core of test items? The primary purpose of the 25% solution is to improve the quality of classroom instruction (Elevate). A secondary purpose is, to the degree possible, to put NCLB test scores from different states onto the same scale so that achievement across states can be more fairly compared (Equate).

How will the Common Core of test items elevate instruction? No single change to NCLB will have as positive effect on student learning as will creating better tests to drive better instruction. The phrase “teaching to the test” is often used in a pejorative manner. But “teaching to” tests with high stakes for administrators and teachers is inevitable—and can be a good thing, as long as the tests are good. For example, if the only way to pass the eighth grade math test is to ensure that students can really understand the concepts and fluently perform the operations we want eighth graders to know, then we want teachers to teach to the test.

The problem is that today high stakes NCLB tests are not designed to improve teaching:

1. Given limited resources and the relatively short time frame for complying with the original NCLB law, states could not dedicate sufficient time, expertise, and public discussion to developing coherent grade-level by grade-level sequences of standards on which to base high stakes NCLB tests.
2. There is no incentive to refine and improve standards. In fact, any state changing its standards must then spend scarce resources replacing pools of already developed test items.

3. Each major objective in a state's standards has many sub-skills that could be tested. Due to limited test development resources test questions often focus only on the subset of skills that are easiest to test reliably—generally procedural skills that don't require a flexible understanding of the most important underlying concepts we want children to learn.
4. The law has created a perverse incentive for states to “dumb down” NCLB tests so that all students can pass them. Not all tests have been “dumbed down”—but many have been.

By working nationally, we can develop a common core of test items that are far better than many of the items currently available. The nation as a whole can dedicate sufficient time, money, and expertise towards developing a coherent sequence of critical Big Ideas to be taught at each grade level, and to developing test items that really test the most central concepts to be taught. Further, the incentive for states to “dumb down” tests in order to meet federal percent passing goals will be irrelevant to standards and test items developed nationally.

How will the Common Core of items equate tests across states? The same statistical techniques used to develop growth models can be used to equate separate tests that are linked by 20 to 30 common core items, called “anchor items”. Such equating will only be valid for states that as part of their Standards implement the Big Ideas which the items measure.

How will the Common Core of items be developed? Step 1 is to develop a national consensus on a coherent sequence of the most important Big Ideas to be taught at each grade level. The sequence would be developed by the best minds—educators and content experts—in the country. A diversity of viewpoints would need to be represented. The effort would build on important work that has already been done by national organizations like the College Board and the National Council of Teachers of Mathematics. Step 2 is to allocate sufficient time for the public and states to react to the early drafts of the 25% Big Ideas, and for revision of the early drafts. Step 3 will be to develop the actual test questions. Measurement experts would work closely with educators and content experts to make sure test questions tapped the central concepts behind the Big Ideas. It will take extra effort and expense to develop items that measure the most important—not the easiest-to-test—ideas students should learn.

Why 25%? Implementing the Common Core as 25% of a state's test should be sufficient to drive instruction in the desired direction—especially since scores across states will be equated based on how well students do on the common core of items. Meanwhile, leaving 75% of the test to be locally developed will maintain the local control, flexibility and opportunity to innovate that are strengths of a federal system.

How can we reliably measure all the critical Big Ideas for a grade level with only 25% of a grade level's test? We can't and don't need to. To drive instruction, the pool of items from which each year's Common Core is selected must cover all the key concepts from the Big Ideas we want teachers to teach. The actual items selected don't need to cover everything—just so long as the only way to make sure students score well on the Common Core is to teach them all the most important ideas for that grade level.

Will states be required to use the Common Core items? No. Participation will be voluntary.
Will the Common Core of items be useful as a diagnostic test for individual students? No. It is not possible even for 100% of a yearly summative test to measure all the most important big ideas in sufficient depth to diagnose what each individual student needs. The goal is to drive instruction: to design a test so that a teacher can only be sure his/her students will score well if the students have a deep understanding of the most important subject matter to be taught at their grade level. While diagnostic and formative assessment are important tools to help teachers move students towards the learning goals the Common Core will test, teachers will need to implement such assessments during the year in ways tailored to their own classrooms.