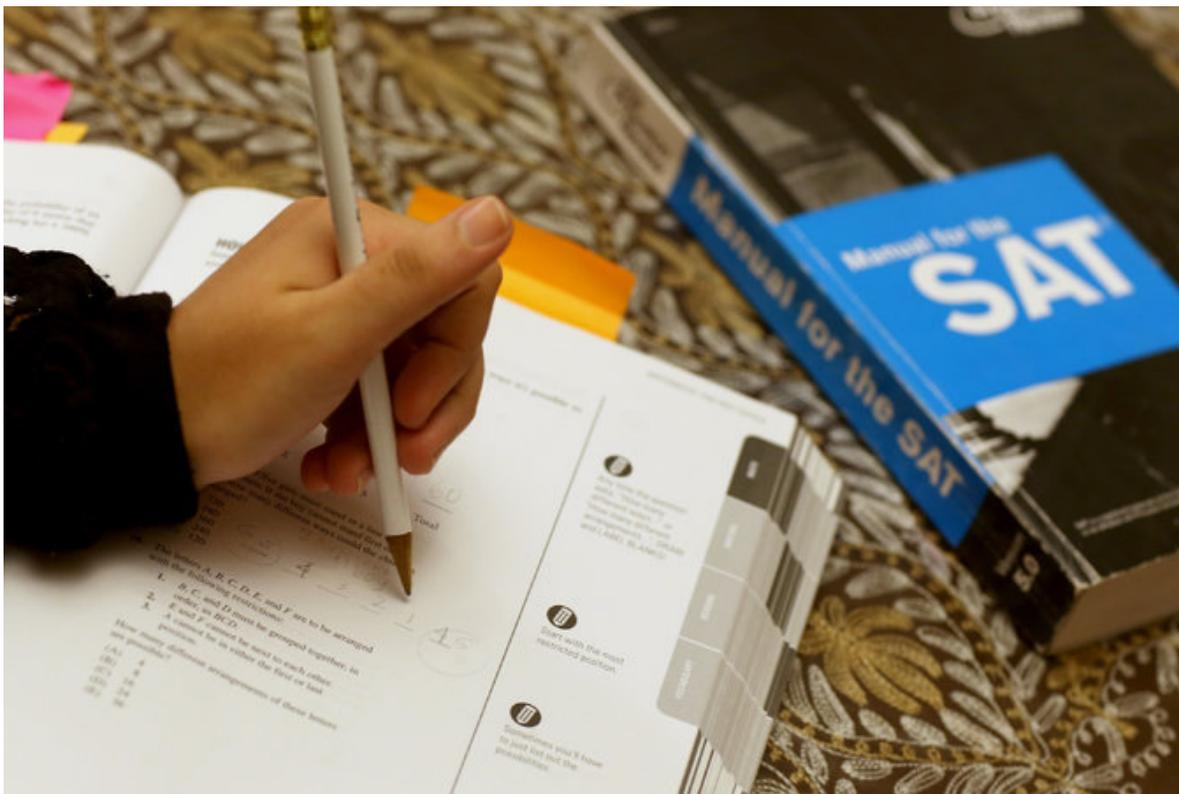


The Big Problem With the New SAT

By RICHARD C. ATKINSON and SAUL GEISER MAY 4, 2015



Credit Joe Raedle/Getty Images

AT first glance, the College Board’s revised SAT seems a radical departure from the test’s original focus on students’ general ability or aptitude. Set to debut a year from now, in the spring of 2016, the exam will require students to demonstrate in-depth knowledge of subjects they study in school.

The revised SAT takes some important, if partial, steps toward becoming a test of curriculum mastery. In place of the infamously tricky, puzzle-type items, the exam will be a more straightforward test of material that students encounter in the classroom. The essay, rather than rewarding sheer verbosity, will require students to provide evidence in support of their arguments and will be graded on both analysis and writing. Vocabulary will move away from the obscure language for which the SAT is noted, instead emphasizing words commonly used in college and the workplace.

While a clear improvement, the revised SAT remains problematic. It will still emphasize speed — quick recall and time management — over subject knowledge. Despite evidence that writing is the single most important skill for success in college, the essay will be optional. (Reading and math will still be required.)

And the biggest problem is this: While the content will be new, the underlying design will not change. The SAT will remain a “norm-referenced” exam, designed primarily to rank students rather than measure what they

actually know. Such exams compare students to other test takers, rather than measure their performance against a fixed standard. They are designed to produce a “bell curve” distribution among examinees, with most scoring in the middle and with sharply descending numbers at the top and bottom. Test designers accomplish this, among other ways, by using plausible-sounding “distractors” to make multiple-choice items more difficult, requiring students to respond to a large number of items in a short space of time, and by dropping questions that too many students can answer correctly.

“Criterion-referenced” tests, on the other hand, measure how much students know about a given subject. Performance is not assessed in relation to how others perform but in relation to fixed academic standards. Assuming they have mastered the material, it is possible for a large proportion, even a majority, of examinees to score well; this is not possible on a norm-referenced test.

K-12 schools increasingly employ criterion-referenced tests for this reason. That approach reflects the movement during the past two decades in all of the states — those that have adopted their own standards, as well as those that have adopted the Common Core — to set explicit learning standards and assess achievement against them.

Norm-referenced tests like the SAT and the ACT have contributed enormously to the “educational arms race” — the ferocious competition for admission at top colleges and universities. They do so by exaggerating the importance of small differences in test scores that have only marginal relevance for later success in college. Because of the way such tests are designed, answering even a few more questions correctly can substantially raise students’ scores and thereby their rankings. This creates great pressure on students and their parents to avail themselves of expensive test-prep services in search of any edge. It is also unfair to those who cannot afford such services. Yet research on college admissions has repeatedly confirmed that test scores, as compared to high school grades, are relatively weak predictors of how students actually perform in college.

By design, norm-referenced tests reproduce the same bell-curve distribution of scores from one year to the next, with only minor differences. This makes it difficult to gauge progress accurately.

Rather than impose higher education’s antiquated regime of norm-referenced tests on K-12 schools, American education would be better served if the kind of criterion-referenced tests now increasingly employed in K-12 schools flowed upward, to our colleges and universities.

And by rewarding students’ efforts in the regular classroom, criterion-referenced exams reduce the importance of test-prep services, thus helping to level the playing field. They signal to students and teachers that persistence and hard work, not just native intelligence or family income, can bring college within reach. They are better suited to reinforce the learning of a rigorous curriculum in our poorest schools.

College admissions will never be perfectly fair and rational; the disparities are too deep for that. Yet the process can be fairer and more rational if we rethink the purposes of college-entrance exams.

The revised SAT takes promising steps away from its provenance as a test of general ability or aptitude — a job it never did well — and toward a test of what students are expected to learn in school. But the College Board should abandon the design that holds it back from fulfilling that promise.

Richard C. Atkinson is [president emeritus](#) of the University of California. [Saul Geiser](#) is a research associate at the Center for Studies in Higher Education at the University of California, Berkeley.

A version of this op-ed appears in print on May 5, 2015, on page A23 of the New York edition with the headline: The Big Problem With the New SAT.